UNIVERSIDAD METROPOLITANA DE CIENCIAS DE LA EDUCACIÓN

FACULTAD DE CIENCIAS BÁSICAS

DEPARTAMENTO DE MATEMÁTICA

EL CONOCIMIENTO MOVILIZADO POR EL PROFESOR DE MATEMÁTICA PARA ENSEÑAR LA DISTRIBUCIÓN NORMAL A PARTIR DEL PROGRAMA DE ESTUDIO DE LA FORMACIÓN DIFERENCIADA DE PROBABILIDADES Y ESTADÍSTICAS DESCRIPTIVA E INFERENCIAL PROPUESTO POR EL MINEDUC

TESINA PARA OPTAR AL TÍTULO DE PEDAGOGÍA EN MATEMÁTICA MENCIÓN EDUCACIÓN EN TECNOLOGÍA Y AL TÍTULO DE PEDAGOGÍA EN MATEMÁTICA MENCIÓN ESTADÍSTICA EDUCACIONAL

AUTORES:

FRANCESCA XIMENA FUENTEALBA AUCAPÁN

JUAN CLAUDIO GALLARDO VALENZUELA

PROFESORA GUÍA:

GIOVANNA VINCENZA TICCHIONE TRONCOSO

SANTIAGO DE CHILE, MARZO DE 2022

UNIVERSIDAD METROPOLITANA DE CIENCIAS DE LA EDUCACIÓN

FACULTAD DE CIENCIAS BÁSICAS

DEPARTAMENTO DE MATEMÁTICA

EL CONOCIMIENTO MOVILIZADO POR EL PROFESOR DE MATEMÁTICA PARA ENSEÑAR LA DISTRIBUCIÓN NORMAL A PARTIR DEL PROGRAMA DE ESTUDIO DE LA FORMACIÓN DIFERENCIADA DE PROBABILIDADES Y ESTADÍSTICAS DESCRIPTIVA E INFERENCIAL PROPUESTO POR EL MINEDUC

TESINA PARA OPTAR AL TÍTULO DE PEDAGOGÍA EN MATEMÁTICA MENCIÓN EDUCACIÓN EN TECNOLOGÍA Y AL TÍTULO DE PEDAGOGÍA EN MATEMÁTICA MENCIÓN ESTADÍSTICA EDUCACIONAL

AUTORES:

FRANCESCA XIMENA FUENTEALBA AUCAPÁN

JUAN CLAUDIO GALLARDO VALENZUELA

PROFESORA GUÍA:

GIOVANNA VINCENZA TICCHIONE TRONCOSO

SANTIAGO DE CHILE, MARZO DE 2022

2022, Francesca Fuentealba Aucapán y Juan Gallardo Valenzuela

Autorizan la reproducción total o parcial de este material, con fines académicos, por cualquier medio o procedimiento, siempre que se haga la referencia bibliográfica que acredite el presente trabajo y sus autores.

Importante

En la presente investigación se utilizan de manera inclusiva términos como “el docente”, “el profesor”, “el estudiante” y sus respectivos plurales junto a sinónimos equivalentes en el contexto educativo para referirse a hombres, mujeres u otros géneros. Se recurre a ello para evitar la saturación gráfica y dificultad de lectura que produce el uso de nomenclaturas como “o/a”, “los/las” y otras.

Dedico con afecto esta tesina:

A Iris, mi hija, mi más sincero amor.

A Sonia, mi mamita, mujer fuerte y resiliente.

A todos los profesores hijos de la Educación Pública.

Francesca Fuentealba Aucapán

Dedico esta tesina:

A mi hermana, quien alegra mis días.

A mis papis, guías de mi camino.

A mi mami y tata.

Con mucho cariño.

Juan Gallardo Valenzuela

Agradecimientos

Al culminar esta etapa tengo infinitos agradecimientos a mi familia, ellos fueron mi apoyo y motivación constante en este largo proceso y siempre serán un pilar fundamental en mi vida. Doy gracias a mis papis, Sonia y Robinson, por entregarme sus valores, por creer y nunca dejarme caer, ¡los amo mis viejitos! Agradezco a mis hermanas y hermano; a Gaby por ser la mejor tía y en ocasiones la segunda mamá para mi monita, a Claudita por cada masajito reponedor y al Ale por todos los cafés que me preparó. Agradezco a mi hija, Iris, por brindarme siempre su comprensión, cariño, amor y ser mi mayor fuente de inspiración para superarme y ser cada día mejor.

Agradezco a todas las personas, compañeros y amigos que estuvieron presentes brindando su ayuda desinteresadamente siempre con palabras de apoyo, logrando que este sueño se haga realidad, en especial a, Maca, Ronny y Javiera.

Imposible no agradecer a Nobuo Uematsu, compositor del OST de FFVII quien me acompañó durante las largas noches de estudio y desarrollo de esta Tesina.

A mi compañero, Juanin, por permitirme terminar este proceso junto a él, por confiar en mí, motivarme y regalarme tantas risas. ¡Muchas gracias, amigo y coleguita!

Quiero finalizar citando John Cotton Dana, cuya frase recordaré por siempre, "quién se atreva a enseñar nunca debe dejar de aprender".

Francesca Fuentealba Aucapán

En primer lugar, agradecer a mi familia, por el apoyo incondicional que siempre me otorgan. Mamá, creo no equivocarme al decir que eres la mejor del mundo, gracias por la confianza que siempre has tenido en mí. Papá, por siempre estar ahí y creer en mí. Hermana, por ayudarme muchísimas veces y dar ese ánimo cuando lo necesitaba. Ustedes siempre me apoyaron en todo lo que quería lograr, acompañaron a cumplir algunos de mis sueños y especialmente nunca me juzgaron por la carrera que elegí estudiar. Sin duda los tres son el pilar fundamental de mi vida y un motivo para terminar la tesina. Gracias por todo eso, y por más. Los quiero mucho.

También agradezco mi grupo de amigos, en especial a la Cata, eres una amiga increíble, gracias por compartir tu amistad conmigo. Aún recuerdo ese primer día de clases cuando me acerqué a conversar con ustedes y me recibieron tan amables y con tanto cariño, junto a Gustavo y Antony fueron un apoyo primordial dentro de la carrera, doy gracias por esos fantásticos momentos que hemos compartido y hacer este paso por la educación superior más ameno. Pauli, tampoco podía olvidarme de ti, eres una persona brillante, agradezco por acompañarme en esa difícil decisión de cambiarnos de universidad, por ser como eres y mantener esta bonita amistad por todo este tiempo, y ojalá mucho más.

Agradezco infinitamente a mi compañera de tesis, Fran, por querer compartir conmigo este proceso tan importante, por tu paciencia en cada momento y haberme ayudado tantas veces. Creo que aprendí a conocerte un poquito más y descubrí que eres una persona excelente. Gracias por todas esas risas, momentos inolvidables en las reuniones durante toda esta difícil y estresante instancia.

A la profe Giova, aprendí demasiado con usted y le debo mi fascinación por la estadística. Gracias por confiar en mis capacidades y creer en mí.

Finalmente, a todos esos compañeros y compañeras que estuvieron presentes en mi transcurso universitario.

Gracias totales, a todos ustedes.

Juan Gallardo Valenzuela

Agradecemos por su ayuda durante estos meses a la profesora Giovanna Ticchione, quien aceptó guiar esta Tesina.

Francesca y Juan

Tabla de Contenido

Lista de Tablas

Lista de Figuras

Resumen

La formación diferenciada Humanista-Científico incorporó oficialmente en el 2021 el electivo de Probabilidades y Estadísticas Descriptiva e Inferencial para los niveles de III° y IV° medio, en él se amplía y profundiza el tratamiento de la Distribución Normal, objeto de este estudio. Esta reciente implementación curricular trajo consigo un escenario complejo para los docentes, comprendiendo que las actividades propuestas en el Programa de Estudio es el único material pedagógico formal a disposición de la comunidad educativa, junto a los desafíos y dificultades propias de la enseñanza de la estadística y probabilidad. Atendiendo esta problemática, el objetivo de investigación corresponde a analizar el conocimiento matemático para la enseñanza de la Distribución Normal, a partir de una revisión crítica a las actividades del Programa de Estudio de la asignatura. Para ello, se utilizó una metodología cualitativa bajo el modelo de investigación-acción. Contrastando las actividades con los Indicadores de Evaluación, los resultados indicaron que es una sólida propuesta curricular, no obstante, algunas actividades requieren cambiar y/o potenciar el tratamiento del contenido asociado a la Distribución Normal. En base a ello se elabora una propuesta pedagógica que contiene subactividades y preguntas complementarias a la propuesta ministerial, un apartado del saber sabio sobre el objeto matemático y la resolución de todos los ejercicios y problemas de las actividades. De esta forma se proporciona un documento que es un aporte a la labor docente.

Palabras Claves: Distribución Normal, Programa de Estudios de Probabilidades y Estadísticas Descriptiva e Inferencial, MTSK.

Introducción

En Chile esta última década la educación ha estado sometida a múltiples cambios curriculares conforme las diferentes reformas educativas que han legislado la educación del país. Entre ellas, el año 2021 se implementa oficialmente los cursos del Plan diferenciado para la educación media Científica-Humanista para III° y IV° medio, donde la formulación de los Objetivos de Aprendizaje se relaciona con habilidades, conocimientos y actitudes, conformando así un currículum centrado en el aprendizaje.

Producto de lo anterior nace la asignatura de Probabilidades y Estadísticas Descriptiva e Inferencial, donde los estudiantes podrán abordar problemas propios de la disciplina, generar propuestas relacionadas con el entorno y familiarizarse con el uso de herramientas digitales especialmente diseñadas para la estadística y las probabilidades. Así, el curso inicia con procedimientos de estadística descriptiva, haciendo énfasis en la interpretación de diversas representaciones de conjuntos de datos y la comparación de las características de muestras y poblaciones. Transita luego desde situaciones modeladas mediante variables aleatorias discretas hacia las que requieren variables aleatorias continuas. Se amplía y profundiza el tratamiento de las distribuciones Binomial y Normal, por sobre lo propuesto en la Formación General, al usar distribuciones como modelos de situaciones o fenómenos del contexto cotidiano, científico y social. Se da cierre al curso con una introducción a los métodos de la estadística inferencial, el uso de intervalos de confianza y la prueba de hipótesis (Ministerio de Educación [MINEDUC], 2021b).

Para llevar a cabo los propósitos planteados en esta nueva asignatura, el MINEDUC (2021b) pone a disposición de la comunidad educativa su Programa de Estudio, en el cual se presenta una serie de actividades como apoyo para efectuar la enseñanza de los contenidos matemáticos bajo los lineamientos definidos por las Bases Curriculares. Producto de la reciente implementación del curso del Plan Diferenciado, se advierte la importancia de este documento, pues actualmente es el único material pedagógico oficial a disposición de los docentes.

Por tanto, el presente trabajo de investigación tiene por objetivo analizar el conocimiento matemático que debe tener el profesor para la enseñanza de la Distribución

Normal, a partir de una revisión crítica al Programa de Estudio de Probabilidades y Estadísticas Descriptiva e Inferencial.

En el primer capítulo se exponen los antecedentes para contextualizar y evidenciar la problemática de esta investigación. En particular, se detalla el trato de la Distribución Normal en las nuevas Bases Curriculares del Plan Diferenciado, se describe las dificultades, falencias y desafíos de los docentes en la enseñanza de la estadística escolarizada y se indican las consecuencias de la operatividad en la enseñanza del objeto matemático en estudio. Por último, se plantea el objetivo general y a partir de él se proponen cuatro objetivos específicos.

En el segundo capítulo se pormenoriza los sustentos teóricos necesarios para realizar esta investigación. Así el marco teórico circula por el desarrollo histórico de la Distribución Normal, el conocimiento profesional del profesor de estadística, de probabilidad y del conocimiento especializado del profesor de matemática (MTSK).

El tercer capítulo presenta el marco metodológico, basado en un enfoque cualitativo, bajo el modelo de investigación-acción, explicando las fases de su ciclo y las tareas realizadas para llevar a cabo cada una ellas, describiendo, además, el contexto bajo el cual esta investigación se efectuó. Para ello se elaboró un instrumento que permitió establecer tanto los parámetros de pertinencia de las actividades de la propuesta ministerial para la enseñanza de la Distribución Normal, como sugerencias y observaciones para resolver las actividades del documento. Además, se diseñaron subactividades y preguntas complementarias a estas.

En el cuarto capítulo se describen los resultados más relevantes obtenidos al aplicar el instrumento pertinente para la revisión del Programa de Estudio, del mismo modo en la validación de subactividades y preguntas complementarias a juicio de expertos se declaran los comentarios más significativos. Por último, producto del ciclo de investigación-acción se exhibe la propuesta de subactividades y preguntas elaborada por los investigadores.

Para finalizar, en el quinto capítulo se plantean las conclusiones obtenidas en función del objetivo general y de cada objetivo específico propuesto, señalando también, algunas proyecciones respecto a esta investigación.

Capítulo I. Antecedentes y Problemática

1.1. Antecedentes

A continuación, se recogen hechos relevantes que permitirán relacionar elementos de la enseñanza de la probabilidad y estadística escolar junto a la preparación docente en esta área de conocimiento, poniendo énfasis en el aprendizaje para la enseñanza de la Distribución Normal.

1.1.1. Contexto nacional: nuevas Bases Curriculares.

El año 2009 se promulga la Ley N° 20.370 General de Educación, estableciendo el marco para una nueva institucionalidad de la educación escolar chilena, decretando principios, obligaciones y promoviendo cambios en la forma en que los niños y jóvenes del país serán educados. En consecuencia, esta última década se han reestructurado progresivamente las Bases Curriculares para los distintos niveles y modalidades educativas. En efecto, la primera modificación se llevó a cabo el año 2012 para los niveles de primero a sexto básico, la segunda, el año 2015 para los niveles de séptimo básico a segundo año medio, en tercer lugar, el año 2019 se realizó en los niveles de tercero y cuarto año medio (MINEDUC, 2019), finalmente, en Julio del 2021 se publicó el decreto en trámite correspondiente a los cambios de las bases curriculares dirigida a la educación para jóvenes y adultos (MINEDUC, 2021a).

Las Bases Curriculares de III° y IV° año de enseñanza media, por medio del Decreto Supremo Nº 193 de 2019, involucró tanto a la reforma curricular del plan común de formación general, como a la del plan diferenciado para la educación Técnico Profesional y Artístico, y Humanístico-Científico. Haciendo hincapié en la nueva formación diferenciada de este último, se estipula que debe ofrecer a los estudiantes diversas asignaturas de profundización vinculadas a las disciplinas de formación general, permitiéndoles acceder a un abanico de opciones que tienen correspondencia con sus intereses y preferencias personales (MINEDUC, 2019).

Dentro de las asignaturas de profundización del área de Matemática, se encuentra el electivo Probabilidades y Estadísticas Descriptiva e Inferencial, donde los estudiantes podrán utilizar las probabilidades y estadística para “analizar diversas situaciones o fenómenos sociales y científicos; extraer conclusiones; tomar decisiones con base en datos cuantitativos; comunicar y argumentar resultados, y validar conclusiones o hallazgos acerca de muestras y poblaciones" (MINEDUC, 2021b, p. 21), de forma que los conocimientos y habilidades adquiridas con este curso sean provechosos para los estudiantes, tanto en su desempeño en la vida diaria como para el estudio de alguna carrera de nivel superior.

El Programa de Estudio del electivo establece solo un Objetivo de Aprendizaje asociado a la Distribución Normal, correspondiente al OA 3, el cual estipula "modelar fenómenos o situaciones cotidianas del ámbito científico y del ámbito social que requieran el cálculo de probabilidades y la aplicación de las Distribuciones Binomial y Normal" (MINEDUC, 2021b, p. 27). Se estima que el tiempo de trabajo de este OA corresponde a diez semanas, fraccionado en seis horas semanales. De igual manera, el documento ministerial señala que en este curso los estudiantes podrán desarrollar y/o potenciar tanto habilidades digitales y tecnológicas como habilidades matemáticas que consisten en argumentar y comunicar, resolver problemas, modelar y representar MINEDUC (2021b).

1.1.2. Preparación docente en la enseñanza de la estadística escolarizada.

El eje temático de Datos y Azar en el currículum chileno se remonta al año 2009, en un principio se esperaba que los estudiantes pudieran "leer, comprender, tomar decisiones y ser críticos con la gran cantidad de información con la que se interactúa diariamente" (Rodríguez, 2021a), para ello, Rodríguez (2021b) indica que los docentes deben tener manejo en tres áreas de la estadística. En primer lugar, la alfabetización estadística, que alude a la comprensión y resultado del aprendizaje de los cursos de Estadística, seguido del pensamiento estadístico, referido a comprender los motivos y formas de hacer investigaciones, las cuales involucran la estadística para el procesamiento de información, por último, el razonamiento estadístico, que corresponde a darle sentido a este quehacer estadístico para su interpretación y explicación.

En ese sentido, se hace sumamente importante que los docentes manejen a cabalidad los contenidos a tratar en el sistema educacional, en especial las tres áreas ya mencionadas de la estadística, por ello, deben tener presente los Estándares Pedagógicos y Disciplinarios, ya que estos contribuyen a su formación universitaria además de considerarse como un instrumento de referencia respecto a los conocimientos esperados que deben adquirir los profesores.

Así, el descriptor 8 del estándar C de Probabilidades y Estadística estipulado en los Estándares Pedagógicos para Carreras de Pedagogía en Matemática, declara que el docente:

> define variables aleatorias y las utiliza para modelar fenómenos aleatorios, describiendo el comportamiento de la variable a través de funciones de probabilidad o densidad, como la Binomial y la Normal, y evalúa la pertinencia del modelo en situaciones de incertidumbre de índole social, cultural o científica. (Centro de Perfeccionamiento, Experimentación e Investigaciones Pedagógicas [CPEIP] y MINEDUC, 2021, p. 88)

Siendo el estándar 11 el que generaliza la importancia de los diez anteriores, señalando que el profesor "comprende el potencial de las Probabilidades y Estadística como herramienta para estudiar fenómenos naturales y sociales, y como ejemplo de la estrecha relación entre la matemática y las otras áreas del conocimiento" (p. 88). En suma, se espera que el docente tenga una formación sólida en esta rama de la matemática. En específico, él debe conocer y aplicar los conceptos básicos de variables aleatorias, funciones de probabilidad y densidad, significado de sus parámetros, esperanza, varianza, percentiles y a través de estas conducir al desarrollo de la habilidad de resolución de problemas, es decir, el profesor en el aula debe abordar la comprensión y elaboración de estrategias que tengan relación con los contenidos estadísticos y probabilísticos, además de diseñar actividades que desciendan el objeto matemático en situaciones realistas de la vida cotidiana de los estudiantes.

En consecuencia, se puede aludir que la formación de los docentes en lo que respecta a probabilidad y estadística, debe ser sólida para llevar a cabo una enseñanza adecuada y aprendizaje propicio a las aulas del territorio nacional. No obstante, lo que estipula la teoría está alejado de la realidad que sucede en el país, ya que al analizar los resultados entregados por Rodríguez, Vásquez y Rojas (2020) sobre el rendimiento en la Evaluación Nacional Diagnóstica (END) del año 2017 de 612 estudiantes de Pedagogía en Matemática pertenecientes a 37 Instituciones de Educación Superior (IES), es posible visualizar que lo que exponen los Estándares Pedagógicos no se cumple a cabalidad. Antes de comenzar a enunciar los resultados de dicho documento, se debe aclarar que la END, es una prueba de carácter censal y obligatorio para cualquier estudiante de pedagogía, la cual tiene como finalidad "diagnosticar la formación de los nuevos docentes, generando información valiosa para los programas académicos" (CPEIP, 2018, p. 05) y es realizada por el CPEIP. Esta evaluación se aplica mediante dos pruebas, la primera correspondiente a los conocimientos pedagógicos generales (PCPG) y la otra prueba es de conocimientos disciplinarios y didácticos (PCDD) (CPEIP, 2018). La evaluación PCDD que rindieron en 2017 los estudiantes de Pedagogía en Matemática corresponde a 60 reactivos cerrados de selección múltiple, de los cuales 13 corresponden al eje de Datos y Azar (Rodríguez et al., 2020).

Según Rodríguez et al. (2020), respecto al análisis de los resultados de la END del 2017 indican que los docentes no poseen la preparación suficiente en el eje temático de Datos y Azar, lo que radica en una "falta de perspectiva pedagógica o didáctica" (p. 150), frente a la enseñanza matemática escolar que realizan en el aula. Esto es atribuible, probablemente a lo abstracto de los contenidos, sin embargo, es fundamental que esto se revierta, para que el docente pueda dirigir el aprendizaje hacía el desarrollo de las habilidades correspondientes a cada eje. Asimismo, hay que señalar que en "las preguntas relacionadas con (…) Datos y azar, la distribución de los puntajes presenta una asimetría positiva", probándose que los docentes tienen problemas en su formación inicial, lo que explica el sesgo presente en los puntajes obtenidos, dejando en evidencia que en lo que respecta a conocimientos disciplinares se está al debe, sin importar la institución de la cual provenga (Rodríguez et al., 2020). Estos resultados obtenidos por los autores se potencian aún más con los estudios nacionales e internacionales, que concluyen que los docentes no terminan de adquirir los conocimientos y habilidades necesarias para la enseñanza de este eje temático, traduciéndose en una mala base en su formación (Rodríguez, 2021b).

Trasladando el enfoque a la implementación y preparación del curso de formación diferenciada, Arenas y Arriagada (2020), en su encuesta titulada Percepción del nuevo electivo de Probabilidades y Estadísticas Descriptiva e Inferencial para III° o IV° medio, destacan que aproximadamente un 57% de los docentes se sienten preparados para dar el electivo. Respecto a la Distribución Normal, se señala que más del 80% de los encuestados, vieron en su formación universitaria docente de pregrado en algún momento contenidos asociados a esta distribución. Además, la Normal, y aproximación de Distribución Binomial a la Normal son el tercer gran contenido del electivo donde los docentes no se sienten preparados para enseñar (quedando por debajo de los intervalos de confianza y test de hipótesis). A pesar de lo anterior, la mayoría de los docentes encuestados están de acuerdo en que se debe enseñar la Distribución Normal y su aproximación.

Además, tres cuartas partes del total de encuestados mencionan que usarán el material que entrega el MINEDUC para el electivo de Probabilidades y Estadísticas Descriptiva e Inferencial, en otras palabras, planificarán y guiarán el electivo a partir del programa del curso que contiene actividades de apoyo a los docentes. Finalmente, uno de los encuestados considera que el electivo "entrega una base mucho más sólida a los y las estudiantes para afrontar la educación superior, ya que en la mayoría de las carreras aparece algún ramo de estadística" (Arenas y Arriagada, 2020, p. 57), por lo que el electivo en la formación escolar de los estudiantes puede ser de gran utilidad para su continuidad de estudios en carreras universitarias del área científica.

1.1.3. Distribución Normal y su enseñanza.

Tauber (2001) indica que la Distribución Normal posee un significado complejo, ya que requiere del apoyo de conceptos previos, como, por ejemplo, los de variables estadísticas y aleatorias, distribución estadística, medidas de posición central y dispersión, simetría, probabilidad, entre otros. A su vez, menciona que el estudiante debe tener estos conceptos afianzados, de forma contraria se obstaculiza su aprendizaje, ahora bien, este escenario ideal es poco probable en la enseñanza de la estadística.

Según González, Ojeda, y Reyes (2018) la principal falencia que pueden presentar los docentes al tratar la Distribución Normal, radica en enfocar su enseñanza en la operatividad de los ejercicios y problemas, es decir, utilizar un algoritmo que consiste en identificar la variable aleatoria, definir lo que se pide, estandarizar y entregar las respuesta en porcentaje, de esta forma el estudiante "sólo repite lo que su profesor le enseña, sea correcto o incorrecto" (González et al., 2018, p. 1771), afectando su comprensión. Adicionalmente, los autores señalan que muchas veces el problema inicia por la interpretación del enunciado o identificación del espacio muestral, seguido de la falta de reflexión en la estandarización de la variable aleatoria y de la corrección de continuidad, error de diferenciación entre una variable aleatoria discreta y continua, relación entre puntos de inflexión de la campana de Gauss y desviación estándar, incluso posibles dificultades en la lectura de las tablas (González et al., 2018).

1.2. Problemática

En base a los antecedentes expuestos anteriormente, los docentes se encuentran en un escenario complejo, comprendiendo los desafíos y dificultades propias de la enseñanza de la estadística y probabilidad, en particular de la Distribución Normal, acompañado de la reciente implementación de las nuevas Bases Curriculares para III° y IV° medio, en específico, del plan diferenciado Humanístico-Científico del área de matemática, donde se advierte la importancia del Programa de Estudio de Probabilidades y Estadísticas Descriptiva e Inferencial, pues actualmente es el único material pedagógico que el MINEDUC ha puesto a disposición de los docentes.

Para la enseñanza de la Distribución Normal, el Programa de Estudio sugiere tres actividades formativas de aprendizaje, y una actividad evaluativa en correspondencia a los indicadores de evaluación en función de los Objetivos de Aprendizajes esperados. Por lo tanto, contemplando el uso que los docentes pueden darle a este documento ministerial, dichas actividades requieren de un análisis crítico respecto a su elaboración que considere el conocimiento especializado del profesor para la enseñanza de la matemática.

Por estas razones, se hace imprescindible elaborar una propuesta didáctica que incluya: (1) el saber sabio de la transposición didáctica acorde al conocimiento matemático necesario que debe poseer el docente para comprender y tratar el contenido asociado al objeto matemático presente en el documento ministerial. (2) La resolución de las actividades propuestas en el Programa de Estudios para cubrir elementos probabilísticos y/o didácticos faltantes, orientaciones docentes y sugerir subactividades en el caso que sea necesario, de esta manera se pretende contribuir tanto a la labor docente, a la gestión curricular y pedagógica, y en consecuencia a la comprensión profunda de la Distribución Normal por parte de los estudiantes.

Finalmente, en vista de los antecedentes y problemática expuesta se plantea la siguiente pregunta de investigación

1.3. Pregunta de Investigación

¿Qué conocimientos requiere movilizar el profesor de matemática para la enseñanza de la Distribución Normal utilizando el Programa de Estudio de la formación diferenciada de Probabilidades y Estadísticas Descriptiva e Inferencial?

1.4. Objetivos

En base a la problemática y la pregunta de investigación surge el Objetivo General y con la finalidad de organizar su estudio se plantean los Objetivos Específicos.

1.4.1. Objetivo general

- Analizar el conocimiento matemático para la enseñanza de la Distribución Normal, a partir de una revisión crítica al Programa de Estudio de Probabilidades y Estadísticas Descriptiva e Inferencial.

1.4.2. Objetivos específicos.

- Caracterizar la Distribución Normal acorde a los dominios del Conocimiento Especializado del Profesor de Matemáticas (MTSK).

Este objetivo consiste en describir y contextualizar los conocimientos del profesor de matemática para un contenido escolar en particular, específicamente para la Distribución Normal, de esta manera cada dominio y subdominio del MTSK es ejemplificado acorde al objeto matemático central de esta investigación. Producto de esta tarea se elabora material en correspondencia al saber sabio – mínimo – que debe considerar el docente para la enseñanza del contenido que trata la Distribución Normal.

- Revisar críticamente la propuesta ministerial sobre el contenido de la Distribución Normal en el Programa de Estudio del plan diferenciado de Probabilidades y Estadísticas Descriptiva e Inferencial.

Por medio de la revisión crítica este objetivo pretende identificar aciertos y desaciertos de las actividades propuestas en el Programa de Estudio del electivo para la enseñanza de la Distribución Normal.

- Describir los indicadores de evaluación que se presentan en menor proporción en el Programa de Estudio de Probabilidades y Estadísticas Descriptiva e Inferencial asociados al contenido de la enseñanza de la Distribución Normal.

A través de un instrumento se identifica si un indicador de evaluación no está plenamente cubierto por las actividades sugeridas en el Programa de Estudio, se debe señalar cuales son las actividades que requieren cambiar y/o mejorar el tratamiento de su contenido respecto a la Distribución Normal, esto con la finalidad de poder guiar las actividades sugeridas y orientaciones docentes, para que haya una cobertura plena del conocimiento probabilístico.

- Resolver las actividades propuestas para la enseñanza de la Distribución Normal en el Programa de Estudio de Probabilidades y Estadísticas Descriptiva e Inferencial, complementando aquellas que requieran de más detalle.

La finalidad de este objetivo consiste principalmente en proporcionar una propuesta pedagógica de las actividades referidas a la Distribución Normal en el Programa de Estudio. Con ella los docentes podrán recordar, complementar y/o mejorar sus propias bases respecto al contenido tratado, de esta forma sustentar sus clases al momento de impartir el electivo, y poder verificar si los procedimientos y resultados obtenidos por ellos y/o los estudiantes son los correctos. Además, se agregan algunas actividades que permitan al estudiante argumentar, visualizar o deducir elementos básicos de la Distribución Normal.

Capítulo II. Marco Teórico

2.1. Distribución Normal

La Distribución Normal, también conocida como Distribución de Gauss o Gaussiana es una de las más importantes dentro de la estadística y probabilidades por sus múltiples aplicaciones en diversos campos de estudio. En efecto, gran parte de los fenómenos biológicos, psicológicos o sociales, se ajustan a unas pocas leyes o distribuciones de probabilidad teóricas, siendo la Distribución Normal una de las tres fundamentales junto a la distribución Binomial, y de Poisson (Dagnino, 2014).

Antes de presentar su definición como se conoce hoy en día, se expondrá un recorrido histórico de los sucesos que llevaron hasta su formalización.

2.1.1. Desarrollo histórico de la Distribución Normal.

El descubrimiento de esta distribución se le adjudica al científico francés Abraham De Moivre (1667 – 1754), alrededor del año 1733 (Anderson, 2019). No obstante, para poder hablar de la historia de la Distribución Normal, se deben distinguir dos caminos seguidos por los científicos y matemáticos de la antigüedad. El primero tiene relación con la búsqueda de una distribución para los errores en las mediciones, principalmente astronómicas, donde se encuentran personajes como: Galileo Galilei, Thomas Simpson y Daniel Bernoulli. El otro camino corresponde a la aproximación del cálculo de probabilidades, donde hubo aportaciones de Luca Pacioli, *Tartaglia*, Girolamo Cardano, Blaise Pascal, Pierre de Fermat, Christiaan Huygens y Jacob Bernoulli (Conde, 2015).

Con respecto al primer camino mencionado, su primera visualización se remonta cerca del 1632 con Galileo Galilei (1564 – 1642) donde apreció que los instrumentos de mediciones astronómicas utilizados en la época tenían un margen de error, dándose cuenta que los "errores pequeños ocurren de manera más frecuente que errores grandes" (Másmela y Serrato, 2007, p. 2), de manera que al graficar estos, la curva formada era simétrica y la mayoría de los errores se concentraban en medio de la gráfica (Másmela y Serrato, 2007). Así, en el siglo XVII se puede hablar de una aproximación bastante abstracta a la Distribución Normal. Posterior a los breves estudios realizados por Galilei, aparece Thomas Simpson (1710 – 1761), quien propone las primeras curvas de error, asociadas a un conjunto de datos, estas funciones graficadas por Simpson tenían formas lineales. En el texto de Conde (2015) se puede evidenciar que una de las funciones asociadas es lo que hoy se conoce como función constante y la otra corresponde a dos funciones lineales (una creciente en el IV cuadrante del plano cartesiano y la otra decreciente en el I cuadrante), que forman un triángulo en el plano XY. Por otro lado, Daniel Bernoulli (1700-1782) también propone curvas de error, sin embargo, en su caso, las propuestas tenían formas circulares.

Por otra parte, el camino de la aproximación del cálculo de probabilidades se inició con el problema de los puntos, del que se tiene registro desde fines del siglo XIV, sin un origen claro, presuntamente árabe. Este problema estipula que:

> A y B participan en un juego de lanzamiento de pelota y acuerdan que el juego terminará hasta que alguno de ellos gane 6 rondas, pero el juego se suspende abruptamente cuando A ha ganado 5 rondas a su favor y B sólo tiene 2 rondas a su favor. ¿Cómo debería distribuirse la apuesta? (Conde, 2015, p. 60).

Matemáticos como Luca Pacioli (c. 1445 – 1517), Niccolò Fontana, más conocido como *Tartaglia* (c. 1499 – 1557) y Girolamo Cardano (1501 – 1576) trataron de darle solución al problema de los puntos, en los años 1494, 1556 y 1565 respectivamente (Conde, 2015), no obstante, sus respuestas fueron consideradas incorrectas, pues los procedimientos entregados no contemplaban las probabilidades que tienen de ganar, sino que se centran en las puntuaciones obtenidas por cada participante. Las confusiones y errores de este problema radican en que los matemáticos de la época los consideraban como un problema de aritmética, y no del cálculo de probabilidades (Vega, 2002).

Posterior a estos aportes, aparecen nombres como Blaise Pascal (1623 – 1662) y Pierre de Fermat (1607 – 1665), junto a Christiaan Huygens (1629 – 1695), que utilizaron métodos más novedosos para la época, como los combinatorios o recursivos ante el problema de los puntos, descubierto casi dos siglos antes (Conde, 2015), estos matemáticos son los primeros en percatarse en la relación que existía entre los juegos de azar y el problema (Vega, 2002). Por último, el libro póstumo publicado en 1713 de Jacob Bernoulli (1654 – 1705), llamado *Ars Conjectandi*, afirmaba que eran más importante las probabilidades de ganar de cada jugador. Asimismo, a partir de este problema, Bernoulli se aproximó a la Distribución Binomial, considerando que los dos jugadores no tuvieran la misma probabilidad para ganar (Conde, 2015).

Sin embargo, a partir de todas las contribuciones entregadas por los matemáticos antes del siglo XVIII sobre probabilidades, aún faltaban los que hicieron Abraham De Moivre, Pierre Simon – Laplace y Carl F. Gauss. De Moivre se dio cuenta que al tomar una Distribución Binomial con parámetro $p = 0{,}5$ y n una constante que aumentaba cada vez más, las barras de la gráfica de la función de probabilidad de la Binomial, seguía un patrón, y mientras más repeticiones del fenómeno Bernoulli se realizaban, la curva graficada se hacía más acampanada, además, la campana que se formaba era simétrica. A pesar de aquello, los aportes de De Moivre eran muy adelantados para la época, por lo que no fueron tomados en cuenta, aunque todo el trabajo que realizó este matemático francés dentro de la estadística, específicamente con la Distribución Normal, marca un hito en esta línea del tiempo, ya que es considerada como un traspaso entre dos eras desarrolladas anteriormente (Conde, 2015). También se destaca la investigación de Abraham De Moivre, porque dio origen al teorema conocido como De Moivre – Laplace, que corresponde a la aproximación de la Distribución Binomial a la Normal, la cual es utilizada dentro de diversas áreas. Cabe mencionar, que el desarrollo de toda su investigación lo adjuntó en la segunda edición de su libro *The Doctrine of Chance* publicado en 1738, donde ya se introducía el concepto de Distribución Normal (Anderson, 2019).

Por su parte, Pierre – Simon Laplace (1749 – 1827) siguió con el estudio de los errores de las mediciones y con el trabajo realizado por De Moivre de aproximar la Distribución Binomial por una Normal, obteniendo una casi nula aceptación de la comunidad matemática de la época, debido a la poca rigurosidad matemática. Es preciso señalar que, la importancia de todas las investigaciones sobre el primer teorema límite de la historia, fueron tomadas en cuenta recién en el año 1901, cuando Aleksandr Lyapunov formalizó y definió el funcionamiento del Teorema Central del Límite (Anderson, 2019). En lo que respecta a Carl F. Gauss (1777 – 1855) utilizó el método de mínimos cuadrados para ajustar la órbita que seguía el planeta enano Ceres (Conde, 2015), precisando en 1809 que la distribución de los errores en su medición correspondía a la **N**ormal (Anderson, 2019). En dicha demostración asume la función de densidad de los errores como $\phi(x)$, afirmando que es “simétrica, diferenciable y solo con un máximo” (Conde, 2015, p. 64), de donde se obtiene la ecuación diferencial $\frac{\phi'(x)}{\phi(x)} = kx$, que al resolverla resulta la siguiente expresión

$$\phi(x) = \frac{h}{\sqrt{\pi}} \cdot e^{-h^2 \cdot x^2}$$

que, si se analiza de forma rigurosa, presenta una forma parecida a la función de densidad de la Normal, con $h \in \mathbb{R}^+$que Gauss definía como “la precisión del proceso de medida” (Másmela y Serrato, 2007, p. 11), asumiendo que la función de probabilidad está determinada por $\mu = 0$ y $\sigma = \frac{1}{\sqrt{2}h}$ (Másmela y Serrato, 2007). Cabe señalar que la Normal, como Distribución de los errores experimentales, se establece a mediados del siglo XIX (Conde, 2015).

Por todo lo anterior, Gauss es reconocido como el autor de esta distribución. Se destaca que el desarrollo llevado a cabo por Gauss para la deducción matemática de la función de densidad de la Distribución Normal se encuentra en el documento de Másmela y Serrato (2007). Finalmente, hay que señalar que la denominación de campana, a la Distribución Normal, fue asignada por Esprit Jouffret, en 1872, quien utilizó el nombre de *bell surface* (superficie de campana) para referirse a una Distribución Normal Bivariada, y el nombre fue otorgado de forma independiente hacia el año 1875 por C. Peirce (1839 – 1914), F. Galton (1822 – 1911) y W. Lexis (1837 – 1914) (Anderson, 2019).

En consecuencia, se puede afirmar que la función de densidad de la Distribución Normal como se conoce hoy en día está dada por

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad con\ x \in \mathbb{R}$$

Esta función es dependiente de los dos parámetros, μ y σ^2, con $\mu \in \mathbb{R}$ y $\sigma^2 > 0$. Además, si X es la variable aleatoria y distribuye Normal, de forma abreviada se escribe $X \sim N(\mu, \sigma^2)$ (Rincón, 2013).

Una vez hallada la función de densidad, surgió la problemática que involucra a la función de distribución acumulada, la cual está definida por:

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$$

De allí se observa que la función de densidad es difícil de integrar puesto que no posee primitiva. Es por esto que los estadísticos a lo largo de los años han trabajado para encontrar valores correspondientes a cada una de las probabilidades acumuladas adjuntándolos en una tabla, llamada Z, pero esta tabla solo considera los valores para una Distribución Normal con media 0 y varianza 1, pues resulta imposible obtener una tabla para cada posible valor de μ y σ^2, por lo que, para poder utilizar dichos valores, se necesita estandarizar la variable aleatoria, dando paso a lo que se conoce como Distribución Normal Estándar, que surge a raíz de una transformación a la variable aleatoria Normal, utilizando la siguiente ecuación para la estandarización

$$Z = \frac{X - \mu}{\sigma}$$

donde $Z \sim N(0,1)$ y los parámetros μ y σ son la media y desviación estándar de la variable aleatoria original, respectivamente. Por lo tanto, la función de densidad de una variable con Distribución Normal Estándar corresponde a

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \qquad con\ z \in \mathbb{R}$$

y su función de distribución acumulada está dada por

$$F_Z(z) = P(Z \leq z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\, dt$$

Así, el cálculo de probabilidades asociado a esta distribución se realiza a través de la estandarización de la variable aleatoria definida junto con el uso de la tabla Z, y no integrando la función de densidad como sucede con otras distribuciones continuas (Rincón, 2013).

2.2. Conocimiento Profesional del Profesor de Estadística

El estudio y literatura sobre la formación profesional del profesor de estadística, probabilidad y su didáctica ha tenido un desarrollo reciente, en consecuencia, un desarrollo menor en comparación al conocimiento del profesor de matemática. Sin embargo, se destacan los siguientes trabajos realizados por distintos autores para determinar modelos que tratan el conocimiento del profesor de estadística y su didáctica.

Uno de ellos fue desarrollado por Batanero, Godino y Roa (2004), quienes proponen un conjunto de dimensiones encaminadas a desarrollar las siguientes capacidades del profesor de estadística, lo primero tiene relación con la capacidad de profundizar en reflexiones de tipo epistemológico sobre el significado de los conceptos que se enseñan, considerando aspectos de tipo histórico, filosóficos y culturales, así como las relaciones con otros dominios de la ciencia. En segundo lugar, se incluye la capacidad para adaptar el conocimiento estadístico a diferentes niveles de enseñanza y a las capacidades de los estudiantes. Además, se considera la capacidad crítica para analizar libros de textos y materiales curriculares. También, se refiere a la capacidad para predecir, interpretar y remediar dificultades, errores, estrategias y obstáculos de los estudiantes al resolver determinados problemas. Por último, la capacidad para diseñar e implementar situaciones didácticas incorporando materiales apropiados como medios.

Otro estudio importante, es el modelo propuesto por Burgess (2008, citado en Rivas, 2014) quien consideró los siguientes componentes del razonamiento estadístico presentado por Wild y Pfannkuch (1999). En primer lugar, el re*conocimiento de la necesidad de los datos*, le siguen la *transnumeración* (cambio de las representaciones de los datos para facilitar la comprensión), *la variación* (el pensamiento estadístico moderno se refiere al aprendizaje y la toma de decisiones bajo incertidumbre, la cual surge de la omnipresente variación), también el *uso de un conjunto de modelos* (como resúmenes de los datos en tablas, gráficos). Por último, *el conocimiento estadístico relacionado con el contexto*. Además, para completar su modelo añade los siguientes elementos: *ciclo de una investigación* (problema, plan, datos, análisis y conclusión), *el ciclo interrogativo* (generar, buscar, interpretar, criticar y juzgar) y *las disposiciones*, tales como escepticismo e imaginación. Por último, en la Figura 1 se aprecia el modelo del conocimiento profesional del profesor de estadística propuesto por Burgess.

<table>
<tr><td colspan="3" rowspan="3"></td><td colspan="4">Conocimiento estadístico para la enseñanza</td></tr>
<tr><td colspan="2">Conocimiento del contenido.</td><td colspan="2">Conocimiento didáctico del contenido</td></tr>
<tr><td>Conocimiento común del contenido</td><td>Conocimiento especializado del contenido</td><td>Conocimiento del contenido y estudiantes</td><td>Conocimiento del contenido y enseñanza</td></tr>
<tr><td rowspan="8">Pensamiento estadístico en investigaciones empíricas</td><td rowspan="5">Tipos de pensamientos</td><td>Reconocimiento de las necesidades de los datos</td><td></td><td></td><td></td><td></td></tr>
<tr><td>Transnumeración</td><td></td><td></td><td></td><td></td></tr>
<tr><td>Variación</td><td></td><td></td><td></td><td></td></tr>
<tr><td>Uso de modelos</td><td></td><td></td><td></td><td></td></tr>
<tr><td>Conocimiento estadístico relacionado con el contexto.</td><td></td><td></td><td></td><td></td></tr>
<tr><td colspan="2">Ciclo investigativo</td><td></td><td></td><td></td><td></td></tr>
<tr><td colspan="2">Ciclo interrogativo</td><td></td><td></td><td></td><td></td></tr>
<tr><td colspan="2">Disposiciones</td><td></td><td></td><td></td><td></td></tr>
</table>

Figura 1: Modelo del conocimiento profesional del profesor de estadística. (Burgess, 2008, como se citó en Rivas, 2014, p. 29).

Por su parte, Garfield y Ben-Zvi (2008) consideran que la habilidad para diseñar actividades de aula integrando herramientas tecnológicas es una competencia que deben desarrollar los profesores de estadística en su formación. Por ello, Lee y Hollebrands (2008) (como se citó en Rivas 2014) proponen un marco sobre el conocimiento profesional en relación a la integración de las Tecnologías de la Información y Comunicación (TIC) del profesor que considera cuatro componentes: Las concepciones sobre cómo enseñar contenidos específicos integrando las TIC, seguido del conocimiento de estrategias de enseñanza e implicancia de las representaciones al estudiar temas específicos con uso de TIC, también se incluye el conocimiento sobre el razonamiento, aprendizaje y comprensión al usar las TIC, y finalmente el conocimiento del currículo y los instrumentos curriculares que promueven el uso de las TIC y el razonamiento estadístico.

2.3. Conocimiento del Profesor para Enseñar Probabilidad

Vásquez y Alsina (2015) realizan un estudio enfocado en determinar las dificultades en el conocimiento didáctico – matemático que presentan los docentes de primaria respecto a dicho objeto matemático. Para ello, elaboraron un cuestionario aplicable a los docentes, conformado por siete ítems que abordan los contenidos que constituyen el significado de probabilidad para los autores, considerando conceptos como experimento y suceso aleatorio, espacio muestral, posibilidad de ocurrencia (seguro, posible, imposible), significados de la probabilidad, cálculo y comparación de probabilidades, independencia de sucesos y equiprobabilidad. Los investigadores señalan que muchos docentes "no han recibido una formación sobre la probabilidad que les permita llevar a cabo una enseñanza idónea en las aulas" (p. 28), esto se ve ratificado por los resultados de su estudio y otros realizados por distintos autores señalados en su investigación, como Begg y Edward (1999) quienes detectan que los profesores encuestados poseen una comprensión insuficiente de conceptos probabilísticos como aleatoriedad, sucesos equiprobables e independencia.

Asimismo, Watson (2001) infiere que los profesores de secundaria presentan mayor confianza respecto a los de primaria al momento de responder una encuesta sobre probabilidad y su enseñanza, añadiendo que los de primaria tienen una mentalidad orientada al cálculo de probabilidades a priori. En esa misma línea, Pereira – Mendoza (2002) ratifica las afirmaciones anteriores.

Por otro lado, Gómez (2014) investiga sobre el conocimiento matemático para la enseñanza de futuros docentes y cómo evoluciona a partir de una acción formativa, obteniendo que los docentes:

> muestran una mezcla de intuiciones y creencias correctas e incorrectas sobre la forma de percibir la aleatoriedad (…), se observan sesgos como la falacia del jugador o el enfoque en los resultados, además de concepciones erróneas sobre equiprobabilidad o falta de comprensión de independencia de sucesos (Vásquez y Alsina, 2015, p. 31).

No obstante, los resultados de Vásquez y Alsina (2015) no distan mucho de los obtenidos por los autores ya mencionados, en este caso, la puntuación media del cuestionario fue de 10,92 puntos (con 1 punto obtenido como mínimo y 28 como puntaje máximo obtenido de un puntaje ideal de 43 puntos), lo que supone que "ningún profesor resolvió correctamente el cuestionario en su totalidad" (p. 35), por lo que se considera que este resultado es "muy bajo dado que dicha puntuación considera tanto las respuestas correctas como las parcialmente correctas" (p. 35)

Al momento de consultarle a los docentes de primaria si se sienten preparados para enseñar probabilidad en el colegio, solo un 5,4% dice sentirse muy preparado, el 60,2% se siente medianamente preparado y un 34,4% no preparado. Además, un 68,8% responde que sí enseña probabilidad en sus cursos, y un 28% señala que no enseña, el porcentaje restante (3,2%) no responde. Uno de los resultados más preocupantes se obtiene cuando se pregunta si han recibido alguna formación en didáctica de la probabilidad y la totalidad de los encuestados declara no recibirla.

En palabras de los autores, el cuestionario presentó:

> un nivel de dificultad bastante alto que osciló entre un 0% y un 62% (…), la distribución de los índices es bastante asimétrica y (…) se encuentran concentrados en los valores más bajos (entre 0 y 0,0305), lo que muestra que el cuestionario resultó bastante difícil para los profesores (p. 34).

De esta forma, concluyen que los docentes poseen un conocimiento muy insuficiente en probabilidad, observando "serias dificultades para resolver correctamente las situaciones problemáticas, manifestándose variados errores y dificultades, como la presencia de heurísticas y sesgos probabilísticos asociados a la recencia negativa y positiva" (Vásquez y Alsina, 2015, p. 41) sumado al "nivel de desempeño obtenido por este grupo de profesores chilenos ha sido demasiado bajo para todos los tipos de conocimientos involucrados" (p. 41), lo que sin duda es alarmante.

El concepto de recencia – positiva o negativa – fue trabajado por Piaget e Inhelder (1951) de donde Vásquez y Alsina (2015) lo interpretaron en su estudio, mencionando que algunos docentes lo presentaban en sus afirmaciones, observando que varias respuestas fueron influenciadas por la secuencia de resultados obtenidos anteriormente en el experimento, considerando los que mayormente se repetían, esto se conoce como recencia positiva. Por el contrario, se encuentra el concepto de recencia negativa, donde las respuestas se ven afectadas por los resultados obtenidos en menor medida, inclinándose por ellos. Para ejemplificar esto, se utiliza una pregunta planteada por los autores mencionados: "Usted ha seleccionado el siguiente problema para sus alumnos de 6º básico: Al lanzar un dado 10 veces han salido los siguientes valores: 3, 6, 2, 3, 4, 4, 3, 2, 6, 2. Si se lanza el dado otra vez, ¿qué número es más probable que salga?" (p. 48). En este caso, la recencia positiva se provoca al responder "el número que es más probable que salga puede ser el 2 y 3, ya que han salido más veces" (p. 38) y la recencia negativa, cuando se responden "el 4 o el 6, porque ha salido menos veces" (p. 38).

2.4. Conocimiento Especializado del Profesor de Matemáticas

El Conocimiento Especializado del Profesor de Matemática (MTSK, por sus siglas en inglés *Mathematics Teacher's Specialised Knowledge*) es un modelo analítico del tipo descriptivo desarrollado en la Universidad de Huelva, el cual surge a partir de las problemáticas detectadas en las limitaciones tanto del modelo dicotómico establecido por Shulman (1986): el conocimiento de las matemáticas y el conocimiento didáctico del contenido (PCK, por sus siglas en inglés *Pedagogical Content Knowledge*), como del modelo del conocimiento matemático para la enseñanza (MKT por sus siglas en inglés *Mathematical Knowledge for Teaching*) planteado por Ball, Thames y Phelps en 2008, donde se identifica la necesidad de un modelo para analizar el conocimiento del profesor de matemática que "debe entender el carácter especializado en su totalidad y no solo en una de las partes y, además, tener a las matemáticas, entendida como objeto de enseñanza y aprendizaje, como elemento nuclear" (Carrillo, Contreras, Climent y Montes, 2017, p. 4).

En consecuencia, el MTSK analiza el conocimiento del profesor de matemática de forma integral, considerando tanto el dominio matemático como el dominio didáctico específico y destaca las diferentes facetas en las que el profesor conoce el contenido matemático, con la finalidad de:

> afrontar con éxito un proceso de enseñanza que acerque a sus alumnos a la naturaleza del conocimiento matemático, es decir, que los lleve a razonar, argumentar, conjeturar, refutar, representar, modelizar y, en definitiva, a hacer un uso con significado del conocimiento matemático y esto, además, en cada etapa educativa que corresponda (Carrillo, Contreras, Climent, Muñoz, Montes y Rojas, 2015, p. 590).

El MTSK posee la característica de ser dual, ya que es una propuesta teórica que modela el conocimiento núcleo del conocimiento profesional del profesor de matemáticas y además es una herramienta metodológica que permite analizar distintas prácticas del profesor de matemática a través de sus dominios y subdominios.

Es así como el MTSK está formado por dos dominios principales que son el Conocimiento Matemático (MK[1]) y el Conocimiento Didáctico del Contenido (PCK[2]) y cada uno de ellos se divide en tres subdominios que mantienen estrecha relación entre ellos. Por un lado, el Conocimiento Matemático analiza el conocimiento que tiene el profesor sobre la matemática como disciplina científica. Por su parte, el Conocimiento Didáctico del Contenido analiza aspectos relacionados con el contenido matemático como objeto de enseñanza-aprendizaje. El MTSK también considera las concepciones y creencias (*beliefs*) que tiene el profesor acerca de las matemáticas y el proceso de la enseñanza-aprendizaje, tal como se aprecia en el centro del hexágono de la Figura 2, donde aparece limitada con líneas discontinuas que representan la permeabilidad entre estas y el conocimiento presente en todos los subdominios del modelo (Carrillo, et al., 2015; Carrillo, et al., 2017; Flores et al., 2016; Yáñez, 2016).

[1] Por sus siglas en inglés Mathematical Knowledge.

[2] Por sus siglas en inglés Pedagogical Content Knowledge.

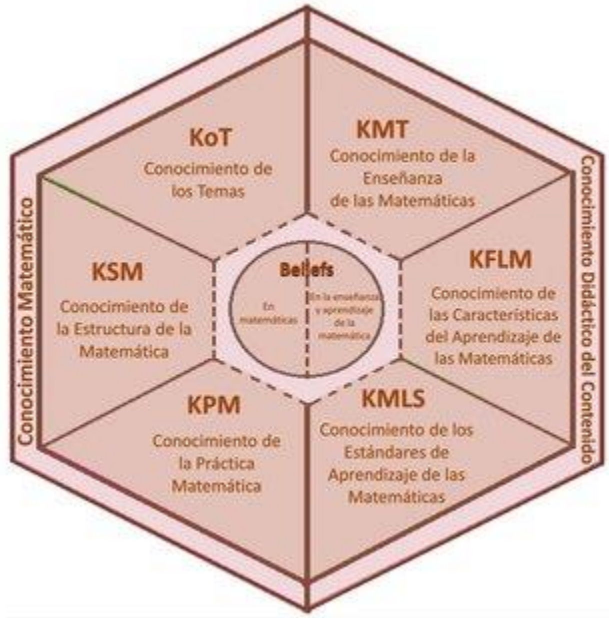


Figura 2: Esquema del modelo Conocimiento Especializado del Profesor de Matemáticas, MTSK. Con siglas en los subdominios en inglés. (Carrillo et al., 2015, p. 596).

2.4.1. Conocimiento Matemático (MK)

Un elemento fundamental del profesor de matemática es el conocimiento de la propia disciplina que enseña. Por lo cual, en el MK es preciso indagar sobre qué conocimientos matemáticos posee el docente para realizar clases idóneas, para ello se determina qué es lo que sabe o lo que debe saber de matemáticas. Entonces, los subdominios propuestos en este dominio están basados en las diferentes formas de conocer matemática disciplinar, escolar y didáctica. Estos son: Conocimiento de los temas Matemáticos (KoT), Conocimiento de la estructura de las matemáticas (KSM), Conocimiento de la práctica matemática (KPM).

2.4.1.1. Conocimiento de los temas (KoT)

El KoT no se limita al contenido que es objeto del proceso enseñanza, aprendizaje porque "es el conocimiento profundo que el profesor debe tener sobre contenido escolar, ya que entendemos que un educador puede y debe conocer el contenido más allá de lo que sus alumnos aprenden" (Carrillo et al., 2015, p. 596). En otras palabras, el KoT se trata de un

> conocimiento fundamentado, que supone conocer el contenido desde un punto de vista más profundo del que corresponde al nivel de aprendizaje en cuestión. Además, se define de modo intrínseco (refiriéndonos exclusivamente al conocimiento matemático del profesor de matemáticas, y no en función de si ese conocimiento es compartido o no con otros profesionales (Flores, Montes, Contreras, Catalán, Mar, 2016, p. 211).

Para caracterizar el Conocimiento de los Temas, Carrillo et al. (2017) lo dividen en cinco categorías. Estas son: Definiciones, Fenomenología y Aplicaciones, Propiedades y Fundamentos, Registros de representación y finalmente, Procedimientos. Con ellas se integra el contenido que se quiere que aprenda el estudiante y permite la consideración de un conocimiento con un nivel de profundización mayor al esperado para los alumnos.

a) Definiciones: Es el conocimiento del conjunto de propiedades que hace definible a un objeto, así como también, las diferentes alternativas que utilice el profesor para definir el o los objetos matemáticos. Para el caso de la Distribución Normal el profesor de enseñanza media debe conocer su definición y las diversas interpretaciones que se le otorgan. A continuación, se describen algunas definiciones para la Distribución Normal.

a.1) La Distribución Normal se entiende que es una distribución de probabilidades o, más propiamente, una densidad de probabilidades. La función de densidad de probabilidad de la variable aleatoria que distribuye Normal siempre dependerá de dos parámetros, μ y σ^2, con $\mu \in \mathbb{R}$ y $\sigma^2 > 0$, está dada por:

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad con\ x \in \mathbb{R}$$

Además, si X es la variable aleatoria y distribuye Normal, de forma abreviada se escribe $X \sim N(\mu, \sigma^2)$ (Rincón, 2013).

a.2) Por otra parte, la Distribución Normal también se puede definir como "un tipo particular de distribución continua de probabilidad. Engloba a todo un conjunto de distribuciones que tienen en común una serie de características que la diferencia de las restantes distribuciones continuas de probabilidad" (Carro Ramos, 1994, como se citó en Tauber, 2001. p. 79).

a.3) Incluso se puede definir como función de la media y de la desviación típica sin especificar su fórmula, indicando que depende de dos parámetros, donde se explica intuitivamente el significado de estos. En efecto,

> la Distribución Normal de una variable X depende de dos parámetros: μ, que puede tomar cualquier valor, y σ, que ha de ser mayor o igual que cero. La distribución suele representarse por $N(\mu, \sigma)$. Se puede probar que μ es la media o valor esperado de la variable y que σ es su desviación típica; es decir, $\mu_x = \mu$ y $\sigma_x = \sigma$. La densidad de la Normal alcanza su valor máximo en la media μ_x, que es una medida de posición y determina la situación del centro de la distribución. Por la simetría, la mediana coincide con la media. La desviación típica σ_x expresa la concentración de la distribución alrededor del valor esperado. Gráficamente, la magnitud de la desviación típica puede determinarse porque los puntos de la densidad donde cambia la concavidad son $\mu_x - \sigma_x$ y $\mu_x + \sigma_x$. (Peña y Romo 1997, como se citó en Tauber, 2001, p. 79-80)

b) Fenomenología y Aplicaciones: Conocimiento que el profesor tiene acerca de modelos atribuibles a un objeto, que son vistos como fenómenos y que pueden ser utilizados para generar conocimiento matemático. Además, para darle sentido al objeto al momento de enseñar, se incluye el conocimiento sobre sus usos y aplicaciones, pues no basta solo con conocer definiciones. Según el objeto matemático en estudio, Tauber (2001) señala que hay dos teoremas que explican la omnipresencia de la Distribución Normal en diversas áreas de aplicación, una de ellas es la propiedad de reproductividad, que implica que la suma de variables Normales es igualmente Normal. La otra indica que la suma de un número grande de variables aleatorias idénticamente distribuidas tiene una Distribución Normal. De esta forma muchos fenómenos físicos, biológicos, psicológicos o sociológicos, pueden ser adecuadamente modelizados mediante la Distribución Normal, por ejemplo, medidas antropométricas, puntuaciones en test y cuestionarios, errores de medición, etc.

c) Propiedades y Fundamentos: Conocimientos de las propiedades y fundamentos atribuibles a las características de un procedimiento, un determinado concepto o definición matemática. Para el caso particular de la Distribución Normal se pueden distinguir las siguientes propiedades:

c.1) Como la Distribución Normal depende de la media y la varianza, se obtiene que $E(X) = \mu$ y $Var(X) = \sigma^2$, cuando $X \sim N(\mu, \sigma^2)$ y $E(X) = 0$, $Var(X) = 1$, para el caso que $X \sim N(0,1)$ (Walpole, Myers R., Myers S., y Ye, 2007).

c.2) En segundo lugar, la gráfica de la distribución es cóncava si $\mu - \sigma < X < \mu + \sigma$ y es convexa en cualquier otro caso, es decir, la gráfica en $x = \mu \pm \sigma$ posee un punto de inflexión. Además de tener una asíntota horizontal con respecto al eje X a medida que se aleja de la media μ, también se puede ver que la gráfica es simétrica con respecto a este mismo parámetro. Asimismo, la moda y mediana coinciden con el valor esperado μ de la Distribución Normal, en otras palabras, la media también es el valor más probable y el que deja un 50% de información por debajo (Walpole et al, 2007).

c.3) Por otro lado, la simetría de la Distribución Normal fundamenta una de las propiedades más importantes y utilizadas, la cual corresponde a $\phi(-z) = 1 - \phi(z)$, con ella, se pueden encontrar los valores que acumulan la probabilidad α y $1 - \alpha$, solo con cambiar el signo del número z (Rincón, 2013). En la Figura 3 se puede apreciar de forma gráfica la propiedad de simetría, donde z_α corresponde al cuantil.

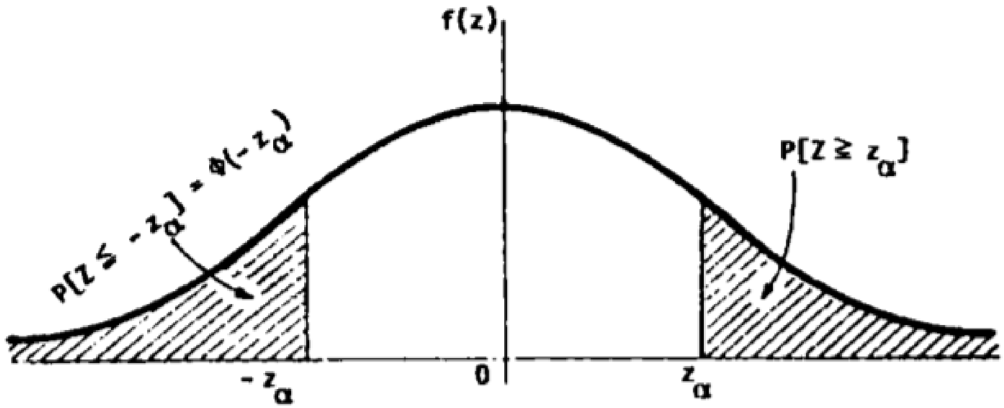


Figura 3: Propiedad de simetría de la Distribución Normal. (Moya y Saravia, 2009, p. 516)

c.4) Por último, se han calculado ciertas probabilidades estándar, que sin importar los valores que se le asignen a μ y σ^2, siempre encierran la misma probabilidad. En efecto, según Cansado (1970), aproximadamente se tiene que:

$$P(\mu - \sigma \leq X \leq \mu + \sigma) = 0{,}683$$

$$P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) = 0{,}954$$

$$P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) = 0{,}997$$

d) Registros de Representación: Es el conocimiento que tiene el profesor de las distintas formas en que se puede representar o comunicar el objeto matemático, este puede ser de forma numérica, gráfica, verbal, analítica, simbólica, algebraica, entre otras.

d.1) En el caso de la Distribución Normal, se puede representar gráficamente a través de la curva Normal, también conocida como la curva con forma de campana. tal como se aprecia en la Figura 4.

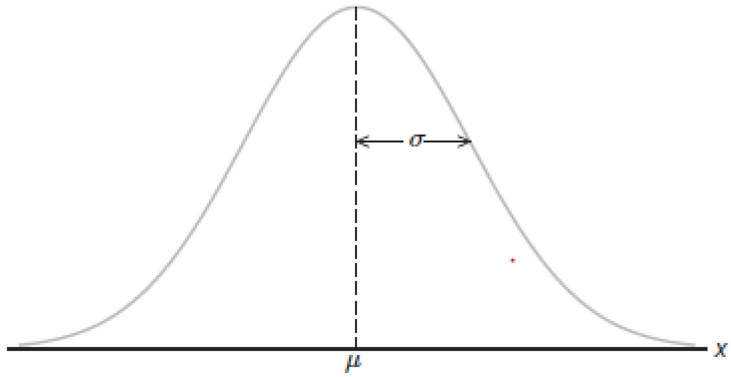


Figura 4: La curva Normal. (Walpole et al., 2007, p. 173).

d.2) Sobre las representaciones simbólicas y el lenguaje utilizado para referirse al objeto matemático en estudio se tiene que la notación usual para una variable aleatoria continua que distribuye Normal es $X \sim N\ (\mu, \sigma^2)$ y si distribuye Normal Estándar se escribe como $X \sim N\ (0{,}1)$, de igual forma se usa la letra Z para denotar la Distribución Normal Estándar y su función de distribución se puede escribir de las siguientes formas:

$$P(Z < z) = F_z(z) = \int_{\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \Phi(z), \forall z \in \mathbb{R}$$ (Rincón, 2013).

d.3) Con respecto a las representaciones numéricas específicas de la Distribución Normal, se tiene que la más utilizada es la tabla de probabilidades Normales o tabla de valores críticos. Además, entre las representaciones numéricas relacionadas con la Distribución Normal se hallan las tablas de datos, tablas de frecuencias y gráficos de tallo y hojas (Tauber, 2001).

e) Procedimientos: Es el conocimiento de algoritmos convencionales y alternativos, para ello el docente debe comprender los fundamentos de estos algoritmos y las características que tendría el objeto resultante asociadas al tema en cuestión, es decir, debe profundizar y saber el porqué de estos, sus deducciones y/o demostraciones y considerar aquellos casos donde se presentan diversos procedimientos alternativos para resolver un mismo ejercicio. Véase a continuación que la demostración del teorema de estandarización está realizada en el libro de Moya y Saravia (2009), cuyo enunciado señala que “si X es una variable aleatoria con Distribución Normal con media μ y varianza σ^2, entonces $Z = \frac{X-\mu}{\sigma}$ es una variable aleatoria con Distribución Normal Estándar” (p. 513), en otras palabras, si $X \sim N(\mu, \sigma^2)$ entonces la transformación

$$Z = \frac{X - \mu}{\sigma} \sim N(0,1)$$

Demostración:

Como $X \sim N(\mu, \sigma^2)$, su función de densidad es

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Para hallar la función de densidad $f_Z(z)$, primero se buscará $F_Z(z)$ en términos de F_X, así:

$$\begin{aligned}
&F_Z(z) = P(Z \leq z) \\
&\Rightarrow F_Z(z) = P\left(\frac{X-\mu}{\sigma} \leq z\right) \\
&\Rightarrow F_Z(z) = P(X - \mu \leq \sigma z) \\
&\Rightarrow F_Z(z) = P(X \leq \mu + \sigma z) \\
&\Rightarrow F_Z(z) = F_X(\mu + \sigma z)
\end{aligned}$$

Luego $F_Z(z) = F_X(\mu + \sigma z)$. Si se tiene la función de distribución acumulada de una variable aleatoria continua, al derivarla se obtendrá la función de densidad, así

$$\begin{aligned}
&f_Z(z) = \frac{d}{dz}\big(F_Z(z)\big) \\
&\Rightarrow f_Z(z) = \frac{d}{dz}\big(F_X(\mu + \sigma z)\big) \\
&\Rightarrow f_Z(z) = f_X(\mu + \sigma z) \cdot \sigma
\end{aligned}$$

$$\Rightarrow f_Z(z) = \left(\frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(\mu+\sigma z-\mu)^2}{2\sigma^2}}\right) \cdot \sigma$$

$$\Rightarrow f_Z(z) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{(\sigma z)^2}{2\sigma^2}}$$

$$\Rightarrow f_Z(z) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{z^2}{2}}$$

Finalmente, se tiene que $Z \sim N(0,1)$, con $Z = \frac{X-\mu}{\sigma}$, ya que su función de densidad está dada por

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{z^2}{2}} \quad con\ Rec(Z) = \mathbb{R}$$

2.4.1.2. Conocimiento de la estructura de la matemática (KSM)

Este subdominio abarca el conocimiento de las matemáticas desde la perspectiva de su integración y relación en estructuras amplias y con mayor capacidad de conexión con otros conceptos, es decir, "integra tanto aquellas relaciones con conceptos más avanzados, como más elementales, que permiten al profesor trabajar la matemática avanzada desde un punto de vista elemental y viceversa, permitiendo al profesor comprender las matemáticas escolares desde un punto de vista superior" (Carrillo, et al., 2015, p. 597).

Carrillo et al. (2017) argumentan que el conocimiento de conexiones entre objetos matemáticos es la esencia del KSM, en efecto, es el conocimiento de las conexiones inter conceptuales que el profesor hace entre los contenidos, ya sean del curso que está impartiendo o de otros niveles educativos. Se trata específicamente de conexiones entre temas matemáticos y sus distintas aristas permitiendo que los contenidos sean vistos como potenciadores para la construcción de otros o bien son potenciados por contenidos anteriores.

Para analizar el conocimiento que corresponde al profesor de matemática sobre estas conexiones se proponen cuatro categorías que Flores et al. (2016) y Carrillo et al. (2013) definen como:

a) Conexiones de Complejización: Son las relaciones que se dan entre los contenidos enseñados con contenidos posteriores. Es el caso de la relación que tiene las Distribuciones Muestrales con la Distribución Normal, pues resulta que “la distribución limitante de promedios muestrales es Normal, lo que brinda una base amplia para la inferencia estadística, que es muy valiosa para el analista de datos interesado en la estimación y prueba de hipótesis” (Meyer, 1992, p. 175)

b) Conexiones de Simplificación: Es la relación que existe entre los contenidos enseñados con los contenidos previos. Por ejemplo, el cálculo de áreas y estimación de probabilidades en un histograma. El cálculo de áreas en histogramas de frecuencias no es específico del tema de la Distribución Normal, pero sí la estimación de probabilidades a partir de las mismas (Tauber, 2001). En efecto, Behar y Grima (2013) señalan que si se pretende determinar la función de densidad de probabilidad como una extensión de la idea de histograma se debe definir el gráfico del histograma de modo que en su construcción se use intervalos de clase, de tamaño desigual, produciendo una ganancia conceptual importante, pues se tiene la obligación de representar al histograma como la construcción de rectángulos que tienen como base el intervalo de clase y su área proporcional o igual a la frecuencia relativa. De esta manera sus ordenadas representan automáticamente la función empírica de densidad, generando el enlace conceptual apropiado con la densidad de probabilidad de una variable aleatoria continua.

c) Conexiones de Contenidos Transversales: No son conexiones de contenidos más simples o complejos entre sí, sino que hay una cualidad común en estos que les relaciona, y los modos de pensamiento asociados a dichos temas contemplan esta característica común. Respecto a la Distribución Normal se tiene que, responde a las leyes de reproductividad, es decir que, la suma de variables aleatorias independientes que distribuyen Normal es una variable aleatoria que también tiene el mismo tipo de distribución (Meyer, 1992).

d) Conexiones Auxiliares: Son las relaciones que se establecen cuando un concepto utiliza otro concepto en sus procedimientos. Por ejemplo, el uso de la Distribución Normal para determinar Intervalos de Confianza, donde X tiene distribución

$N(\mu, \sigma^2)$, cuyo valor de σ es conocido, mientras que μ es el parámetro desconocido (Meyer, 1992).

2.4.1.3. Conocimiento de la práctica matemática (KPM)

Este subdominio corresponde a aquellas formas de hacer y proceder en matemáticas que sin duda un profesor ha de conocer para desarrollar su clase, los aspectos de comunicación, argumentación, validación y demostración en matemática, el significado de definición, las condiciones necesarias y suficientes para generar definiciones matemáticas, axioma o teorema como elementos constituyentes de la matemática o el conocimiento de la sintaxis matemática (Carrillo et al., 2015; Carrillo et al., 2017). También es considerado como el conocimiento heurístico en la resolución de problemas. Entonces, este subdominio plantea que el docente no solamente debe conocer los resultados de los contenidos planteados en el KoT, sino que también debe saber las formas de llegar a dichos resultados y las características del trabajo matemático (Flores et al., 2016).

En efecto, Escudero (2015) señala seis indicadores que están formulados como conocimientos que responden al cómo se desarrolla y trabaja genéricamente en matemática, independientemente del tema abordado, estos son:

- Jerarquización y planificación como forma de proceder en la resolución de problemas matemáticos.
- Formas de validación y demostración.
- Papel de los símbolos y uso del lenguaje formal.
- Procesos asociados a la resolución de problemas como forma de producir matemática.
- Prácticas particulares del quehacer matemático, por ejemplo, la modelación.
- Condiciones necesarias y suficientes para generar definiciones.

2.4.2. Conocimiento Didáctico del Contenido (PCK)

Otro elemento fundamental del MTSK es el conocimiento didáctico del contenido (PCK), que corresponde al conocimiento particular del docente de matemáticas en su labor de enseñanza. En el PCK se vincula la importancia de que el profesor conozca los puntos de vista desde el contenido a aprender, a enseñar y como una visión global de los estándares por lograr con el conocimiento matemático. Se compone por tres subdominios, estos son: Conocimiento de la Enseñanza de las Matemáticas (KMT), Conocimiento de las Características del Aprendizaje de las Matemáticas (KFLM) y Conocimiento de los Estándares de Aprendizaje de las Matemáticas (KMLS) (Yáñez, 2016).

2.4.2.1. Conocimiento de la enseñanza de la matemática (KMT)

Este subdominio tiene como foco la enseñanza, y se concentra en el conocimiento potencial del profesor que tiene relación con las vías, recursos o formas que se poseen para enseñar la matemática, se espera que esta interrelación sea intrínseca sobre los conocimientos que son de una naturaleza matemática, pues esta relación condiciona la enseñanza, y dejará excluida las visiones pedagógicas generales (como el trabajo en equipo). Este subdominio está compuesto por las siguientes tres categorías que varios autores (Carrillo et al., 2015; Carrillo et al., 2017; Flores et al., 2016; Pizarro, 2015; Yáñez, 2016) definen como:

a) Conocimiento de las teorías institucionales o personales de enseñanza, que consideran el conocimiento del profesor sobre teorías de enseñanza de la educación matemática, y conocimientos potenciales de las actividades utilizadas por ellos, también incorpora los ejemplos típicos, metáforas, analogías, entre otros, que los docentes consideren potentes para el abordaje del contenido matemático, y que aporten de forma significativa a los estudiantes en un momento particular de la enseñanza. Escudero (2015) señala que esta categoría, “trata de la inclusión de elementos teóricos que derivan directamente de los estudios específicos de la Didáctica de la Matemática como disciplina científica, correspondientes a teorías de enseñanza” (p. 49). Hasta el momento no existe una teoría didáctica exclusiva diseñada para el proceso enseñanza-aprendizaje en estadística y probabilidad, sin embargo, Escudero (2015), sugiere como ejemplo, el conocimiento de la estructura general propuesta en la Teoría de Situaciones Didácticas de Brousseau (2007), que contempla las etapas de Acción, Formulación, Validación e Institucionalización, bajo las cuales pueden diseñarse actividades para el aula y ambientes de trabajo matemático adhoc a dichas actividades.

b) El conocimiento de recursos (físicos o virtuales), materiales y formas de presentar el contenido, para ilustrar de forma concreta y/o pictórica lo que sucede con el conocimiento matemático, y que favorecen el proceso de enseñanza de la matemática, teniendo una estrecha relación con el contenido matemático, por ello, Pizarro (2015) destaca que la motivación del estudiantado no se considera en esta categoría del KMT, pues cualquier fundamento pedagógico que relacione el uso del recurso en la sala de clases es imprescindible. Como ejemplo de la categoría se encuentran distintos software matemáticos, libros de textos, calculadoras, entre otros. Dentro de la web se pueden encontrar diversos archivos provenientes de GeoGebra, y otras aplicaciones sobre la Distribución Normal, que apoyan la comprensión de esta, ya sea a partir de su forma, simetría y estandarización.

   b.1) Por ejemplo, Tauber (2001) y Batanero (2001) mencionan que, para realizar una aproximación intuitiva y experimental – inductiva de la Distribución Normal, es preciso utilizar el aparato de Galton para captar las simetrías de las trayectorias posibles de las bolitas al caer en el aparato, donde al dejar caer algunas bolitas por el orificio superior, se dispersan al azar, recogiéndose en unas casillas colocadas al final del plano inclinado. Las casillas centrales reciben más bolas que las extremas y en consecuencia la disposición es simétrica y su representación matemática de la distribución de las bolitas corresponde a la curva de Gauss, es decir, a una Distribución Normal. Esto es posible apreciarlo en la Figura 5. Se tiene dos formas de utilizar el aparato de Galton, la primera, construyendo material concreto, en particular un Quincux, que tiene forma de plano inclinado de madera donde se han colocado unos clavos regularmente espaciados o utilizando un simulador del aparato, tal como lo muestra Romero, M. (2020).

Figura 5: Aparato de Galton. (Aparicio, 2018)

b.2) Adicionalmente, dentro de la web se pueden encontrar distintos archivos provenientes de GeoGebra sobre la Distribución Normal, que apoyan la comprensión de esta, ya sea a partir de sus parámetros y cómo influyen en la forma de la campana, como el creado por Giselle (2015) o para calcular probabilidades de la forma $P(a < X < b)$, mostrando en la gráfica la porción de campana encerrada entre a y b, como lo muestra Morales. R, (2014), o uniendo ambos archivos en uno solo como lo presentan Yague, A., Morón, J., y López, I. (2020).

c) Conocimiento de las actividades, tareas, ejemplos y ayudas, que tienen una intencionalidad de enseñanza frente al tema en juego, por ello, estas siempre deben poseer un objetivo acorde al contenido matemático y marco curricular, considerando el enfoque del KMT, que es la enseñanza. Flores et al. (2016) diferencian el foco que tienen el KoT y KMT sobre el punto de vista de las actividades, señalando que "los ejemplos y representaciones del contenido son considerados desde el punto de vista de su potencial para el aprendizaje (a diferencia de las representaciones consideradas en el conocimiento del tema, desde el punto de vista de su potencial matemático)" (p. 213). Respecto a esta categoría en la Distribución Normal, se puede encontrar las aplicaciones en la vida real, y considerarla como una distribución que modela una infinidad de fenómenos cotidianos como problemas físicos (distribución de los errores de observación o medida), problemas biológicos (distribución de medidas físicas de un conjunto de personas), datos psicológicos (distribución del coeficiente de inteligencia, puntuaciones de exámenes o pruebas), problemas técnicos (distribución de las medidas de piezas manufacturadas), entre otros. (Batanero y Godino, 2001). Otro ejemplo, puede ser el siguiente:

> El nivel de colesterol en sangre es importante porque un nivel alto puede incrementar el riesgo de ataque al corazón. La distribución del nivel de colesterol en sangre en una población de personas de la misma edad y sexo es aproximadamente Normal. Para niños de 14 años, la media es $\mu = 170\, mg$ de colesterol por dl de sangre (mg/dl) y la desviación estándar $\sigma = 30\, mg/dl$. Niveles cercanos a $240\, mg/dl$ pueden requerir atención médica. ¿Qué porcentaje de niños tienen más de $240\, mg/dl$ de colesterol? (Moore 1995, como se citó en Tauber, 2001, p.77).

2.4.2.2. Conocimiento de las características del aprendizaje de las matemáticas (KFLM)

Este subdominio tiene su base en el "contenido matemático como objeto de aprendizaje, por ello evita mirar al estudiante en sí" (Pizarro, 2015, p. 56), ya que su idea es examinar características de la comprensión por parte de los estudiantes sobre el contenido, gracias al KFLM los docentes pueden anteponerse al comportamiento de los estudiantes frente a la interacción con el conocimiento, es importante recalcar que este subdominio solo incluye el cómo los estudiantes piensan y aprenden el contenido matemático y las formas de interactuar con él (Carrillo et al., 2015). Las cuatro categorías que componen este subdominio según Carrillo, et al. (2015), Carrillo, et al. (2017), Flores et al. (2016), Pizarro (2015) y Yáñez (2016) son:

a) El conocimiento de teorías de aprendizaje asociadas a un contenido matemático, que refiere a las estructuras, o teorías (personales o institucionales) sobre la apropiación de contenido de los estudiantes, ya sea de los contenidos específicos como de la matemática general. Se trata del conocimiento que permite al profesor organizar y dar sentido a la enseñanza del contenido. Por ejemplo, "el conocimiento de la secuencia de Acciones, Procesos, Objetos y Estructuras (Arnon, et al., 2014) como explicación del desarrollo cognitivo del estudiante cuando este se enfrenta a la tarea de aprender en matemáticas" (Escudero, 2015, p. 44).

b) El conocimiento de fortalezas y dificultades asociadas al aprendizaje, que se enfocan en las posibles inconvenientes, errores o capacidades potenciales que se pueden dar frente al contenido específico y a la matemática en general, por parte de los estudiantes. Como expresa Tauber (2001), las dificultades que se asocian a la Distribución Normal se pueden clasificar en tres apartados.

b.1) Desarrollo de la idea de distribución y convergencia, donde se abordan dos conceptos, referidos a la intuición de azar, y la idea de distribución centrada, concluyendo que además de requerir un conjunto de conceptos, es necesaria una madurez conceptual por parte de los estudiantes.

b.2) Comprensión de propiedades del objeto matemático, que considera conceptos como:

i. La estandarización, que conlleva a confundir el rango de variación entre las puntuaciones Normales, y las estandarizadas.

ii. La Distribución Normal como aproximación de distribuciones discretas, en el que se encuentra la dificultad en diferenciar las distribuciones discretas de las continuas, y la utilización de la corrección de Yates de forma mecánica sin saber bajo qué circunstancias debe usarse.

b.3) Hay dos factores que se relacionan e influyen en la comprensión del objeto matemático, el primero considera la ansiedad epistemológica de los sujetos frente al concepto de Distribución Normal. El segundo contempla el efecto que tiene en el proceso de aprendizaje de este objeto, ya que esta no cuenta con un medio o un material de apoyo que sea suficientemente potente, diverso y ayude a comprender cómo emerge.

c) El conocimiento de las formas de interacción de los estudiantes con el contenido matemático, donde se consideran los posibles procesos o estrategias que utilizan los estudiantes, dentro de la cual también tiene cabida el lenguaje o vocabulario usado por los estudiantes para el contenido. Por ejemplo, una de las estrategias que es posible observar respecto a la estandarización de la Distribución Normal, es descrita por Moore, 1995 (como se citó en Tauber, 2001), señalando que

> En vez de querer encontrar una proporción, puede suceder que queramos conocer el valor observado con una proporción dada de las observaciones por arriba o debajo de ese valor. Para hacer esto, usamos la tabla de la Normal típica a la inversa, es decir, buscamos la proporción dada en el cuerpo de la tabla, y leemos el correspondiente valor de z en la columna izquierda y en la fila superior, luego se "desestandariza" para obtener el valor observado. Para "desestandarizar" utilizamos la fórmula de tipificación, pero despejando X, de la siguiente forma: $X = \mu - z \cdot \sigma$. (p. 75)

En otras palabras, corresponde a tener una probabilidad en la escala Z, y querer encontrar su valor respectivo en la escala X.

d) El conocimiento de concepciones de los estudiantes sobre matemáticas, que se refiere a los saberes que tiene el docente respecto a las expectativas e intereses que poseen los estudiantes sobre el contenido matemático. Respecto a la Distribución Normal, se debe considerar que en la vida cotidiana existe un conocimiento básico de lo que refiere la normalidad, y esto debería llamar la atención de los estudiantes y facilitar su comprensión, sin embargo, los estudiantes suelen presentar problemas en el eje Datos y Azar, específicamente en el área de Probabilidades, variable aleatoria, Distribución Normal y Medidas de Tendencia Central.

2.4.2.3. Conocimiento de los estándares de aprendizaje de las matemáticas (KMLS)

Está enfocado en el conocimiento curricular del docente respecto al contenido matemático bajo dichas normativas o en documentos e investigaciones que aporten recomendaciones o estándares de aprendizaje al proceso educativo del objeto matemático. Este foco viene de la mano del intento de desarrollar una visión más amplia a partir de los diferentes grados de dificultad y profundidad que el profesor puede otorgarle al conocimiento, considerando la componente matemática del nivel y vigencia del contenido a raíz del territorio donde imparte las clases, es decir, que el docente le dé una ubicación contextual y temporal al contenido abordado. También considera los conocimientos previos y posteriores al objeto matemático. Dentro del subdominio se encuentran tres categorías que varios autores (Carrillo, et al., 2015; Carrillo, et al., 2017; Flores et al., 2016; Yáñez, 2016) las definen como se detalla a continuación.

a) El conocimiento sobre las expectativas de aprendizaje dentro de un nivel específico, relacionado con la finalidad de conocimiento matemático en el nivel que se debe impartir, y referido a las metas y objetivos estipulados para ello. Para la Distribución en el curso del electivo, el MINEDUC (2021b) indica que el objetivo de aprendizaje OA 3 consiste en "modelar fenómenos o situaciones cotidianas del ámbito científico y del ámbito social que requieran el cálculo de probabilidades y la aplicación de las distribuciones Binomial y Normal" (p. 27), y los indicadores de evaluación asociados al OA son los que se señalan a continuación: (p. 152)

- Identifican las principales características de una Distribución Normal de probabilidades.
- Interpretan información estadística que involucra distribuciones de probabilidad Binomial y el Normal.
- Resuelven problemas que involucran los modelos Binomial y Normal.
- Argumentan cuándo se puede modelar una situación o fenómeno con una distribución Binomial o Normal.

- Modelan fenómenos o situaciones cotidianas, científicas y sociales mediante distribuciones Binomiales y Normales.
- Modelan situaciones que involucran aproximar una distribución Binomial mediante el modelo Normal.

b) El conocimiento sobre el nivel de desarrollo conceptual o procedimental esperado, para el contenido y nivel en concreto, en otras palabras, que el enfoque que tenga el objeto matemático sea adecuado para el curso donde se impartirá la clase. De acuerdo con el currículum chileno, el contenido referido a la Distribución Normal, este aborda en el electivo del plan diferenciado Científico-Humanista en los niveles de tercero o cuarto medio (MINEDUC, 2021b).

c) La secuenciación de diversos temas, donde se relaciona el contenido matemático con temas anteriores y posteriores, haciendo su respectiva conexión. El vínculo que posee la Distribución Normal con los contenidos curriculares anteriores o posteriores a la enseñanza de esta se evidencia en la Tabla 1. Cabe señalar, que los contenidos y secuenciación respecto al objeto matemático fueron extraídos desde el Programa de Estudio del electivo, según lo que se señala el MINEDUC (2021b).

Tabla 1: Conexión entre la Distribución Normal y diversos conceptos estadísticos.

Contenido	**Es anterior o posterior a la Distribución Normal**	**Conexión entre el contenido y la Distribución Normal**
Variable aleatoria	Anterior	La Distribución Normal es una variable aleatoria continua, considerada especial, ya que se usa para el tratamiento de variables aleatorias estudiadas con anterioridad.
Cálculo de probabilidades de una variable aleatoria continua	Anterior	Los teoremas que se trabajan a nivel de una variable aleatoria cualquiera, se heredan para la Distribución Normal.

Aproximación intuitiva de la Distribución Normal a partir de una Distribución Binomial.	Anterior	Se induce la forma de la Distribución Normal a partir de un gráfico de barras que indica las probabilidades de una Distribución Binomial que realiza cada vez más repeticiones del experimento Bernoulli.
Distribución de las medias muestrales (Teorema Central del Límite)	Posterior	En promedio todas las distribuciones (sean discretas o continuas) convergen a una Distribución Normal, bajo ciertas condiciones $(n > 30)$, es decir $\overline{\mathrm{X}} \longrightarrow N\left(\mu, \frac{\sigma^2}{n}\right)$
Aproximación de la Binomial a la Normal (Teorema de De Moivre - Laplace)	Posterior	Considera que una variable con Distribución Binomial, que cumple ciertas condiciones, tiende en distribución a una Normal $N(np, npq)$.
Intervalos de Confianza para la media con varianza poblacional conocida	Posterior	Se asume que la muestra aleatoria obtenida proviene de una población $N(\mu, \sigma^2)$. Además, para la construcción del intervalo, se usa la función pivotal $Q(\mu, \overline{\mathrm{X}}) = \left(\frac{(\overline{\mathrm{X}} - \mu)}{\sigma} \cdot \sqrt{n}\right) \sim N(0,1)$
Test de Hipótesis para la media con varianza poblacional conocida.	Posterior	Se asume que la muestra aleatoria obtenida proviene de una población $N(\mu, \sigma^2)$, con μ desconocido. Asimismo, se considera que $\overline{\mathrm{X}} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$ para la construcción del criterio de aprobación o rechazo de la hipótesis nula.

Fuente: Elaboración Propia.

Capítulo III. Marco Metodológico

Esta investigación se centró en dos tareas principales, el primero de ellos consistió en elaborar una escala estimativa para evaluar críticamente las actividades propuestas en el Programa de Estudio del plan diferenciado de Probabilidades y Estadísticas Descriptiva e Inferencial, específicamente en el contenido de la Distribución Normal. Posteriormente, a raíz del análisis y reflexión de los resultados obtenidos, el estudio se enfocó en la elaboración de una propuesta pedagógica que incluye: (1) preguntas y subactividades formuladas por los investigadores para complementar las actividades planteadas en el Programa de Estudio, (2) desarrollo paso a paso de todas las actividades de la propuesta ministerial y de las propuestas en la investigación, (3) realizar la "bajada del contenido", es decir, el saber sabio de la transposición didáctica dirigida al docente.

3.1. Metodología de la Investigación

El objetivo de esta investigación consiste en analizar el conocimiento matemático del profesor para la enseñanza de la Distribución Normal a partir de una revisión crítica al Programa de Estudio de Probabilidades y Estadísticas Descriptiva e Inferencial, por ende exige que se desarrolle bajo un procedimiento de carácter cualitativo, pues resulta necesario usar prácticas interpretativas que permitan visualizar el objeto en estudio, transformarlo y convertirlo en una serie de representaciones en forma de observaciones, anotaciones y documentos (Hernández, et. al, 2006).

A su vez, la temporalidad de esta investigación fue transversal ya que, tanto la descripción de las variables como el análisis de su incidencia ocurre en un momento dado (Hernández, et. al, 2006), en otras palabras, es limitada por el contexto y el tiempo.

Por último, la investigación-acción es el diseño que mejor se ajusta a los requerimientos propios del fenómeno de interés. Efectivamente, Sandín (2003) afirma que la mayoría de los autores ubican este modelo de investigación en los marcos referenciales interpretativos y críticos. Por su parte, Hernández, et. al (2006) señalan que su "propósito fundamental se centra en aportar información que guíe la toma de decisiones para programas, procesos y reformas estructurales" (p. 706). Además, promueve el conocimiento, por lo que no solo es entendida como un método de investigación, sino como una herramienta epistémica orientada hacia el cambio educativo (Sandín, 2003).

3.1.1. Diseño de Investigación-Acción.

Desde la perspectiva educativa, Suarez Pazos (2002, citado en Colmenares y Piñero, 2008) refiere que la investigación acción es “una forma de estudiar y de explorar una situación social, en nuestro caso educativa, con la finalidad de mejorarla, en la que se implican como *indagadores* los implicados en la realidad investigada” (p. 104).

Con el fin de mejorar y transformar la problemática en estudio, los investigadores asumen el rol de agentes externos ajustando su trabajo a la modalidad técnica en la cual se inscribe esta metodología de investigación, la que se fundamenta en diseñar y aplicar un plan de intervención eficaz para la mejora de habilidades profesionales y resolución de problemas. En ella los agentes externos actúan como expertos responsables de la investigación, estableciendo las pautas que se deben seguir (Colmenares y Piñero, 2008).

Particularmente el desarrollo de esta investigación se ajustó al modelo procedimental propuesto por Kemmis en 1989, el que se constituye por cuatro fases que siguen un patrón cíclico. La primera de ellas es la planificación, aquí se toma de manera consciente y crítica la información que se conoce, previo diagnóstico de la situación problemática y la formulación de los objetivos deseables de alcanzar. Cabe señalar que la planificación se programa con cierta flexibilidad y adaptabilidad. Luego, sigue la fase de acción donde se aplican las acciones del plan en forma controlada y organizada. Posteriormente se tiene la fase de observación de la acción, con el fin de recoger evidencias que ayuden a evaluarla. En último lugar se tiene la fase de reflexión, ella se plantea sobre la acción registrada durante el momento de la observación y se desarrolla por la discusión con los participantes.

La ejecución de estas cuatro fases conduce a generar una nueva situación cuya consecuencia es posiblemente la necesidad de planificar una nueva etapa para el proceso de mejora continua, requiriendo volver a repetir el ciclo procedimental. No obstante, es importante aclarar que para efectos de esta investigación se llevó a cabo un solo ciclo de investigación-acción.

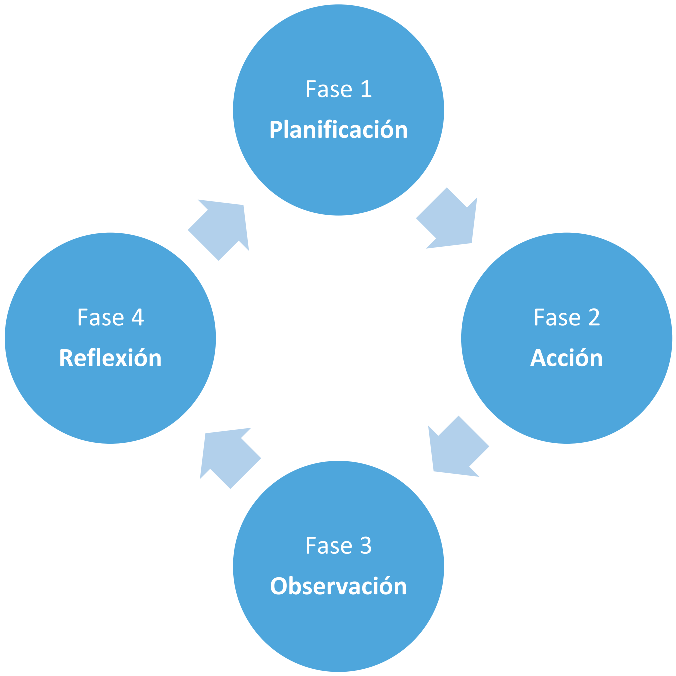


Figura 6: Modelo Procedimental para la metodología de Investigación-Acción. (Elaboración propia a partir de Kemmis, 1989, citado en Colmenares y Piñero, 2008).

3.2. Contexto del Objeto de Estudio

El objeto de estudio de esta investigación se focalizó en la enseñanza-aprendizaje de la Distribución Normal en la educación escolar media chilena, específicamente en el curso Probabilidades y Estadísticas Descriptiva e Inferencial, el cual está constituido por cuatro unidades:

- Unidad 1: ¿Qué dicen los gráficos? Análisis crítico de la información.
- Unidad 2: Comprender la media muestral, las medidas de dispersión y la correlación
- Unidad 3: Modelaje de fenómenos mediante las probabilidades las distribuciones Binomial o Normal
- Unidad 4: Hacer inferencia estadística.

La estructura de cada Unidad se compone de cuatro elementos bien definidos, tal como se detalla en la Tabla 2.

Tabla 2: Estructura de las Unidades del plan diferenciado de Probabilidades y Estadísticas Descriptiva e Inferencial.

Elemento	Descripción
Propósito de la Unidad	Posee un resumen del objeto formativo detallando lo que se espera que el estudiante aprenda, actuando como una guía para el conjunto de actividades que se desarrollan en la unidad.
Objetivos de Aprendizaje	Detallan los aprendizajes asociados a la unidad.
Actividades de Aprendizaje	Movilizan de forma integrada las actitudes, habilidades y conocimientos para el lograr la cobertura del Objetivo del Aprendizaje asociado al contenido tratado según corresponda. Cabe mencionar que estas sirven de guía para el diseño de los ejercicios a desarrollar en clase, por lo que es opcional el uso de estas. Además, al principio de cada actividad se señala el Propósito de ellas junto con las actitudes a trabajar en el desarrollo de esta. Asimismo, al final de la actividad se detallan las orientaciones para el docente, se señalan indicadores de evaluación que permite medir el alcance de los contenidos, por último, está el apartado *recursos o sitios web* a los que puede recurrir el docente para realizar y/o complementar su clase
Actividad de Evaluación	Plantea una serie de preguntas que abordan todos los contenidos de la Unidad. A su vez, se proporcionan los respectivos *indicadores de evaluación* globales, a partir de lo que se ha tratado en el transcurso de la unidad.

Fuente: Elaboración propia a partir de MINEDUC (2021b).

3.2.1. Descripción de las actividades del Programa de Estudio.

Cada Unidad se compone por cinco actividades, cuatro de aprendizaje y una evaluativa. La investigación se enfocó en la Unidad 3, exclusivamente en aquellas actividades referentes a la Distribución Normal. En consecuencia, se excluyó la primera actividad de esta Unidad, titulada *Experimentos aleatorios con modelos de Bernoulli y Binomial*, puesto que no se vincula con el objeto de estudio. Cabe señalar que las actividades para la enseñanza de la Distribución Normal comienzan desde la segunda en la propuesta ministerial, no obstante, para efectos de organización del análisis y estudio de cada una de ellas se enumeran de 1 a 4 (ver Tabla 3), donde la primera corresponde a la segunda de la propuesta ministerial, análogamente se mantiene esta relación con las siguientes. Además, se denominó actividad 4 a la actividad evaluativa de la unidad de trabajo.

Tabla 3: Organización de las Actividades de la Unidad 3 para la investigación.

Título de la Actividad	**N° de la Actividad en la propuesta ministerial**	**N° de la Actividad en la investigación**
Comprender el modelo Normal de probabilidades	2	1
Aplicar el modelo Normal en el transporte de personas	3	2
Aproximar la Distribución Binomial por la Distribución Normal	4	3
Actividad de Evaluación	Sin número	4

Fuente: Elaboración Propia

3.3. Ciclo Investigación-Acción

A continuación, se describe cada una de las fases del ciclo de investigación-acción en función de la problemática que se aborda en este estudio.

3.3.1. Fase de Planificación y Fase de Acción

La Fase de Planificación inicia con la identificación de un problema educativo sobre el cual puede realizarse un cambio. De esta forma, la problemática planteada aborda los contenidos para la enseñanza de la Distribución Normal en las actividades de la propuesta ministerial para el electivo de Probabilidades y Estadísticas Descriptiva e Inferencial. Para ello, se diseñó un instrumento que permitió a los investigadores junto a la profesora guía de este trabajo, Giovanna Ticchione, evaluar críticamente las actividades en relación al objetivo de aprendizaje y los indicadores de evaluación definidos en la Unidad 3 (ver Anexo 1), cuya aplicación evidenció no solo los contenidos que están presentes en menor proporción sino también reveló aciertos y desaciertos, adicionando otras problemáticas específicas que requieren ser atendidas en el ciclo de investigación. (ver Anexo 2).

En general la revisión permitió identificar errores de notación y cohesión en algunas preguntas. También se obtuvo que tanto los parámetros, como la propiedad simétrica, la estandarización y desestandarización de una variable aleatoria continua con Distribución

Normal no son tratados eficazmente teniendo presente la línea deductiva que presentan estas actividades, dificultando su aprendizaje.

Comprendiendo que este es un curso de profundización del objeto matemático con respecto al curso del plan de formación general, surge la necesidad de plantear subactividades que complementen las ya propuestas por el MINEDUC en relación con aquellos contenidos de mayor complejidad para los estudiantes, como lo son el teorema del Límite Central y el teorema de De Moivre-Laplace.

Considerando los antecedentes observados, el plan de acción consiste en elaborar subactividades y preguntas complementarias a las de la Unidad 3, estas abordarán contenidos asociados a la Distribución Normal tratados en menor proporción o que requieran mayor profundidad, incorporándolas a la propuesta ministerial en correspondencia al propósito de cada actividad. Además, se plantea remediar los problemas de notación y cohesión entre las preguntas, gráficas y contenidos matemáticos. Adicionalmente, se especifica realizar la resolución tanto de las actividades de la propuesta ministerial como la de los investigadores, de esta forma se presentarán procedimientos que permitirán al profesor identificar diferentes maneras de dirigir el conocimiento.

En la Fase de Acción se implementan todas las tareas del plan de acción mencionadas anteriormente. En efecto, se llevó a cabo la resolución de las cuatro actividades de la propuesta ministerial asociadas a la Distribución Normal, añadiendo los algoritmos para el uso de las herramientas digitales, las gráficas de la curva normal, la tabla Z, histogramas, tablas de frecuencia y notas al docente en casos particulares. Igualmente, se incorporó un total de cinco subactividades y trece preguntas para complementar las actividades de la Unidad 3. Finalmente, el plan de acción y Fase de Acción de cada actividad de la propuesta ministerial se detallan posteriormente según la necesidad de cada una de ellas.

Por otro lado, considerando la transposición didáctica de la Distribución Normal se elaboró un documento con el saber sabio que encierra todos los contenidos vinculados al objeto matemático que son tratados en la Unidad 3, haciendo uso del conocimiento matemático del MTSK se abordó: (1) la definición de variable aleatoria, (2) función de densidad, (3) definición de Distribución Normal y sus propiedades, (4) Distribución de las medias muestrales, o distribución de $\overline{X}$, (5) Aproximación de la Distribución Binomial a la Normal.

La aplicación de todas las tareas realizadas en la Fase de Acción, resultan en una propuesta pedagógica que puede consultar en el Anexo 5.

3.3.1.1. Plan de Acción y Fase de Acción de la Actividad 1.

La Actividad 1: *Comprender el modelo Normal de probabilidades*, se subtitula *Desarrollo y significado de la Distribución* Normal. A su vez cuenta con cinco subactividades que la propuesta ministerial denomina *Pasos.* Se debe aclarar que nombrar de esta forma a las subactividades es una particularidad que ocurre únicamente en la Actividad 1. El plan de acción para cada Paso se describe en la Tabla 4.

Tabla 4: Descripción y Plan de Acción Actividad 1.

N° Paso	Título	Descripción del Paso	Plan de Acción
1	El concepto de una variable aleatoria continua	Aborda la introducción a este concepto matemático por medio de la comparación entre la frecuencia relativa y probabilidad clásica de un experimento con números aleatorios. Se divide en preguntas de la a) hasta la f).	➢ Reemplazar la gráfica de la pregunta a) por una gráfica lúdica. ➢ En la pregunta d) elaborar un nuevo enunciado bajo cierto contexto para su cohesión con la pregunta anterior y posterior.
2	Sin título	Comienza mostrando gráficos de barras de Distribuciones Binomiales, manteniendo constante la probabilidad de éxito y variando la cantidad de repeticiones, luego se presenta una serie de preguntas que, a partir del análisis de la forma de los gráficos anteriores introducen la idea de campana de Gauss desde la Binomial. Se divide en preguntas de la a) hasta la c).	➢ Arreglar el error conceptual de la gráfica presentada en el enunciado inicial y añadir los títulos a los ejes. Además de eliminar la pregunta c) para incorporarla a la primera subactividad propuesta. ➢ Titular la Actividad 2 como: *Aproximándonos a la Distribución Normal*.
3	Significado de la	Trata la notación ϕ para hacer	➢ Entre el Pasos 2 y 3 incorporar

	Distribución Normal	uso de la tabla Z, calculando las probabilidades icónicas de la Distribución Normal (que se encuentran a 1, 2 y 3 desviaciones estándar de la media). Se divide en preguntas de la a) hasta la c).	tres subactividades elaboradas por los investigadores, donde en la primera, permite reconocer los parámetros de la Distribución Normal, la segunda, ejercitar el cálculo de probabilidades con la tabla Z y la tercera, visualizar la propiedad simétrica. ➢ Se integra el ítem D, que se subdivide en tres preguntas, con ellas se resuelven problemas usando las probabilidades icónicas para determinar parámetros de distintas situaciones.
4	Aproximación de una Distribución Binomial por la Distribución Normal Estandarizada	La aproximación se lleva a cabo mediante un experimento dicotómico que se realiza una cantidad de veces considerable. Se divide en preguntas de la a) hasta la d).	➢ Entre los Pasos 3 y 4 se añade una actividad para el cálculo de probabilidades y cuantiles a través del uso de la tabla Z en forma inversa.
5	Aplicación de la Distribución Normal	Aplica los conceptos aprendidos de la Distribución Normal con dos problemas puestos en contexto. Se compone por dos ítems A y B. El ítem A se divide en dos preguntas a) y b) y el ítem B se divide en preguntas desde la a) hasta la c).	➢ Tanto en el ítem A como en el B se integran dos preguntas acordes al contexto en el cual se aplica la desestandarización de la Distribución Normal. Adicionalmente se añade un ítem C que permite calcular probabilidades utilizando los intervalos icónicos y desestandarizando. ➢ Eliminar enlaces caducados.

Fuente: Elaboración Propia

En la Fase de Acción, en total se incorporaron diez preguntas distribuidas en los Pasos 3 y 5 (ver Tabla 5) de la propuesta ministerial. A su vez se elaboraron cuatro subactividades (ver Tabla 6) elaborada por los investigadores entre algunos Pasos de la Actividad 1.

Tabla 5: Preguntas complementarias para la Actividad 1 de la propuesta ministerial.

N° de Paso	Título del Paso	Cantidad de preguntas añadidas	Ubicación de las preguntas en la propuesta ministerial	Objetivo de las preguntas
3	Significado de la Distribución Normal	3	Se agrega la pregunta d) que se subdivide en i) y ii), además se incorpora la pregunta e).	Resolver problemas que involucren el uso de los intervalos icónicos y estandarización para el cálculo de probabilidades.
5	Aplicaciones para la Distribución Normal	4	Al ítem A se añaden las preguntas c) y d). Del mismo modo el ítem B agrega las preguntas d) y e).	Aplicar la desestandarización, extrayendo información de los enunciados para plantear lo que se pide para su posterior resolución.
		3	A los ítems A y B del Paso 5 se incorpora el ítem C que se subdivide en tres preguntas a), b) y c).	Calcular probabilidades estandarizando, utilizando los intervalos icónicos y desestandarizando

Fuente: Elaboración Propia

Tabla 6: Subactividades complementarias para la Actividad 1 de la propuesta ministerial.

Ubicación en la Actividad 1	Título Subactividad	Descripción de la Subactividad
Entre el Paso 2 y el Paso 3	Identificando los parámetros de una Distribución Normal	La subactividad se divide en tres ítems A, B y C. El ítem A permite identificar cuando un parámetro es mayor, menor o igual que otro en una gráfica, a partir de los colores señalados para cada campana. Además, comprenden de qué forma influye el valor del parámetro en la gráfica de la Distribución Normal. Este ítem tiene preguntas desde la a) hasta la f). En el ítem B se debe aplicar la identificación de parámetros en tres problemas contextualizados a), b) y c), donde no solo deben comparar los valores que dan forma a las campanas, sino que también debe extraer

		información del enunciado para plantear su resolución y tomar decisiones a partir de su respuesta. El ítem C permite descubrir qué pasa al variar el valor de los parámetros en la campana de Gauss, y como el valor de estos va a afectar en la forma y escala que adopta la campana. Este ítem tiene preguntas desde la a) hasta la g).
	Un teorema bastante utilizado	Se presentan seis preguntas de la a) hasta la f) para ejercitar el cálculo de probabilidades con la tabla Z. Se enfatiza en la necesidad de resolver diferentes expresiones de probabilidad, considerando que los estudiantes de tercero medio (y algunos de cuarto) probablemente no conozcan la Distribución Normal, ni su estandarización.
	Descubriendo una nueva propiedad	A partir de ocho preguntas de la a) hasta la h) se induce la propiedad simétrica de la Distribución Normal, por medio de la observación de los valores de la tabla y de las gráficas.
Entre el Paso 3 y el Paso 4	Desestandarizando una variable aleatoria Normal	Presenta dos ítems A y B, donde ambos se subdividen desde la a) hasta la f) para resolver ejercicios que involucran el cálculo de probabilidades y cuantiles de la Distribución Normal a través del uso de la tabla Z en forma inversa.

Fuente: Elaboración Propia

Mientras se incorporaron las subactividades, estas y los Pasos de la Actividad 1 tuvieron que ser reorganizados en correspondencia al orden de los contenidos presentes en la propuesta ministerial, por dicho motivo en la propuesta pedagógica las subactividades fueron designadas a un número de Paso específico. Finalmente, se obtuvieron nueve Pasos para la Actividad 1, cuya nueva distribución se aprecia en la Tabla 7.

Tabla 7: Organización de los Pasos para la Actividad 1.

N°	Nombre del Paso	Origen
Paso 1	El concepto de una variable aleatoria continua	Propuesta ministerial
Paso 2	Aproximándonos a la Distribución Normal	Propuesta ministerial
Paso 3	Identificando los parámetros de una Distribución Normal	Subactividad incorporada

Paso 4	Un teorema bastante utilizado	Subactividad incorporada
Paso 5	Descubriendo una nueva propiedad	Subactividad incorporada
Paso 6	Significado de la Distribución Normal	Propuesta ministerial
Paso 7	Desestandarizando una variable	Subactividad incorporada
Paso 8	Aproximación de una Distribución Binomial por la Distribución Normal Estandarizada	Propuesta ministerial
Paso 9	Aplicación de la Distribución Normal	Propuesta ministerial

Fuente: Elaboración Propia

3.3.1.2. Plan de Acción y Fase de Acción de la Actividad 2.

En la Actividad 2, *Aplicar el modelo Normal en el transporte de personas*, se aplican conceptos que se involucran y relacionan con la Distribución Normal a través de dos subactividades que se centran en un mismo problema de la vida cotidiana, cuyo enunciado es descrito al inicio de la Actividad. El plan de acción se describe en la Tabla 8.

Tabla 8: Plan de acción Actividad 2 de la propuesta ministerial.

N° sub-actividad	Título subactividad	Descripción de la subactividad	Plan de Acción
1	Transporte en teleférico	Vincula el transporte de la ciudad de Santiago con un problema de carga máxima que puede soportar, en ella los estudiantes determinan la probabilidad de solo un dato a través de la Distribución Normal Estándar. La subactividad tiene seis preguntas, enumeradas del 1) al 6). Cabe señalar que la pregunta 4 se subdivide en a), b) y c).	➢ En el enunciado se agrega el valor de $\sigma = 12{,}9\ kg$ para resolver las preguntas que requieran conocer el valor de la desviación estándar. ➢ En 4.a cambiar la pregunta "*¿Por qué es necesario convertir la masa en su puntuación z correspondiente para buscar la probabilidad buscada? Señalen los valores*" por "*Calcule la probabilidad pedida haciendo uso de la tabla Z*".
2	Determinar probabilidades que impliquen el uso del Teorema del Límite Central	Bajo el mismo contexto del teleférico, aborda el concepto de Teorema Central del Límite (TCL) utilizando Excel, permitiendo a los estudiantes conjeturar resultados a partir de esta	➢ Corregir el error de tipeo en la pregunta 1 cambiando el 6 de "77,6" por un 3. ➢ Corregir error de tipeo en la pregunta 7 eliminando la palabra "total" y el dígito 1 de "180".

	herramienta digital. Esta subactividad tiene siete preguntas que se organizan de la siguiente forma: -Las preguntas 1 y 4 no tienen subdivisiones. -Las preguntas 2 y 6 se subdividen en a) y b). -Las preguntas 3, 5 y 7 se subdividen en a), b) y c).	➢ Para complementar y sacar provecho al uso de la herramienta de datos de Excel, se incorpora una subactividad que incluya la simulación de una muestra que no provenga de una Distribución Normal y evidenciar que, de todas formas, en promedio tiende a una Distribución Normal por medio del TCL.

Fuente: Elaboración Propia

Tal como señala el plan de Acción, en la Fase de Acción se elaboró una subactividad que al igual que la segunda subactividad de la Actividad 2 usa Excel para tratar el Teorema Central del Límite, por ello se decidió incorporar la nueva subactividad como pregunta 8 en continuación a la de la propuesta ministerial, sin embargo, bajo su naturaleza de complemento deja a un lado el contexto del teleférico para recurrir a otra problemática.

La subactividad elaborada por los investigadores se subdivide en puntos desde la a) hasta la g), cuyos objetivos son los siguientes:

- Comprender que el Teorema del Límite Central también se cumple cuando la población de origen no distribuye Normal.
- Aplicar el Teorema Central del Límite a partir de un estudio de caso que distribuye Binomial utilizando la herramienta digital Excel.

3.3.1.3. Plan de Acción y Fase de Acción de la Actividad 3.

La última actividad de aprendizaje de la propuesta ministerial se titula *Aproximar la Distribución Binomial por la Distribución Normal,* y se compone por dos subactividades, el primero llamado *aproximando una Distribución Binomial mediante una Distribución Normal*, y el segundo *cálculo de probabilidades mediante la aproximación Normal de una Distribución Binomial.* La descripción de estas y el detalle del plan de acción se aprecia en la Tabla 9.

Tabla 9: Plan de acción Actividad 3 de la propuesta ministerial.

N° sub-actividad	Título subactividad	Descripción y organización de la subactividad	Plan de Acción
1	Aproximando una Distribución Binomial mediante una Distribución Normal	Parte desde lo más intuitivo de la aproximación, apoyando el proceso con el software matemático GeoGebra, para identificar el comportamiento de las gráficas a medida que la cantidad de repeticiones de un experimento dicotómico crece. Esta subactividad tiene cinco preguntas que se organizan de la siguiente forma: -Las preguntas 1 y 2 se subdividen en a) y b). -Las preguntas 3, 4 y 5 se subdividen en a), b) y c).	➢ En el enunciado de la pregunta 2 añadir el dato de la independencia de las variables. ➢ En la pregunta 3.a) agregar los títulos a los ejes del gráfico.
2	Cálculo de probabilidades mediante la aproximación Normal de una Distribución Binomial	Hace la relación entre los parámetros de la Distribución Binomial con la Normal, además de pedirles a los estudiantes que conjeturen una regla para aplicar la aproximación y la corrección por continuidad. Esta subactividad tiene cinco preguntas que se organizan de la siguiente forma: -La pregunta 1 se subdivide en a), b) y c). -La pregunta 2 se subdividen en a) y b). -La pregunta 3 y 4 en a), b), c) d) y e). -La pregunta 5 no tiene subdivisiones.	➢ Incorporar preguntas desde 1.b) cuya finalidad sea determinar la relación entre los parámetros de la Distribución Binomial y Normal a partir de la visualización de gráficas en GeoGebra. ➢ En las preguntas 2.b), 4.a) y 5 corregir el error de notación, pues la propuesta ministerial usa x en vez de X para la variable aleatoria continua.

Fuente: Elaboración Propia

Durante la Fase de Acción se incorporan tres preguntas elaboradas por los investigadores para complementar la subactividad *Cálculo de probabilidades mediante la aproximación Normal de una Distribución Binomial* de la propuesta ministerial.

El objetivo de estas preguntas permite determinar la relación entre los parámetros de la Distribución Binomial y Normal a partir de la visualización de gráficas en el software GeoGebra.

3.3.1.4. Plan de Acción y Fase de Acción de la Actividad 4.

Se culmina con la Actividad de Evaluación, que reúne todos los conceptos aplicados en el transcurso de la Unidad 3 para realizar una síntesis de los contenidos aprendidos por los estudiantes, esta se compone por tres subactividades que se organiza de la siguiente forma:

I. *Identificar el sentido de contar con datos distribuidos mediante una Distribución Binomial o una Normal*. Esta subactividad posee cinco preguntas enumeradas del 1 al 5, sin embargo, solo desde la cuarta se trata la Distribución Normal, de esta forma la pregunta 4 está subdividida en a), b) c), d y e) y la pregunta 5 en a), b) y c).

II. *Modelar situaciones con una Distribución Binomial y una Distribución Normal*. Esta subactividad tiene tres preguntas enumeradas, la primera sin subdivisiones, la segunda de la a) hasta la g) y la tercera de la a) hasta la d).

III. *Modelar situaciones, usando la Distribución Normal y el Teorema del Límite Central, y aproximar una Distribución Binomial mediante una Normal*. Esta subactividad tiene tres preguntas enumeradas, la primera de ellas se subdivide de la a) hasta la f), la segunda pregunta de la a) hasta la e) y la última se subdivide en a), b) y c).

La Actividad 4 es la que menos acción requirió, particularmente solo se procedió a permutar las preguntas 3.b y 3.c de la subactividad III, con la finalidad de tener mayor coherencia en el paso a paso que los estudiantes realizan al resolver la pregunta.

3.3.2. Fase de Observación.

Por medio de esta fase se registran situaciones importantes de la Fase de Acción, obteniendo datos que permitirán más adelante reflexionar y evaluar las acciones que se llevaron a cabo para abordar las problemáticas observadas. En consecuencia, para efectos de validar las preguntas y subactividades elaboradas por los investigadores, estas serán sometidas al juicio de expertos proporcionando información valiosa para la posterior reflexión.

3.3.3. Fase de Reflexión.

La Fase de Reflexión se efectúa durante todo el proceso de la investigación, tal como señala Latorre (2015) “no es una fase aislada en el tiempo, ni algo que ocurre al final de la investigación, sino una tarea que se realiza mientras persiste el estudio” (p.82). Sin embargo, se hace énfasis en los hechos observados en la Fase de Acción interpretando y analizando tanto la resolución de las actividades, como las observaciones de los expertos en relación a las preguntas y subactividades propuestas por los investigadores.

3.4. Instrumentos para Recolección de Información

Para la investigación-acción se utilizaron dos instrumentos que permitieron recoger información para la Fase de Planificación y la Fase de Observación del modelo utilizado. Para tal caso, ambas técnicas de recolección se construyeron a partir de Escalas Estimativas.

La Escala de Estimación, es un instrumento que se emplea con objeto de lograr cierta igualdad en la recopilación e interpretación de datos. Estas observaciones sirven para evaluar las conductas, productos, procesos o procedimientos de cierto individuo o objeto en estudio. Para ello se elaboran una serie de indicadores o descriptores que son evaluados marcando el grado de intensidad en el cual la característica o cualidad está presente. Así, los niveles de desempeño de este instrumento permiten discriminar con un mayor grado de precisión el comportamiento a observar o el contenido a medir, dado que pueden ser expresados en una escala verbal, numérica, gráfica o descriptiva (Pimienta, 2008).

Entonces, su construcción comienza con la recopilación o elaboración de un listado de enunciados que expresan diversas actitudes positivas y negativas hacia un objeto o acontecimiento específico, en otras palabras, estos son indicadores que representan las variables que el investigador está interesado en medir por medio de categorías que expresan el nivel de logro de cada una de ellas (Lewis, 2003).

3.4.1. Instrumento 1.

El propósito del primer instrumento consiste en aportar información útil para la Fase de Planificación y a partir de su aplicación tomar decisiones para el plan de acción, determinando cuales son los indicadores de evaluación que se presentan en menor proporción en el Programa de Estudio de Probabilidades y Estadísticas Descriptiva e Inferencial asociados al contenido de la enseñanza de la Distribución Normal.

Para su construcción se seleccionó los descriptores que serán evaluados en la escala estimativa, aquellos se elaboraron considerando ocho indicadores de evaluación presentes en el Programa de Estudio para la enseñanza de la Distribución Normal, porque el logro de estos indicadores demuestra que un estudiante ha aprendido y transitando desde lo más elemental hasta lo más complejo, aludiendo a procesos cognitivos, estimulando el pensamiento lógico y al procesamiento y organización de la información (MINEDUC, 2021b).

Además, en coherencia con los estándares de aprendizaje del MINEDUC y el conocimiento de recursos digitales se complementó el listado incorporando al final de él dos elementos fundamentales para la enseñanza y aprendizaje del objeto matemático. Efectivamente, un indicador implica el uso de herramientas digitales que favorecen el proceso de enseñanza de la Distribución Normal, a su vez, se agregó el indicador que involucra la comprensión de la distribución de las medias muestrales, es decir, del Teorema del Límite Central. Cabe destacar que, utilizar los indicadores de evaluación distinguidos en la propuesta ministerial para los descriptores de la escala estimativa infieren una validación a juicio de expertos, no obstante, la naturaleza del instrumento no requiere necesariamente del recurso de los jueces (Lewis, 2013).

Tabla 10: Lista de Indicadores utilizados en el Instrumento 1.

Nº Indicador	**Descripción del Indicador**
1	La actividad permite identificar las principales características de una Distribución Normal, tales como la varianza o desviación estándar y la media.
2	La actividad permite interpretar información que involucra la Distribución Normal y sus parámetros.

3	En la actividad se resuelven problemas que involucran la Distribución Normal.
4	Por medio de la actividad los estudiantes determinan la probabilidad de intervalos en distribuciones normales, estandarizando y utilizando la tabla Z.
5	La actividad propicia la toma de decisiones, basándose en la Distribución Normal, y en el cálculo de probabilidades.
6	Por medio de la actividad los estudiantes argumentan cuando se puede modelar un fenómeno con una Distribución Normal.
7	En la actividad se pueden modelar fenómenos o situaciones cotidianas, científicas y sociales mediante la Distribución Normal.
8	En la actividad se pueden modelar situaciones que involucren la aproximación de la Binomial a la Normal.
9	La actividad contribuye efectivamente al aprendizaje de la Distribución Normal por medio de herramientas digitales.
10	La actividad contribuye a la comprensión de la distribución de las medias muestrales (Teorema Central del Límite)

Fuente: Elaboración propia a partir de MINEDUC (2021b).

Por otra parte, el grado de concordancia entre el indicador de cada enunciado y la opinión de quienes utilicen este instrumento está en una escala de tres puntos, la que contiene un polo negativo, otro positivo y una opción neutra, dicho de otra forma, el investigador tiene que justipreciar en qué medida se aprecia cada indicador en las actividades para la enseñanza de la Distribución Normal propuestas por el MINEDUC, clasificando su grado de conformidad en: (1) *totalmente en desacuerdo*, (2) *parcialmente de acuerdo* o (3) *totalmente de acuerdo*. Adicionalmente, se añadió la opción (4) *no aplica*, ya que cada actividad del Programa de Estudio tiene sus propios objetivos específicos de aprendizaje, en cuyo caso, se hace imposible evaluar aquellos indicadores que están en función de objetivos distintos.

Adicionalmente, cada descriptor de la escala apreciativa incluyó un apartado para justificar las actividades cuyos indicadores fueron clasificados en las casillas parcialmente de acuerdo o totalmente en desacuerdo con el propósito de especificar el o los criterios observados para tomar dicha decisión.

Por último, considerando la naturaleza cualitativa del instrumento, junto a las apreciaciones subjetivas de los investigadores y de la profesora Giovanna Ticchione en torno a la lectura y revisión del programa de estudio, se realiza una tabla de observaciones y sugerencias para las actividades en la propuesta ministerial (ver Anexo 2).

Los resultados obtenidos al aplicar el instrumento 1 pueden observarse en el apartado Anexo 1, sin embargo, en la Figura 7 se aprecia un esbozo del instrumento.

N°	**Indicador**	**Totalmente en desacuerdo**				**Parcialmente de acuerdo**				**Totalmente de acuerdo**				**No Aplica**				**Justifica**
		A1	A2	A3	A4	A1	A2	A3	A4	A1	A2	A3	A4	A1	A2	A3	A4	
1																		
2																		
3																		

Figura 7: Bosquejo del instrumento 1. Elaboración Propia.

3.4.2. Instrumento 2.

Para su elaboración se consideró dos aspectos que son fundamentales para los objetivos de esta investigación, es decir, obtener la percepción y juicio de los expertos acerca de las actividades de la propuesta ministerial, pero principalmente sobre las subactividades elaboradas por los investigadores. Por lo tanto, el objetivo de este instrumento consiste en validar las subactividades y preguntas complementarias elaboradas por los investigadores.

Se confecciona una escala apreciativa, la cual busca conocer el juicio profesional de profesores de matemática en ejercicio, respecto a la propuesta de actividades incorporadas a la Unidad de aprendizaje para la enseñanza de la Distribución Normal, para ello se listó un total de cuatro descriptores de valoración, cuyo grado de apreciación de los descriptores están dados por las categorías: *de acuerdo*, *ni en acuerdo ni es desacuerdo* y *en desacuerdo*.

Luego, se añaden tres preguntas abiertas, donde dos de ellas se realizan para que los jueces expresen su apreciación tanto positiva como negativa acerca de las subactividades y preguntas hechas por los autores de esta investigación. La tercera pregunta es para obtener su opinión crítica sobre las actividades de la propuesta ministerial.

Dado el contexto sanitario por Covid-19 mientras se realizó esta investigación, el segundo instrumento se aplicó a través de un formato digital, utilizando la plataforma Google Forms (ver Anexo 3).

Por último, cabe recalcar que la información aportada por los expertos contribuye a la fase de reflexión del modelo de investigación-acción.

3.4.3. Validación de las Actividades Complementarias.

Las subactividades elaboradas por los investigadores, con la finalidad de complementar las ya propuestas por el MINEDUC, fueron validadas por tres profesores de matemática quienes emitieron su juicio de experto. Además, uno de ellos ya ha realizado docencia en el curso del plan de formación diferenciada de Probabilidades y Estadísticas Descriptiva e Inferencial.

Por otro lado, es imprescindible mencionar que dos de los entrevistados llenaron el formulario de Google Forms, sin embargo, el otro solo entregó sus apreciaciones sobre las subactividades y preguntas por medio de un documento Word.

Capítulo IV. Resultados

Para efectos del ciclo del modelo investigación-acción en la fase de planificación, a partir del uso del instrumento 1, se evaluaron las actividades de la propuesta ministerial para determinar el plan de acción a seguir, sin embargo, en esta etapa no se detallan los resultados obtenidos en la aplicación del instrumento ni en la revisión del Programa de Estudio, por ello estos serán abordados en este capítulo.

Al contrastar los descriptores del primer instrumento con las actividades de la propuesta ministerial se obtuvo que, entre las cuatro actividades revisadas, todas tuvieron cierto grado de correspondencia positiva. A pesar de ello, la Actividad 1, *Comprender el modelo Normal de probabilidades*, presentó menor cobertura y en consecuencia es la que requiere mayor atención. Esta decisión se fundamenta porque, se estuvo parcialmente de acuerdo con el nivel de logro de dos de los indicadores del instrumento 1. Singularmente respecto al indicador N°1, en la actividad no se apreció la comprensión de la desviación estándar, ni se indicó que característica aporta a la campana de Gauss, igualmente tampoco se enseña a utilizar la tabla Z. En relación con el indicador N°2 se observó que, si bien se interpreta la información que es necesaria para realizar los cálculos pertinentes al contenido, no se pide la interpretación de las probabilidades determinadas en dicha actividad, ni la de los parámetros de la Distribución Normal. Al mismo tiempo se cree que para potenciar la actividad se debe incorporar herramientas digitales, como GeoGebra, para que los estudiantes puedan visualizar y comprender las propiedades de la Distribución Normal.

De los ocho indicadores del instrumento 1 asociados a la Actividad 2, *Aplicar el modelo Normal en el transporte de personas*, se estuvo de acuerdo con siete de ellos, pues el indicador N° 9 fue cubierto parcialmente, su justificación radica en el uso de las herramientas digitales, pues fueron utilizadas para comprender la distribución de medias muestrales, más no para la Distribución Normal como tal. Cabe destacar que el uso de este recurso influyó positivamente en la pregunta 3 de la segunda parte de la Actividad 2, ya que efectivamente permitió determinar y observar el comportamiento de las medias muestrales. En general, a partir de la revisión expresamos que la actividad no presenta errores de notación, no obstante, la segunda parte de esta actividad, *Determinar probabilidades que* *impliquen el uso del Teorema del Límite Central,* presenta un error de precisión conceptual pues no se cumple la condición $n > 30$ para realizar la estandarización. De igual forma se encuentran dos pequeños errores de tipeo.

La Actividad 3, *Aproximar la Distribución Binomial por la Distribución Normal,* a diferencia de sus antecesoras cumple totalmente con los nueve indicadores del instrumento 1 vinculadas a su desarrollo, no obstante, es importante señalar que en su revisión se observa un error de tipeo, pues no se entrega el valor de σ donde corresponde. Además, conceptualmente pasaron por alto, señalar la independencia de las variables para poder utilizar el TLC y declarar x en vez de X para referirse a la variable aleatoria continua. Por último, podemos decir, que la actividad destaca por ser deductiva, por ejemplo, se aprecia cómo a través de una serie de preguntas los estudiantes deducen la regla de la corrección de continuidad en vez de introducir directamente la fórmula.

Respecto a la Actividad 4, *Actividad de Evaluación,* cumple en totalidad su objetivo de recorrer todos los contenidos tratados en la Unidad 3, ya que las preguntas que la componen comprenden nueve indicadores asociados a la enseñanza de la Distribución Normal señalados en el instrumento 1, excluyendo el indicador N°9 pues según lo observado la naturaleza de esta actividad implica ser utilizada en un aula tradicional.

Por lo tanto, producto de la revisión crítica realizada sobre la propuesta ministerial resulta que las Actividad 2 y 3 están bastante completas, porque ambas inician y desenvuelven su desarrollo a partir de una problemática real que permite contextualizar la aplicación del conocimiento respecto al objeto matemático, de esta forma se propicia un aprendizaje significativo, además las preguntas están diseñadas de forma tal que conducen el razonamiento deductivo, sin embargo, ambas actividades poseen pequeños errores de tipeo y precisión conceptual. Por su parte, a diferencia de las Actividades 2 y 3, la Actividad 1 se compone por una serie de pasos que parecen aisladas una de las otras, es decir carecen de cohesión, presentando también errores de precisión conceptual y en los gráficos adjuntos.

El instrumento 2 es relevante pues cumple un papel fundamental para la fase de observación del ciclo investigación-acción, ya que no solo válida la propuesta de los investigadores, sino también recoge las impresiones negativas y positivas de los jueces en relación con él, es decir, actúan como agentes evaluadores del resultado de la aplicación de la fase de acción.

Una vez que se aplica el instrumento 2, cuyos resultados se detallan en Anexo 4, se recolectan las respuestas de tres docentes, quienes para efectos de la investigación son considerados *expertos*, de ahí que a continuación se denominan *experto 1, experto 2* y *experto 3*. Cabe mencionar que solo los *expertos 1* y *2* respondieron el formulario de Google, por su lado el *experto 3* hace entrega de sus respuestas en un documento Word aludiendo solo a las preguntas abiertas sobre su apreciación crítica frente a las subactividades y preguntas propuestas por los investigadores.

De los tres, solamente el *experto 2* ha realizado docencia del electivo y declara haber utilizado las actividades propuestas del Programa de Estudio, sin embargo, no alcanzó a abordar el programa completo, afirmando que "los estudiantes no desarrollaron actividades de Distribución Normal presentes en el programa ministerial".

Respecto a las respuestas otorgadas en las preguntas cerradas sobre la apreciación de los expertos en función del grado de correspondencia que atiende a las afirmaciones allí presentadas, se tiene que, el *experto 1* está de acuerdo con todas ellas, por el contrario, el *experto 2* indica estar ni en acuerdo ni en desacuerdo con las afirmaciones 1, 3 y 4, en tanto, con la 2 está en desacuerdo. El *experto* 3 no respondió estas preguntas.

En relación con el apartado de preguntas abiertas, específicamente la de aspectos positivos acerca de la propuesta elaborada por los investigadores, se encuentran las siguientes respuestas:

El *experto 1* indica que la propuesta de los investigadores "tienen un enfoque contextualizado y cercano a lo cotidiano, permitiendo dar un uso real a los conocimientos adquiridos sobre cada parámetro estudiado, facilitando así la comprensión y aplicación de estos". Referido a los recursos digitales utilizados en la propuesta, él considera que contribuyen "a procesos indagatorios en el proceso de enseñanza-aprendizaje a través de la comparativa", además, estos ayudan "a que los estudiantes valoren el uso de tecnologías en estos procesos probabilísticos". Concluye mencionando que "las orientaciones al docente son clarificadoras y acordes a lo que las actividades permiten lograr".

Asimismo, el *experto 2* señala que “las actividades y preguntas tenían por objetivo que el estudiante pudiese interpretar los resultados obtenidos y/o presentados, más que simplemente calcular alguna medida en particular”. Por último, recalca en la importancia que tiene el “uso de GeoGebra para estudiar el comportamiento de la media en distribuciones simétricas y asimétricas”

Finalmente, el *experto 3* menciona que las actividades “sin duda que son un aporte”, agregando que “la idea es plantear diversas situaciones, ojalá relacionadas con otra ciencia o situación cotidiana, de forma que se trabaje la habilidad de resolución de problemas (aunque el problema del teleférico trabaja esta habilidad)”. Concluye que "está muy presente la habilidad de argumentar en todas las actividades”, lo que desde luego contribuye a la propuesta ministerial.

En relación con los aspectos negativos, el *experto 1* y *experto 3* indican que “la mayoría de las actividades que se plantean, de alguna forma, mecanizan el aprendizaje” agregando que falta “una introducción de carácter motivante y atractiva para el estudiante”. El *experto 3* adiciona que las actividades “en su mayoría son ejercicios que tratan sobre resolver problemas rutinarios, similares unos de otros, donde con unos pocos de ellos bastaría, sin necesidad de revisar todos los casos (…). Se quedan en una habilidad básica, sin avanzar dentro de la misma habilidad con problemas de mayor complejidad”, por lo que propone “presentar diversas situaciones que se vinculen al contenido permitiendo a los estudiantes enfrentarse a distintos escenarios donde no es tan inmediato aplicar el contenido, de modo que estos tendrán que reflexionar un poco para llegar a la respuesta”. También, corrige un error de tipeo y sugiere cambiar la precisión conceptual que observó en el documento entregado a los validadores. Cabe destacar que solo él advirtió de los desaciertos. Por su parte, el *experto 2* expone que “el planteamiento de algunas preguntas era poco claro y carecía de coherencia”.

El instrumento 2 culminó con una pregunta final que recogía las apreciaciones críticas acerca de la propuesta ministerial, teniéndose que:

El *experto 1* considera que la “propuesta curricular es bastante sólida, ya que abarca puntualidades del contenido de probabilidad y estadística descriptiva e inferencial de forma efectiva, haciendo uso de ejemplos, contextualizaciones, recursos digitales en cada concepto a trabajar, de manera que el objetivo de aprendizaje a lograr se va construyendo clase a clase de manera coherente entre estas, dando lugar a una linealidad conceptual bastante clara". Además, destaca la "conexión interdisciplinaria con ciencias para la ciudadanía en las actividades", pues "permite a los estudiantes valorar el alcance de lo que están aprendiendo". En último lugar, enjuicia al instrumento de validación pues este "apunta a factores demasiado generales sobre la propuesta, impidiendo una valoración más técnica sobre otros factores importantes tales como; tiempos destinados a cada actividad calzando con un tiempo realista de las clases, también factores como estrategias didácticas específicas a utilizar en alguna clase o actividad".

La última respuesta corresponde al *experto 2*, quien indica que el programa "ayuda a orientar los objetivos y metodología del electivo, con preguntas que buscan el análisis, la reflexión e interpretación de resultados obtenidos en contextos reales. Vincula de buena forma los contenidos con el uso de herramientas digitales". Sin embargo, menciona algunos arreglos que debieran considerarse, como "la claridad en la formulación de algunas preguntas y que las orientaciones al docente sean más específicas en base a lo que se busca en cada pregunta", sumado a que "algunos links propuestos en el programa estaban caducados (recuerdo el link para construir diagramas de cajas)".

Considerando cierto grado de flexibilidad y adaptabilidad con la que se planifica, junto a los comentarios negativos del *experto 3* se decidió añadir al Paso 3 de la Actividad 1, *Identificando los parámetros de la Normal*, un ítem B que incluye tres situaciones problemáticas cuyo propósito define que:

- Los estudiantes identifican los parámetros de la Distribución Normal, con ellos deben comparar los valores que dan forma a las campanas, de tal forma que a partir de sus resultados y la información entregada en cada enunciado tomen decisiones fundamentadas.

Finalmente, a causa de lo descrito anteriormente, resulta la propuesta de subactividades y preguntas que tienen como propósito apoyar el proceso de aprendizaje de los estudiantes, complementando a las actividades del Programa de Estudio, además de ayudar al quehacer docente en los contenidos relacionados con la Distribución Normal del curso de Probabilidades y Estadísticas Descriptiva e Inferencial para los niveles de III° y IV° medio.

La propuesta se compone de:

- El objetivo de la subactividad, que justifica la finalidad y utilidad de esta.
- Las preguntas de la subactividad, que complementan a la propuesta ministerial.
- Las notas y orientaciones docentes, que son sugerencias para el oportuno desarrollo de las subactividades y preguntas.

A continuación, se presentan solo los enunciados de las subactividades y preguntas. La solución y cohesión de ellas junto a las actividades de la propuesta ministerial se pueden apreciar en el Anexo 5. Es preciso indicar que las preguntas complementarias del Paso 9 de la Actividad 1, desprenden de los problemas planteado en los ítems A y B, por este motivo para contextualizarlas se incorpora el enunciado de A y B que son parte de la propuesta ministerial.

ACTIVIDAD 1: COMPRENDER EL MODELO NORMAL DE PROBABILIDADES

DESARROLLO Y SIGNIFICADO DE LA DISTRIBUCIÓN NORMAL

PASO 1: EL CONCEPTO DE UNA VARIABLE ALEATORIA CONTINUA

d. Lee la siguiente situación y luego responde:

El arroz es uno de los principales cereales usados en la alimentación humana en el mundo. Las distintas características que poseen los granos de arroz se asocian al hecho de que los consumidores de distintos países demanden tipos de arroz diferentes en términos de variedad, calidad y aspecto. Es así, como estas industrias deben considerar ciertos estándares de calidad, uno de ellos se relaciona con la clasificación del arroz de acuerdo con sus proporciones, ya sea según su longitud o ancho, como también por su masa. Por una parte, la masa de cada granito de arroz está entre $26{,}1\ mg$ y $30\ mg$. Además, en Chile, país el arroz de grano largo es aquel que posee una longitud media superior a $6{,}5mm$.

Considerando la información anterior, supongan que tienen un kilo de arroz de grano largo y seleccionan al azar un granito de arroz del paquete.

i. ¿Qué variables aleatorias podrías definir a partir del enunciado anterior? ¿Qué valores pueden tomar estas variables aleatorias?

ii. ¿Cuál es la probabilidad de encontrar un granito de arroz con una masa exacta de $28,\overline{3}\ mg$? Argumenten su respuesta.

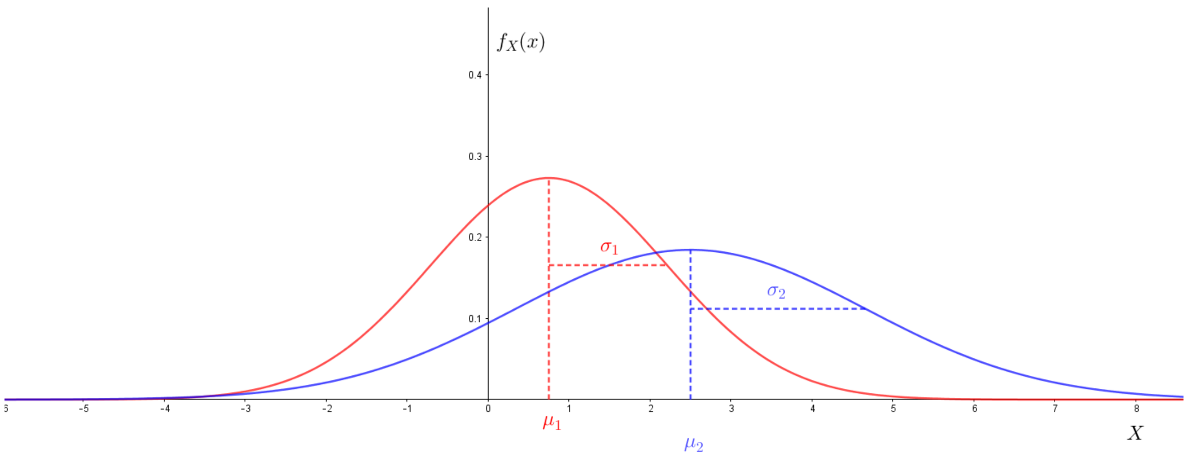
iii. ¿Cuál es la probabilidad de encontrar un granito de arroz de longitud $6{,}777778\ mm$? Argumenten su respuesta.

PASO 3. IDENTIFICANDO LOS PARÁMETROS DE UNA DISTRIBUCIÓN NORMAL

La media μ y varianza σ^2 son los dos parámetros de los que depende una distribución Normal, y estos condicionan la ubicación y forma de la campana de Gauss.

A. Analiza las siguientes gráficas, y reconoce la relación (<, > o =) que existe entre la media y desviación estándar de ambas. Para ello considera que μ_1 y σ_1 es la media y desviación estándar de la campana de color azul y que μ_2 y σ_2 son la media y desviación estándar de la campana de color rojo.

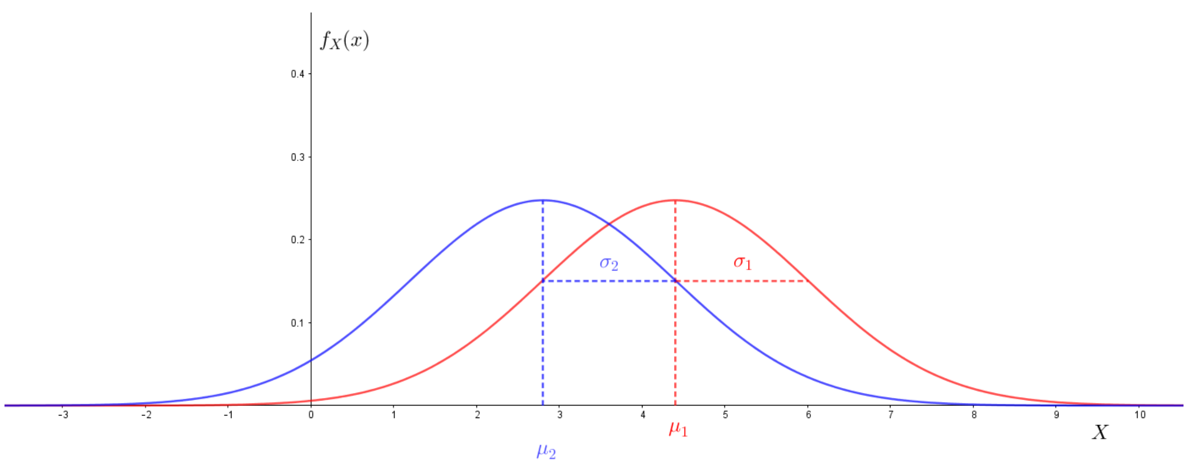
a.



μ_1 ___ μ_2

σ_1^2 ___ σ_2^2

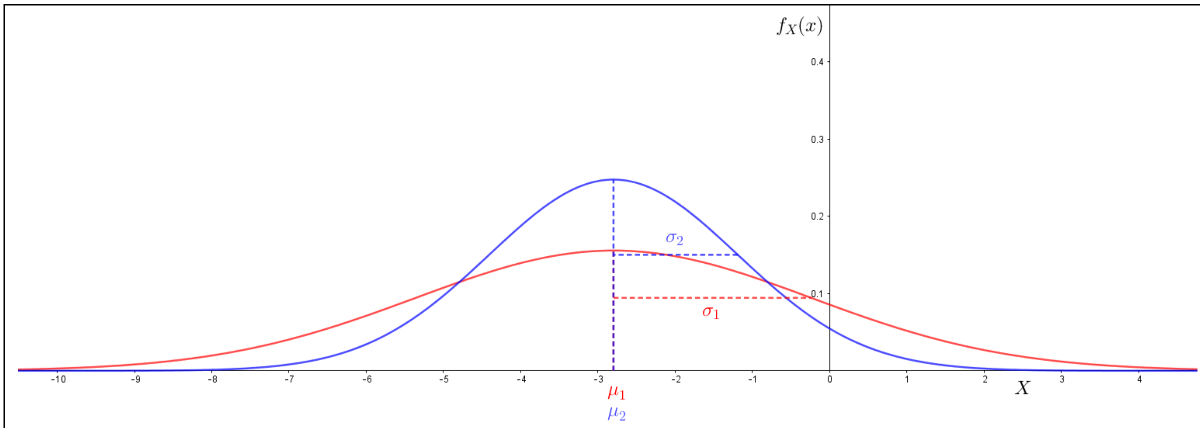
b.



μ_1 ___ μ_2

σ_1^2 ___ σ_2^2

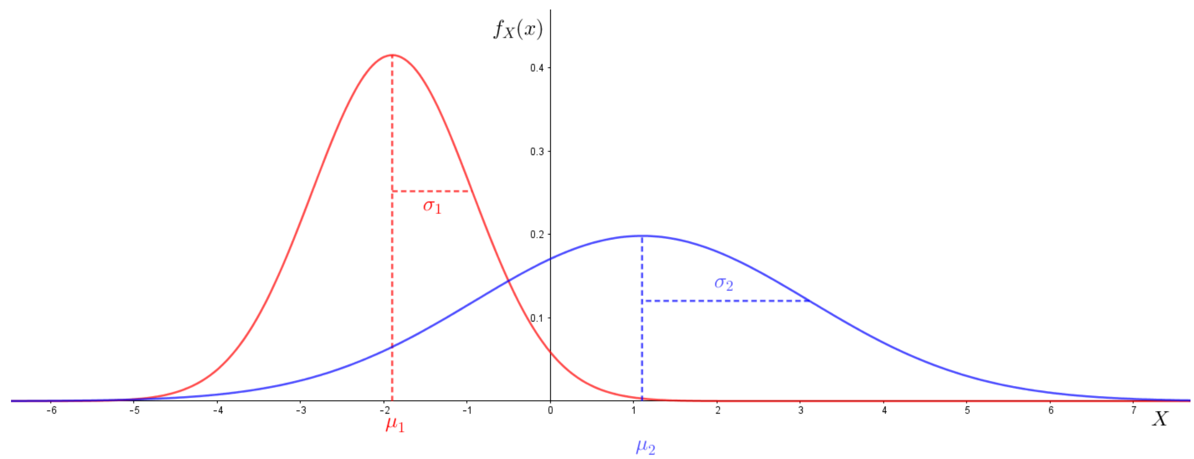
c.



μ_1 ___ μ_2

σ_1^2 ___ σ_2^2

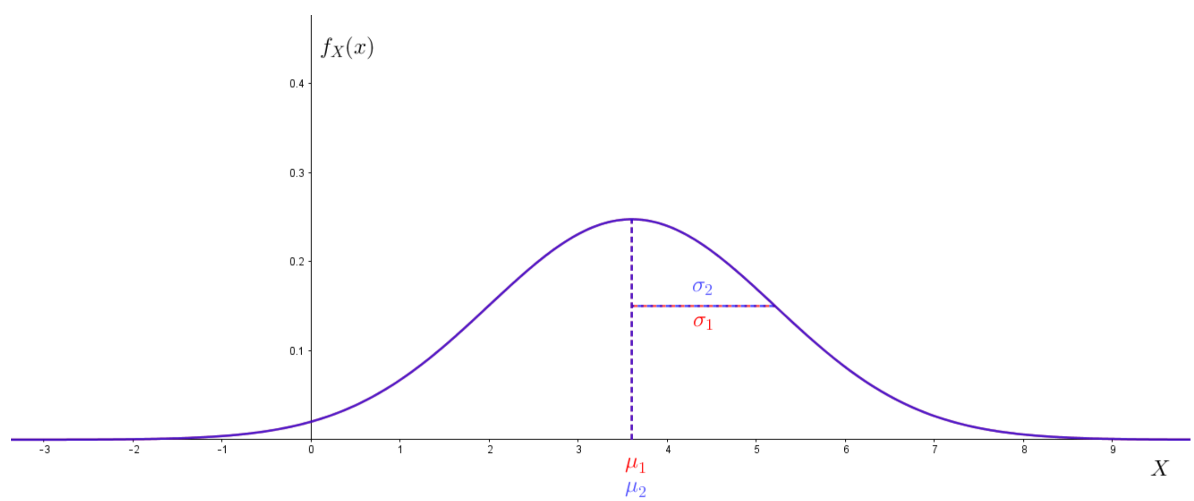
d.



μ_1 ___ μ_2

σ_1^2 ___ σ_2^2

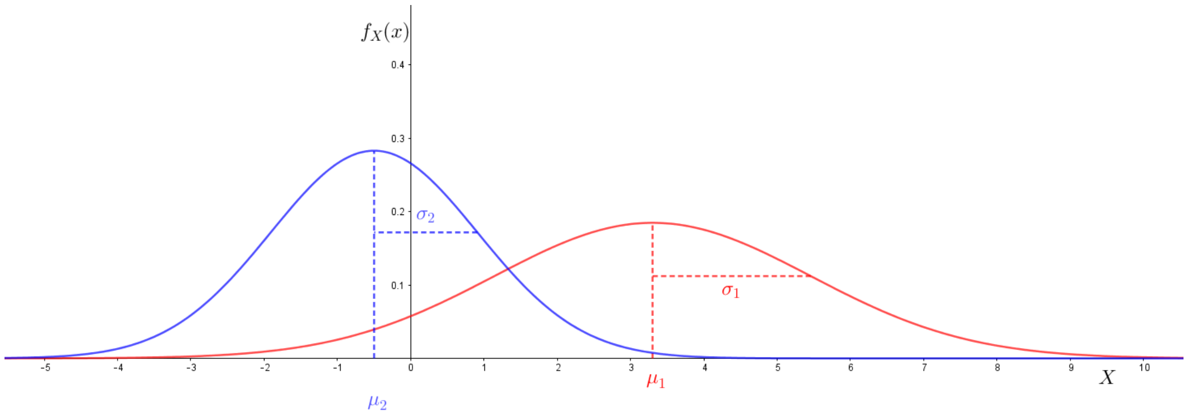
e.



μ_1 ___ μ_2

σ_1^2 ___ σ_2^2

f.



μ_1 ___ μ_2

σ_1^2 ___ σ_2^2

B. Resuelve los siguientes problemas.

a. El animal más alto del mundo es la jirafa, vive en sabanas, pastizales y bosques abiertos repartidos en casi todo el continente africano y su dieta alimenticia se compone principalmente de las hojas de la acacia. La subespecie denominada *jirafa de Sudáfrica* habita principalmente en esa región, también vive en Botsuana, Zimbabue y Mozambique. Su altura sigue una Distribución Normal de parámetros $\mu = 5{,}22\ m$ y $\sigma = 0{,}2\ m$. Por otro lado, la *jirafa Masai* o *jirafa del Kilimanjaro* habita solamente en Kenia y Tanzania, su altura igualmente distribuye Normal, con media $\mu = 5{,}12\ m$, y desviación estándar $\sigma = 0{,}25\ m$.

   i. En tu cuaderno, esboza las campanas de ambas poblaciones en un mismo eje.

   ii. ¿Qué población de jirafas posee mayor variabilidad en su altura, respecto de la media?

   iii. En media, ¿qué población de jirafas tiene la menor altura?

b. Josefa y sus amigos desean comer pizza, quieren que sea de entrega rápida pues tienen bastante hambre. No tienen claro a qué pizzería pedir, sin embargo, cerca de su casa hay dos locales que realizan delivery el primero es "La pizzería de la esquina", tiene un tiempo de producción y entrega que sigue una Distribución Normal de parámetros $\mu = 57\ min$ y varianza de $\sigma^2 = 25\ min^2$, mientras que el tiempo del otro local, "Pizzamanía", también sigue una distribución Normal de parámetros $\mu = 54\ min$ y desviación estándar $\sigma = 5{,}92\ min$.

   i. En tu cuaderno, esboza ambas campanas en un mismo eje.

   ii. En media, ¿qué pizzería demora menos tiempo en realizar el delivery? ¿Qué elementos te permitieron identificarlo? Argumenta tu respuesta.

iii. ¿Qué pizzería tiene menor variabilidad respecto al tiempo medio de producción y entrega? ¿Cómo identificaste que era ese local? Argumenta tu respuesta.

iv. Considerando que la pizzería favorita de Josefa es la que demora más tiempo, porque elabora pizzas más ricas. ¿A qué local encargarías tú, si fueses amiga o amigo de Josefa? Argumenta tu respuesta.

c. Tres grupos de distintos géneros musicales realizan un concierto el mismo día en Santiago de Chile. Los tiempos de duración del evento de las tres bandas siguen una distribución Normal, de forma que, el tiempo de duración del grupo de rock tiene parámetros $\mu = 150\ min$ y desviación estándar de $\sigma = 10\ min$. Por su parte, el evento del grupo de pop posee una media $\mu = 146\ min$ y varianza de $\sigma^2 = 64\ min^2$. Por último, el tiempo de duración del grupo de cumbia distribuye Normal con parámetros $\mu = 152\ min$ y $\sigma^2 = 36\ min^2$.

i. Esboza las tres campanas en un mismo eje.

ii. Identifica qué grupo tiene el menor tiempo medio de duración del concierto.

iii. ¿Qué grupo tiene mayor variabilidad respecto al tiempo medio de duración del concierto?

iv. Supón que a Matías le gustan los tres grupos que se presentan y no tiene claro a cuál evento asistir, solo sabe que quiere ir al concierto que sea de mayor duración media. Indica a qué concierto le aconsejarías ir. Argumenta tu respuesta.

C. A partir de la información analizada en los ejercicios anteriores, y utilizando el enlace https://www.geogebra.org/m/fdvyufzw responde:

a. Si $\mu_1 = \mu_2$ ¿se puede concluir que las campanas tendrán el mismo eje de simetría? Argumenta tu respuesta.

b. Al variar el valor de μ, en cualquiera de las dos campanas, se produce un cambio en la gráfica. ¿En qué afecta el valor de la media a la gráfica de la distribución?

c. Al variar el valor de σ en cualquiera de las dos campanas produce un cambio en la gráfica. ¿En qué afecta el valor de la desviación estándar a la gráfica de la distribución?

d. ¿Qué sucede con la distancia entre el eje de simetría $X = \mu$ hasta la primera desviación estándar de las campanas cuando $\sigma_1 > \sigma_2$?

e. Posiciona los deslizadores μ_1 y μ_2 en el mismo valor, que sea distinto de 0. Realiza el mismo procedimiento para σ_1 y σ_2, sin que el valor escogido sea 1. Una vez se tengan las campanas superpuestas, y considerando que ahora las campanas son una sola, varía el valor de μ_2 hacia el $X = 0$, ¿qué cambio observas en la campana? Argumenta tu respuesta.

f. Ahora, sin mover los deslizadores para la media, varía el deslizador σ_2 hasta el valor 1, ¿Qué cambio observas en la campana? Argumenta tu respuesta.

g. A partir de lo anterior, ¿qué transformación algebraica se debe realizar para que el valor esperado μ cambie a la posición $X = 0$?

PASO 4. UN TEOREMA BASTANTE UTILIZADO

Estandariza las siguientes variables aleatorias, expresándolas en puntuación Z, y luego calcula la probabilidad indicada en cada caso. Además, realiza el registro gráfico de lo que se está calculando. asociada a ese valor.

a. $P(X \leq 90)$ si $X \sim N(\mu = 82; \sigma^2 = 17{,}5)$

b. $P(X > 22)$ si $X \sim N(\mu = 12; \sigma^2 = 2{,}5)$

c. $P(X \geq 41)$ si $X \sim N(\mu = 30; \sigma^2 = 25)$

d. $P(450 \leq Y \leq 700)$ si $Y \sim N(\mu = 600; \sigma^2 = 10\,000)$

e. $P(0 < Y < 6{,}3)$ si $Y \sim N(\mu = 7{,}7; \sigma^2 = 3)$

f. $P(|X - 39| < 2)$ si $X \sim N(\mu = 40; \sigma^2 = 16)$

PASO 5. DESCUBRIENDO UNA NUEVA PROPIEDAD

Considera los valores de las siguientes tablas, y responde las preguntas.

Z	0	0.01	0.02
-1.7	0.0446	0.0436	0.0427
-1.6	0.0548	0.0537	0.0526
-1.5	0.0668	0.0655	0.0643
-1.4	0.0808	0.0793	0.0778

Z	0	0.01	0.02
1.3	0.9032	0.9049	0.9066
1.4	0.9192	0.9207	0.9222
1.5	0.9332	0.9345	0.9357
1.6	0.9452	0.9463	0.9474

a. Si $P(Z \leq -1{,}5)$ ¿Cuál es la probabilidad pedida?

b. Si $P(Z \leq 1{,}5)$ ¿Cuál es la probabilidad pedida?

c. Dibuja las campanas asociadas a las siguientes probabilidades:

- $P(Z \leq -1{,}5)$
- $P(Z \leq 1{,}5)$
- $P(Z \geq 1{,}5)$

d. A partir de las gráficas encuentra alguna relación entre ellas y conjetura una regla para determinar la probabilidad de $P(Z > 1{,}5)$ en función de la $P(Z \leq 1{,}5)$

e. ¿Encuentras alguna relación entre el resultado anterior y $P(Z \leq -1{,}5)$?

f. A partir de los resultados anteriores, conjetura una regla general para encontrar la probabilidad $P(Z \leq -z)$.

g. Expresa con tus palabras lo que significa la regla encontrada en la pregunta f. Además, grafícala en una campana.

h. Utiliza las siguientes probabilidades para aplicarlas en la regla descubierta en la pregunta f. y verifica que el valor obtenido coincide con el de la tabla Z.

- $P(Z \leq 2{,}25) = 0{,}9878$
- $P(Z \leq 0{,}97) = 0{,}834$

PASO 6. SIGNIFICADO DE LA DISTRIBUCIÓN NORMAL

d. Una máquina que rellena botellas de bebidas sufrió un desperfecto técnico con la variabilidad del líquido respecto al volumen medio a vaciar en el envase, por ello la empresa ha decidido realizar un estudio para diagnosticar el nuevo valor de la desviación estándar. Los técnicos solo conocen que el volumen de líquido depositado sigue una distribución Normal de parámetros $\mu = 591\, ml$ con una desviación estándar desconocida.

i. Si se sabe que aproximadamente 3415 botellas de 5000 tienen un volumen entre $588{,}5\, ml$ y $593{,}5\, ml$ ¿cuál es la desviación estándar σ del volumen del líquido?

ii. ¿Cuál es la probabilidad que el contenido de la botella supere los $600\, ml$ de líquido?

e. La longitud diametral de las *donuts* que consume Homero Simpson siguen una distribución Normal de parámetros μ desconocido y varianza $\sigma^2 = 2\, cm^2$. Suponiendo que anualmente consume 1000 *donuts* de las cuales aproximadamente 683 tienen un diámetro entre $6{,}79$ y $9{,}61\, cm$. ¿Cuál es el valor de la media de la variable en cuestión?

PASO 7. DESESTANDARIZANDO UNA VARIABLE ALEATORIA NORMAL

A. Calcula los siguientes cuantiles de una variable aleatoria $N(0{,}1)$, esbozando la gráfica donde se identifique la probabilidad dada.

a. $P(Z \leq z) = 0{,}9$
b. $P(Z < z) = 0{,}95$
c. $P(Z \geq z) = 0{,}33$
d. $P(Z > z) = 0{,}9$
e. $P(z < Z < 1{,}7) = 0{,}9108$
f. $P(-0{,}46 < Z \leq z) = 0{,}6142$

B. Encuentra el valor de la variable original que encierra la probabilidad dada.

a. $P(X \leq x) = 0{,}9772$, si $X \sim N(\mu = 30; \sigma^2 = 25)$

b. $P(Y < y) = 0{,}2514$, si $Y \sim N(\mu = 87; \sigma^2 = 500)$

c. $P(X \leq x) = 0{,}9066$, si $X \sim N(\mu = 217; \sigma^2 = 625)$

d. $P(Y \leq y) = 0{,}1379$, si $Y \sim N(\mu = 90; \sigma^2 = 68)$

e. $P(X \geq x) = 0{,}6915$, si $X \sim N(\mu = 26; \sigma^2 = 4)$

f. $P(Y > y) = 0{,}0027$, si $Y \sim N(\mu = 6; \sigma^2 = 8)$

PASO 9. APLICACIÓN DE LA DISTRIBUCIÓN NORMAL

A. Según el estudio antropométrico en párvulos atendidos por el Sistema Educativo Público Chileno para el diseño de mobiliario, publicado en la revista científica Scielo, la estatura media de niños en el grupo de 25 a 36 meses es de $\mu = 91{,}30\ cm$, con una desviación estándar de $\sigma = 4{,}27\ cm$.

c. ¿A partir de qué estatura los niños se encuentran sobre el 12% superior de la población?

d. ¿Hasta que estatura, los niños se encuentran en el $80{,}5\%$ de las estaturas?

B. La vida útil de los motores de un automóvil de cierta marca se distribuye según una distribución Normal con el valor esperado de $\mu = 105\,000\ km$ y una desviación estándar de $\sigma = 10\,000\ km$.

d. La empresa que fabrica los motores hace constantemente pruebas con ellos, saca los que se encuentran con una vida útil de un 10% o menos. ¿Hasta qué valor de vida útil el motor no pasa las pruebas?

e. Se considera un riesgo que el motor tenga una vida útil en el $2{,}5\%$ más alto de la distribución, ¿desde qué valor se encuentra en esa categoría?

C. El consumo promedio de pan en Chile ronda los $90\ kg$ por persona al año, para su elaboración se utiliza principalmente harina. Esta materia prima posee distintos niveles de calidad basado de la cantidad de proteína presentes en ella. Específicamente en Santiago se realizó un estudio con 1500 panaderías sobre la calidad del pan producido en función de la cantidad de harina que usan. La categorización se hizo respecto a la ración de harina requerida para 1 kilogramo de pan, teniéndose los niveles bajo, medio y alto, de modo que el primero concentra el $26{,}11\%$ de las panaderías, el nivel medio considera el $66{,}54\%$ y el resto pertenece al último nivel. Además, se sabe que la cantidad promedio de harina para 1 kilogramo de pan sigue una distribución Normal de parámetros μ desconocido, con una variabilidad de $\sigma = 40\ g$.

a. Si se sabe que 1431 panaderías utilizan entre 525 y 685 gramos de harina, ¿cuál es el promedio de la cantidad de harina utilizado por las panaderías de Santiago para elaborar 1 kilogramo de pan?

b. Si se escoge una panadería al azar, ¿cuál es la probabilidad de que utilice más de 590 gramos de harina para la producción de $1\ kg$ de pan?

c. Si una panadería pertenece al nivel de calidad medio, ¿Entre qué valores se encuentra la cantidad de harina utilizada por el local? Bosqueja las campanas correspondientes a los tres niveles de calidad según la cantidad de harina que usan.

ACTIVIDAD 2: APLICAR EL MODELO NORMAL EN EL TRANSPORTE DE PERSONAS

DETERMINAR PROBABILIDADES QUE IMPLIQUEN EL USO DEL TEOREMA DEL LÍMITE CENTRAL

8. Se desea realizar una prueba con la intención de medir las habilidades matemáticas adquiridas por los estudiantes hasta segundo medio a 200 establecimientos educacionales de la Región Metropolitana. Esta evaluación consta de 40 preguntas que tienen 3 alternativas y solo una de ellas es correcta. Se han escogido al azar 30 personas de los colegios para rendir dicha evaluación. Suponga independencia entre los estudiantes que rinden la prueba.

    a. Define la variable aleatoria X que se está midiendo e identifica su distribución y parámetros respectivos.

    b. Con los parámetros encontrados en a. calcula la esperanza y desviación estándar de la población.

    c. Ahora, en Excel, generen 200 muestras de 30 individuos cada una con la herramienta *Análisis de datos ⟶ Generación de números aleatorios* y completen los datos como se muestra en la figura adjunta considerando para la sección *Parámetros* la información contestada en a.

    d. Calcula los promedios de cada muestra con la función $PROMEDIO$ de Excel.

    e. Realiza el histograma de los promedios muestrales y describe qué forma tiene la gráfica.

    f. Obtén la media y la desviación estándar de los promedios muestrales, utilizando la función $PROMEDIO$ y $DESVEST.P$ respectivamente.

    g. Comparen los valores obtenidos en la pregunta f. con los valores teóricos del Teorema Central del Límite. Concluyan si la diferencia entre ambos valores es significativa o no (considere como significativa la diferencia cuando es mayor a $0{,}02$).

ACTIVIDAD 3: APROXIMAR LA DISTRIBUCIÓN BINOMIAL POR LA DISTRIBUCIÓN NORMAL.

CÁLCULO DE PROBABILIDADES MEDIANTE LA APROXIMACIÓN NORMAL DE UNA DISTRIBUCIÓN BINOMIAL

1. ¿Cuáles son los parámetros que se necesita para determinar probabilidades según cada distribución?

    c. En GeoGebra, abran la vista cálculos de probabilidad, simulen una gráfica Binomial con $p = 0{,}5$ y los siguientes valores de n.

    - Con $n = 5$
    - Con $n = 10$
    - Con $n = 20$
    - Con $n = 30$

    d. Luego seleccionen la opción superposición de curva Normal y estimen donde se encuentra el centro de la campana superpuesta.

    e. Ahora, calcula la esperanza y desviación estándar de la Distribución Binomial y compárenlos con los parámetros dados de la Distribución Normal.

- Con $n = 5$, los parámetros de la Normal son $\mu = \frac{5}{2}$ y $\sigma = \frac{\sqrt{5}}{2}$.
- Con $n = 10$, los parámetros de la Normal son $\mu = 5$ y $\sigma = \frac{\sqrt{10}}{2}$.
- Con $n = 20$, los parámetros de la Normal son $\mu = 10$ y $\sigma = \sqrt{5}$.
- Con $n = 30$, los parámetros de la Normal son $\mu = 15$ y $\sigma = \frac{\sqrt{30}}{2}$.

Capítulo V. Conclusión y Reflexiones

Tal como define el MINEDUC (2021b), las actividades propuestas para la enseñanza de la Distribución Normal del Programa de Estudio de Probabilidades y Estadísticas Descriptiva e Inferencial son un apoyo para la labor docente bajo los lineamientos de las Bases Curriculares. Considerando que actualmente es el único material oficial, nuestro estudio se enfocó en potenciar el material pedagógico ministerial vinculado a la Distribución Normal, para ello resultó necesario analizar el conocimiento matemático para el trato de este objeto matemático y hacer una revisión exhaustiva y crítica del documento para luego elaborar una propuesta complementaria a la ministerial. Atendiendo la problemática, en este apartado se concluye y reflexiona detalladamente sobre cada uno de los objetivos específicos planteados para llevar a cabo esta investigación.

- OE 1: Caracterizar la Distribución Normal acorde a los dominios del Conocimiento Especializado del Profesor de Matemáticas (MTSK).

Describir y ejemplificar cada dominio y subdominio del MTSK acorde al conocimiento matemático que idealmente debería tener el profesor de matemática sobre la Distribución Normal es un ejercicio que no solo fue de gran utilidad para esta investigación sino también contribuyó a nuestro propio rol docente, ya que enriqueció nuestros conocimientos y nos permitió enfrentar con mirada amplia y de expertos la revisión del Programa de Estudio. Sin embargo, a partir de la indagación bibliográfica para llevar a cabo esta tarea, no podemos dejar de mencionar que falta explotar investigaciones nacionales sobre la enseñanza de la Distribución Normal en educación media, del mismo modo, faltan otras que se enfoquen en el conocimiento del profesor de matemática respecto a este objeto matemático comprendiendo que es un contenido que de a poco se ha posicionado con mayor fuerza en las aulas chilenas y requiere ser tratado con la cautela que demanda la más famosa distribución de probabilidad. De igual forma la acción de este objetivo específico nos llevó a concluir que, la elaboración de un apartado en la propuesta pedagógica que incorporó el saber sabio en base a todos los contenidos vinculados al objeto matemático que son tratados en la Unidad 3, es un aporte considerable a la labor docente, pues en un solo documento se pone a disposición el conocimiento matemático que los profesores deben poseer para afrontar la enseñanza de la Distribución Normal acorde a los lineamientos del MINEDUC.

- OE 2: Revisar críticamente la propuesta ministerial sobre el contenido de la Distribución Normal en el programa de estudio del plan diferenciado de Probabilidades y Estadísticas Descriptiva e Inferencial.

Revisar las actividades de la propuesta ministerial no fue una tarea sencilla, en ocasiones resultó que no se comprendía el propósito específico de algunas preguntas, por lo que tuvimos que deducirlas asumiendo ciertos hechos. Por ejemplo, para transitar de una pregunta asumimos que el docente realizó previamente la institucionalización del contenido, del mismo modo nos preguntamos ¿qué se supone que se debe hacer si en la primera parte de la actividad se pide convertir el puntaje original en puntaje z, pero no es hasta la segunda parte dónde aparece la desviación estándar? En este caso el documento no especificaba en ninguna parte que esperaba que hiciera el estudiante en esta situación. Es preciso señalar que la mayoría de estas deducciones se realizaron en la Actividad 1. Por ello concluimos la importancia de las orientaciones docentes a la hora de elaborar actividades a disposición pública, pues es en este apartado donde la información entregada debe ser clara y precisa, convirtiéndose en un aporte para planificar y ejecutar satisfactoriamente la propuesta pedagógica acorde al objetivo de aprendizaje declarado.

A medida que extendimos la revisión de las actividades, la idea de que fueron elaboradas por personas distintas aumentaba, porque la forma en que se trataban ciertos elementos no se encontraban en cohesión, pues al revisar las orientaciones docentes de cada actividad, las diferencias entre su redacción y presentación afirmaron nuestra hipótesis. Es evidente que participan varios autores en la elaboración de la propuesta ministerial, sin embargo, es importante que los encargados de la revisión final de ellas lo hagan críticamente, ya que no es favorable para los profesores y estudiantes que existan diferencias notables entre el diseño y modelo de una actividad con otra.

Es la Actividad 1 la que presenta más problemas de cohesión, porque subdivisión de Pasos parecían aisladas una de las otras y el aprendizaje no era conducido deductivamente a diferencia de sus pares que incluso trataban los contenidos a partir de una problemática bajo ciertos contextos. En particular, el Paso 4 de la propuesta ministerial revisa la aproximación de la Distribución Binomial a la Normal y Corrección de Continuidad. A nuestro criterio estos contenidos no deberían ser parte de la Actividad 1, porque su solución utiliza insumos que son tratados en la Actividad 3, incluso el Paso 5 del Programa de Estudios, deja de lado estos conceptos matemáticos, considerando que plantea problemas que incluyen los contenidos de los otros pasos. Esta observación podría ser resulta en otro ciclo de investigación-acción.

A pesar de que las actividades presentaron leves errores conceptuales, de notación y de tipeo, señalamos que en general las Actividades 2 y 3 son bastante adecuadas, pues la forma en que van tratando los contenidos permite la comprensión, reflexión y por sobre todo la construcción del conocimiento por parte de los estudiantes, por ello es necesario que el docente tenga bien claras las actividades para apoyar y conducir este proceso de aprendizaje.

Por último, en base a la validación realizada por los expertos, comprobamos que los resultados de nuestra apreciación sobre las actividades del Programa de Estudio no estaban tan alejados de lo que responden ellos, pues mencionan que el documento abarca puntualidades del contenido de estadística y probabilidades, convirtiéndolo en una sólida propuesta curricular, aun así, presenta algunos elementos que mejorar tales como, la claridad en orientaciones docentes y formulación de preguntas. Si bien, la propuesta no tiene errores garrafales, ya que estos se pueden abordar y solucionar de manera simple, no deben estar presentes, porque existe la posibilidad de problematizar la labor docente de aquellos que no cuentan con el saber sabio apropiado, afectando la forma en que trata los contenidos en el aula.

- OE 3: Describir los indicadores de evaluación que se presentan en menor proporción en el Programa de Estudio de Probabilidades y Estadísticas Descriptiva e Inferencial asociados al contenido de la enseñanza de la Distribución Normal.

En primer lugar, queremos manifestar nuestra opinión sobre los dos indicadores de evaluación que fueron propuestos adicionalmente para llevar a cabo la tarea de este objetivo. Si bien el uso de herramientas digitales no es parte de los indicadores de evaluación declarados en el Programa de Estudio, sí es una habilidad que el propio MINEDUC fomenta en las nuevas Bases Curriculares y de igual forma es considerado un recurso dentro del conocimiento de la enseñanza de la matemática (KMT), por ello resulta natural que usen software como GeoGebra y Excel para promover el aprendizaje, no obstante, a nuestro juicio y respondiendo a las competencias del siglo XXI opinamos que se puede sacar mayor provecho al uso de estas herramientas, pero si este ideal se aterriza a la realidad muchos docentes en ejercicio no poseen las habilidades básica para implementarlas y de igual forma varios establecimientos del país siquiera poseen sala de computación, por lo que es un desafío pendiente poner fin a estas problemáticas.

Respecto al segundo indicador de evaluación propuesto por nosotros, *la actividad contribuye a la comprensión de la distribución de las medias muestrales*, evidenciamos que en el Programa de Estudio se aborda el Teorema Central del Límite en la segunda parte de la actividad 2, sin embargo, ni en los indicadores de evaluación formativos de las orientaciones docentes de cada actividad, ni en los indicadores de cierre de unidad existe un descriptor para evaluar el aprendizaje del contenido. Creemos que esto no puede pasar por alto, ya que el TCL es uno de los teoremas importantes dentro de la estadística y no es intuitivo en términos de aprendizaje, por lo que se recalca la importancia que el MINEDUC oficialice un indicador que evalúe el logro de dicho contenido.

A partir de los resultados obtenidos se observa que las actividades de la propuesta ministerial responden casi en su totalidad tanto a los indicadores de evaluación planteados en el Programa de Estudio, como a los incorporados por nosotros, La actividad 1 fue la más descendida en comparación a sus pares, por lo que requirió ser atendida con mayor cautela, las demás solo debieron potenciarse. Se concluye que efectivamente fueron diseñadas bajo los lineamientos de las Bases Curriculares, entregando así un aprendizaje centrado en el aprendizaje.

- OE 4: Resolver las actividades propuestas para la enseñanza de la Distribución Normal en el Programa de Estudio de Probabilidades y Estadísticas Descriptiva e Inferencial, complementando aquellas que requieran de más detalle.

A la hora de resolver las actividades, presentamos dificultad en aquellas que involucran al Teorema Central del Límite, ya que olvidamos -en menor grado- los elementos asociados a este, debido a que requiere de bastantes conceptos previos que se deben tener en cuenta para realizar la transposición didáctica del TCL, como lo son variable aleatoria y sus transformaciones, distribución muestral, convergencia, estandarización, entre otros. No resulta extraño que esta situación haya sucedido pues como señala Méndez (1991, como se citó en Tauber, 2001) quien estudió la relación entre la comprensión del TCL junto con diversos conceptos y procedimientos que están implícitos en él, resultó que los estudiantes de pregrado con y sin formación estadística no evidenciaron una comprensión profunda del TCL, pues observó que justificaban erróneamente el teorema por medio de los axiomas de probabilidad. Por lo tanto, estas falencias pueden significar problemas de aprendizaje o comprensión del objeto matemático. Del mismo modo, quienes no cuenten con una formación sólida no podrán guiar correctamente a sus estudiantes, por ello destacamos el aporte que entrega el documento con la resolución de actividades, para todos los docentes que no tienen la formación profesional matemática suficiente para tratar eficazmente todos los contenidos para la enseñanza de la Distribución Normal. Sin duda el TCL es un concepto matemático que requiere un trato minucioso, siendo así resulta probable pasar por alto ciertos elementos de este. Uno de ellos es la condición que debe cumplir el tamaño de cada muestra ($n > 30$) que hace posible el uso del teorema, sin embargo, la Actividad 2 de la propuesta ministerial utiliza $n = 6$, para estos casos se utiliza la suma de Normales para muestras pequeñas, no obstante, no es un contenido escolar. Por lo tanto, corregir este error conceptual escapa de los insumos escolares, además la Actividad 2 gira en torno a un contexto real que implica la cantidad de personas que pueden subir a un teleférico, haciendo imposible editar el valor de las muestras a menos que se trabaje bajo supuestos irreales. De igual forma se realizó el desarrollo de esta Actividad, pero queda pendiente un plan de mejora para un segundo ciclo de investigación de acción.

Otro posible aporte, involucra al conocimiento de la enseñanza de las matemáticas, pues al poder observar los posibles desarrollos de las actividades implica también idear distintas maneras de presentar el contenido, por ejemplo, las Actividades 2 y 3 podrían implementarse de tal forma que se adapten a la Teoría de Situaciones Didácticas (TSD) de Brousseau.

Al resolver las actividades pudimos evidenciar nuestros propios errores de notación, por ejemplo, cuando aproximamos la Distribución Binomial a la Normal, consideramos la igualdad entre algunas expresiones en vez de utilizar el símbolo de aproximación para conectarlas, olvidando que estamos aproximando una variable aleatoria discreta a una continua. Por tal motivo, enfatizamos en la relevancia de las notaciones, pues sin importar el contexto la rigurosidad matemática no puede pasar por alto, mucho menos cuando hablamos de un curso que no solo profundiza en los contenidos de la formación general, sino también entrega herramientas para enfrentar carreras de educación superior, de manera que, los docentes somos responsables de transmitir correctamente los contenidos matemáticos y del mismo modo debemos seguir y acompañar el proceso de aprendizaje de los estudiantes para que no incurran en errores de este tipo.

Elaborar la propuesta de actividades en línea a lo que estipula el MINEDUC, fue una tarea que no estuvo exenta de complejidades, si bien el primer instrumento proporcionó los objetivos y propósitos de las subactividades y preguntas que elaboramos, tuvimos dificultades porque no queríamos que todos los ejercicios propuestos fuesen iguales, al contrario, la intención siempre fue otorgar variedad de interrogantes y problemas. A su vez, estas se confeccionaron de manera que los estudiantes descubran y construyan por su propia cuenta los conocimientos, de esta forma se convierten en protagonistas de su aprendizaje y queda para el docente realizar la institucionalización de los contenidos.

En lo que respecta a la aplicación del instrumento de validación de la propuesta de actividades de esta investigación, hubo varios percances con relación al juicio de los expertos. Por un lado, considerando las respuestas del *experto 2* pareciera que no comprendió que, las primeras dos preguntas abiertas consideraban solo las subactividades y preguntas elaborada por nosotros, porque al momento de mencionar aspectos positivos indica que algunos links propuestos en el programa estaban caducados y las orientaciones al docente eran muy generales, sin embargo, nosotros no propusimos algún enlace para la construcción del diagrama de cajas, ya que no forma parte del contenido en cuestión, además de que nuestras orientaciones y/o notas docentes iban dirigidas a cada una de las actividades que propusimos, lo que las hacía singulares para cada propuesta. Por su parte el *experto 3*, no respondió el Google Forms, a pesar de ello, sabemos que lo revisó pues su primer comentario alude a la gramática del formulario. Estas situaciones dificultaron validar nuestra propuesta ya que finalmente tuvimos dos respuestas para la escala de estimación, así mismo tuvimos dos respuestas abiertas y una incompleta referidas a nuestro trabajo. Finalmente, después de todo, nos quedamos con los aspectos positivos que en suma son suficientes para corroborar su contribución.

Por otro lado, recogiendo la apreciación de los jueces sobre la propuesta de esta investigación, se tiene que, el *experto 1* comenta que solo consideramos factores generales, dejando fuera los tiempos destinados a cada actividad y estrategias didácticas específica, además existe ausencia de introducción motivante y atractiva, sumado a lo que mencionó el *experto 3*, que debíamos plantear diversas situaciones, para trabajar la habilidad de resolución de problemas, queremos destacar que desde un principio nuestro enfoque era agregar ejercicios que complementaran y suplieran aquellos que el documento ministerial no presentaba, por ello, no se agregan elementos como los que menciona el *experto 1*. Tampoco fue nuestra intención generar actividades que solo mecanicen el aprendizaje, pues consideramos que eso no es provechoso para el proceso enseñanza-aprendizaje, a pesar de ello en la actividad 1 introdujimos pasos para que el estudiante ejercite, como se dice coloquialmente "suelten la mano" comprendiendo que al ser el inicio del trato del contenido en la Unidad, se da el caso que estudiantes de IV° medio deben recordar el contenido o los de III° medio se están recién familiarizado con ellos, por lo mismo queda a criterio del profesor utilizar o no las actividades, lo que es claro que después de esta propuesta tienen más opciones que facilitan su labor a la hora de realizar clases. Bajo el mismo contexto, el *experto 3* señaló que las preguntas propuestas eran demasiadas y que bastaría solo un poco de aquellas, sin necesidad de revisar todos los casos. Sin embargo, quisimos dar cobertura a todos los casos, para ayudar a los profesores a mostrar cada uno de ellos, de esta forma al conocer el contexto de los estudiantes él pueda seleccionar los que más se acerquen a este, o elija los que considere idóneos a desarrollar. Comprendemos que, por la extensión del documento y el tiempo total destinado a la implementación de la Unidad, es prácticamente imposible abordar todo lo que estipula el Programa de Estudio, ya que se debe tener en cuenta que se constituye de tres unidades más, que requieren la misma importancia que la Unidad 3. Por ello, la propuesta viene a complementar lo que más pueda.

Recordando que seguimos un modelo de investigación-acción, se consideran los aspectos negativos de las actividades complementarias propuestas como elementos para abordar en otro ciclo de investigación, de manera que proyectamos los resultados del nuestro para contribuir en nuevas investigaciones que aborden desde otra perspectiva la problemática de este trabajo, así se podría plantear, por ejemplo, la elaboración de una propuesta didáctica que sea utilizada tanto por el profesor, como por los estudiantes. De igual forma se pueden implementar en las aulas las actividades que menos se atienden según el contraste con los indicadores de evaluación, de esta forma se obtiene su comprobación empírica.

En términos generales, concluimos que nuestra propuesta está en concordancia con el Programa de Estudio, ya que las subactividades y preguntas incorporadas cumplen con los objetivos asignados y aportan elementos clave para la propuesta ministerial, pues los tres jueces coincidieron en que, ya sea con el uso de herramientas digitales o con los contenidos faltantes al programa, nuestra propuesta suma información importante, sobre todo porque se enfatiza en desarrollar la habilidad matemática de argumentar. Del mismo modo el documento final de esta investigación es un aporte al docente, porque contar con los contenidos del saber sabio y la resolución de las actividades facilita su labor, pues fue adrede que se incluyeran las gráficas acampanadas, imágenes de la tabla Z y los algoritmos para el uso de las herramientas digitales, esperamos que estos insumos sean útiles en la enseñanza.

Finalmente, por medio de esta investigación fortalecimos los conocimientos didáctico matemático alcanzados en nuestra formación profesional, además, nos instruyó sobre la actualización curricular de la formación diferenciada de Probabilidad de Probabilidades y Estadísticas Descriptiva e Inferencial.

Referencias Bibliográficas

Anderson, I. (2019). Abraham De Moivre, Pierre-Simon Laplace y Carl Friedrich Gauss: La nueva ciencia matemática sobre el concepto teórico de la belleza particular concreta y material en la que los objetos, artefactos y productos del diseño artesanal e industrial influencian sobre el gusto de los usuarios/consumidores (casos). *ArtyHum*, 63, 1-29.

Arenas, F. y Arriagada F. (2021). *Propuesta de guía docente para el contenido y resolución de actividades sugeridas por el MINEDUC en los contenidos de estadística inferencial para el electivo Probabilidades y Estadísticas Descriptiva e Inferencial para III° o IV° medio*. Tesina para optar al grado de Licenciado en Educación Matemática. Universidad Metropolitana de Ciencias de la Educación. Santiago, Chile.

Batanero, C. (2001). *Didáctica de la Estadística*. Granada: Universidad de Granada.

Batanero, C. y Godino, J. (2001). *Análisis de Datos y su Didáctica*. Granada, España.

Batanero, C., Godino, J. y Roa, R. (2004). *Training teachers to teach probability*. Journal of Statistics Education, 12.

Behar, R. & Grima, P. (2013). El histograma como un instrumento para la comprensión de las funciones de densidad de probabilidad. Actas de las Jornadas Virtuales en Didáctica de la Estadística, Probabilidad y Combinatoria (pp. 229-23). Granada, Universidad de Granada.

Cansado, E. (1970). *Estadística General*. CIENES. Santiago, Chile.

Carrillo, J., Contreras, L., Climent, N., y Montes, M. (2017). *Introducción al modelo MTSK: origen e investigaciones realizadas*. Universidad de Huelva. España.

Carrillo, J., Contreras, L., Climent, N., Muñoz, M., Montes, M., y Rojas, N., (2015). Conocimiento especializado del profesor de matemáticas (MTSK): un modelo analítico para el estudio del conocimiento del profesor de matemáticas. *La Gaceta de la RSME*, 18(3), 589-605.

Colmenares, A y Piñero, M. (2008). LA INVESTIGACIÓN ACCIÓN. Una herramienta metodológica heurística para la comprensión y transformación de realidades y prácticas socio-educativas. *Laurus*, *14*(27), p. 96-114.

Conde, G. (2015). La Distribución Normal una rápida revisión histórica. *Heurística*, 17, 59-65.

CPEIP. (2018). *Informe de Resultados Institucionales Evaluación Nacional Diagnóstica de la Formación Inicial Docente*. Santiago, Chile.

CPEIP y MINEDUC. (2021). *Estándares Pedagógicos y Disciplinarios para Carreras de Pedagogía en Matemática*. Santiago, Chile.

Escudero, E. (2015). *Una caracterización de conocimiento didáctico del contenido como parte del conocimiento especializado del profesor de matemática de secundaria*. Tesis doctoral. Universidad de Huelva, España.

Flores, E., Montes, M., Contreras, L., Catalán, M., y Mar, M. (2016). El Papel del MTSK como Modelo de Conocimiento del Profesor en las Interrelaciones entre los Espacios de Trabajo Matemático. *Bolema*, 30(54), 204-221.

Giselle. (2015). *GeoGebra* [Free Software Foundation]. Recuperado 18 de noviembre de 2021, de https://www.geogebra.org/m/J6KfvXw5

González, Y., Ojeda, A., y Reyes, D. (2018). Comprensión de profesores de la Distribución Normal. *Acta Latinoamericana de Matemática Educativa*, 31(2), 1764-1772.

Hernández, Fernández y Baptista. (2006). *Metodología de la Investigación*. (5ta ed.). McGraw-Hill.

Latorre, A. (2005). *La investigación- acción: Conocer y cambiar la práctica educativa.* (3ª ed.) Barcelona, España: Graó.

Ley N°20.370. (2009). Ley General de Educación. MINEDUC. Diario Oficial de la República de Chile.

Lewis, A. (2003). *Test Psicológicos y Evaluación*. (11 Ed.). México: Pearson Educación.

Másmela, L.A. y Serrato, J.C. (2009). Una aproximación histórica a la evolución de la curva normal. Trabajo presentado en el Segundo Congreso Internacional las matemáticas un lenguaje universal, ALAMMI 2009. Asociación Colombiana de Maestros de Matemáticas.

Meyer, P. L. (1992). Probabilidad y Aplicaciones Estadísticas. Addison-Wesley Iberoamericana.

MINEDUC. (2019). *Bases Curriculares para 3° y 4° Medio*. Santiago, Chile.

MINEDUC. (2021a). *Bases Curriculares de Educación para personas y jóvenes adultas.* Santiago, Chile.

MINEDUC. (2021b). *Programa de Estudio 3° o 4° Medio Formación Diferenciada Matemática, Probabilidades y Estadísticas Descriptiva e Inferencial*. Santiago, Chile.

Morales, R. (2014). *GeoGebra* [Free Software Foundation]. Recuperado 18 de noviembre de 2021, de https://www.geogebra.org/m/k6xFEq7W

Moya, R. y Saravia, G. (2009) *Probabilidad e inferencia estadística.* Editorial San Marcos. Perú.

Pimienta, J. (2008). Evaluación de los aprendizajes. Un enfoque basado en competencias. México: Pearson.

Pizarro, R. (2015). *Estimación de medida: el conocimiento didáctico del contenido de los maestros de primaria*. Tesis doctoral. Universitat Autònoma de Barcelona. Barcelona, España.

Rincón, L. (2013). *Introducción a la probabilidad*. Ciudad de México, México.

Rivas, H. (2014). *Idoneidad didáctica de procesos de formación estadística de profesores de educación primaria*. Tesis doctoral. Universidad de Granada, España.

Rodríguez, F. (2021a). *El eje de Datos y Azar es un desafío pendiente en la Formación Inicial Docente después de más de una década en el currículo escolar chileno* [Diapositivas de PowerPoint]. Chillán, Chile.

Rodríguez, F. (2021b). *El eje de Datos y Azar: un desafío pendiente en la Formación Inicial Docente después de más de una década en el currículo escolar chileno*. Recuperado el 31 de agosto de 2021 de https://cmmedu.uchile.cl/events/el-eje-de-datos-y-azar-un-desafio-pendiente-en-la-formacion-inicial-docente-despues-de-mas-de-una-decada-en-el-curriculo-escolar-chileno/

Rodríguez, F., Vásquez, C. y Rojas, F. (2020). Formación inicial docente en profesores de matemática: una mirada desde la evaluación nacional diagnóstica. *Estudios pedagógicos*, 45, 141-153.

Romero, M. (2020). *GeoGebra* [Free Software Foundation]. Recuperado 18 de noviembre de 2021, de https://www.geogebra.org/material/show/id/ya5ygkgu

Sandín, M. (2003). *Investigación cualitativa en Educación: Fundamentos y Tradiciones*. McGraw-Hill. Madrid, España.

Tauber, L. (2001). *La construcción del significado de la distribución normal a partir de actividades de análisis de datos*. Tesis Doctoral. Universidad de Sevilla. Sevilla, España.

Vásquez, C. y Alsina, A. (2015). El conocimiento del profesorado para enseñar probabilidad: Un análisis global desde el modelo del Conocimiento Didáctico – Matemático. *Avances de Investigación en Educación Matemática*, 7, 27 – 48.

Vega, Ó. (2002). Surgimiento de la Teoría Matemática de la Probabilidad. Apuntes de Historia de las Matemáticas, 1, 54-62.

Walpole, R., Myers, R., Myers, S. y Ye, K. (2007). *Probabilidad y estadística para ingeniería y ciencias*. México: Pearson Educación.

Yague, A., Morón, J., y López, I. (2020). *GeoGebra* [Free Software Foundation]. Recuperado 18 de noviembre de 2021, de https://www.geogebra.org/m/ypu8kpvz

Yáñez, J. (2016). *El conocimiento especializado del profesor de matemática sobre la resolución de problemas de optimización de funciones aplicando el concepto de derivada. Una investigación-acción.* Tesis para optar al grado de Magíster en Didáctica de la Matemática. Pontificia Universidad Católica de Valparaíso. Valparaíso, Chile.

Anexos

Anexo 1. Instrumento 1: Contraste entre los Indicadores de Evaluación y las actividades para la enseñanza Distribución Normal del Programa de Estudio.

N°	Indicador	Totalmente en desacuerdo				Parcialmente de acuerdo				Totalmente de acuerdo				No Aplica				Justificación
		A1	A2	A3	A4	A1	A2	A3	A4	A1	A2	A3	A4	A1	A2	A3	A4	
1	La actividad permite identificar las principales características de una Distribución Normal, tales como la varianza o desviación estándar y la media.					x					x	x	x					A1: Estamos parcialmente de acuerdo, porque en esta actividad no se aprecia la comprensión de la desviación estándar, o qué característica aporta a la campana de Gauss. Tampoco se enseña a utilizar la tabla Z.
2	La actividad permite interpretar información que involucra la Distribución Normal y					x					x	x	x					A1: Estamos parcialmente de acuerdo. La decisión se fundamenta en que, si bien se interpreta

	sus parámetros.																	información que es necesaria para realizar los cálculos pertinentes, no se pide la interpretación de las probabilidades determinadas en dicha actividad, como tampoco la de los parámetros de la Distribución Normal.
3	En la actividad se resuelven problemas que involucran la Distribución Normal.									x	x	x	x					
4	Por medio de la actividad los estudiantes determinan la probabilidad de intervalos en distribuciones Normales, estandarizando y utilizando la tabla Z.									x	x	x	x					
5	La actividad propicia la toma de decisiones, basándose en la								x		x	x		x				

	Distribución Normal, y en el cálculo de probabilidades.																	
6	Por medio de la actividad los estudiantes argumentan cuando se puede modelar un fenómeno con una Distribución Normal.											x	x	x	x			
7	En la actividad se pueden modelar fenómenos o situaciones cotidianas, científicas y sociales mediante la Distribución Normal.									x	x	x	x					
8	En la actividad se pueden modelar situaciones que involucren la aproximación de la Binomial a la Normal.									x		x	x		x			

9	La actividad contribuye efectivamente al aprendizaje de la Distribución Normal por medio de herramientas digitales.						x					x		x			x	A2: Parcialmente de acuerdo, ya que el uso de las herramientas digitales sirvió para comprender la distribución de medias muestrales, pero no para la Distribución Normal como tal.
10	La actividad contribuye a la comprensión de la distribución de las medias muestrales (Teorema Central del Límite)										x		x	x		x		

Anexo 2. Tabla de Observaciones y Sugerencias para las actividades propuestas en el Programa de Estudio para la enseñanza de la Distribución Normal

N° Actividad	N° de Pregunta, Paso o problema	Observaciones	Sugerencias
1	1	➢ La gráfica del ítem a) es poco lúdica y su eje no presenta notación. ➢ No hay un hilo conductor entre el ítem c) y d). ➢ Los ítems d) y e) son análogos, resultan lo mismo. ➢ Para realizar el ítem f) se asume que el docente institucionalizó las variables aleatorias continuas.	➢ Mejorar el enunciado del ítem d) tal forma que se mantenga el hilo inductor de la actividad. Además, agregar una pregunta para tratar las variables aleatorias continuas.
	2	➢ El Paso 2 no presenta título. ➢ El inicio general del segundo paso comete un error al señalar como histograma a un gráfico que es discreto, pues es un gráfico de barras, además en su elaboración no se consideró la notación de los ejes. ➢ Los ítems a) y b) tienen concordancia entre ellas para que el estudiante pueda evidenciar cómo las variables aleatorias discretas se convierten en continuas. ➢ Para el ítem c) se asume que el docente ya ha tratado el valor esperado, la variable aleatoria y la distribución Binomial.	➢ Buscar un título acorde al Paso 2. ➢ Como alternativa al Paso 2, se sugiere realizar una actividad que involucre el uso del aparato de Galton por su utilidad para mostrar cómo la Distribución Binomial converge a la Z. ➢ Trasladar la pregunta c, para que se incorpore a la primera actividad propuesta.
	3	➢ El ítem a) utiliza una expresión poco familiar para los estudiantes. Además, la función ϕ no se define y	➢ Se sugiere al docente que para seguir la línea intuitiva con la que se trabajó los pasos anteriores, debe

		tampoco se menciona a qué refiere o de dónde aparece, por lo tanto, se asume que previo a esta actividad el docente ya institucionalizó la Distribución Normal Estándar. ➢ Considerando las orientaciones docentes, aquí se señala que interpreten el significado de los intervalos icónicos de la Distribución Normal, pero no se aprecia en ninguno de los ítems que esto sea tratado. ➢ El paso 3 pierde esa naturaleza intuitiva que tenían los pasos anteriores.	comenzar el Paso 3 trabajando con la dispersión de los datos a través de gráficos. ➢ En a) agregar una imagen para mostrar y explicar la propiedad de simetría, asimismo, explicar quién es ϕ. Por lo tanto, con este recurso se puede determinar las probabilidades pedidas según el intervalo dado. ➢ Añadir preguntas donde los estudiantes puedan resolver problemas usando las probabilidades icónicas para determinar parámetros de distintas situaciones bajo un contexto dado.
	4	➢ El ítem d) se puede abordar de dos formas. La primera opción, conecta este ítem con el Paso 1, así los estudiantes pueden reflexionar e indicar que la probabilidad es 0. En segundo lugar, se puede utilizar la corrección de continuidad en caso de que se haya enseñado con anterioridad.	➢ Anterior a los ítems del Paso 3, se sugiere elaborar una pregunta donde se solicite calcular la probabilidad pedida usando la Distribución Binomial, de esta forma se van a encontrar con una dificultad, pues la calculadora no puede determinar el número combinatorio que necesitan.
	5	➢ Los dos problemas propuestos en el Paso 5, son de aplicación, sin embargo, no se apoyan en recursos gráficos. ➢ No se aplica la desestandarización ni los intervalos icónicos para determinar probabilidades. ➢ En ambos problemas de este Paso no se pide interpretar los resultados. ➢ Para resolver los ejercicios de este Paso se deduce que ya se comprendió completamente el conocimiento de los objetos matemáticos involucrados.	➢ Integrar otro problema que permita calcular probabilidades utilizando los intervalos icónicos y la desestandarización.

	General	➢ El Paso 1, 3, 4 y 5, presentan un título a modo de guía para identificar el objeto del contenido a tratar, no obstante, el Paso 2 no presenta esta sugerencia. ➢ Falta más aplicación de GeoGebra para que los estudiantes puedan visualizar y comprender las propiedades de la Distribución Normal. ➢ En ninguna de las preguntas se visualiza que traten la importancia de σ, en cambio, se apunta solo a la media. ➢ La actividad sugiere varios enlaces caducados.	➢ Solicitar que los estudiantes dibujen a priori el gráfico de campana, pues no requiere ser graficada a escala, porque a través de ella se muestra la información sin la necesidad de ser precisos. ➢ Incorporar herramientas digitales, como GeoGebra para que los estudiantes puedan visualizar y comprender las propiedades de la Distribución Normal. ➢ Entre el Paso 2 y 3 incorporar tres pasos más, de tal forma que en el primero se reconozcan los parámetros de la Distribución Normal, en el segundo se visualice la propiedad simétrica y en el último se pida calcular probabilidades por medio de la estandarización. ➢ Añadir otro Paso entre el 3 y el 4 para ejercitar la desestandarización. ➢ Incorporar una actividad que incorpore la comprensión de la desviación estándar bajo el contexto trabajado. Por ejemplo, se pide que comparen gráficamente dos distribuciones de igual media y distinta varianza identificando cómo cambia la curva y de qué forma influye la desviación en ella.
2	1	➢ Se clasifica como problema N° 1 a la agrupación de las primeras seis preguntas que giran en torno a la problemática del teleférico. ➢ Se considera que la pregunta 1 permite abordar los parámetros de la Distribución Normal. ➢ No se entiende el objetivo de la pregunta 3, pues ni siquiera apunta a identificar la media. ➢ La pregunta 4 pide determinar la probabilidad de un intervalo, sin embargo, se infiere que	➢ Para dar un sentido a la pregunta 3 se sugiere cambiar su orientación. ➢ Para la pregunta 4 se sugiere al docente promover la construcción de la campana de Gauss para visualizar la probabilidad pedida. ➢ En 4.a cambiar la pregunta “¿Por qué es necesario convertir la masa en su puntuación z correspondiente para buscar la probabilidad buscada? Señalen los valores” por “Calcule la probabilidad pedida haciendo

		intuitivamente identifiquen la simetría de la Distribución Normal respecto a la media, verificando la probabilidad teóricamente. ➢ Se asume que el x pedido en a) corresponde a 77,3, además se pide identificar el valor de la desviación estándar sin haber entregado este dato, por lo que se considera irrelevante para el cálculo pedido, ya que se está solicitando la probabilidad de que su masa sea menor a la media. En consecuencia, el apartado 4c no puede determinarse realizando un cálculo, sin embargo, si se puede resolver a través de la observación de la visualización del diagrama de la Distribución Normal. ➢ No se puede realizar la pregunta 5, ya que no es posible calcular la probabilidad pedida sin el valor de σ. ➢ Para evaluar la pregunta 6, se asumió que σ es un dato dado.	uso de la tabla Z" ➢ Para la pregunta 4b) se sugiere que el docente entregue las orientaciones que considere pertinente a los estudiantes respecto al aporte de la tabla de Distribución Normal estándar. ➢ Agregar el valor de σ para resolver las preguntas 4.a) y 5.
	2	➢ Se clasifica como problema N° 2 a la agrupación de siete preguntas que giran en torno a *determinar probabilidades que impliquen el uso del teorema del límite central*, usando el contexto del problema del teleférico. ➢ El enunciado del problema presenta un error de precisión conceptual pues no se cumple la condición $n > 30$ para realizar la estandarización. ➢ Hay un error de tipeo en la pregunta 1, pues el 6 de "77,6" debe ser un 3. ➢ Se destaca positivamente la pregunta 3, ya que	➢ Para calcular la desviación estándar de la pregunta 5.b, hay que utilizar la desviación estándar poblacional. ➢ Se sugiere añadir en 3.b) el gráfico del histograma allí señalado. ➢ En la pregunta 6, a modo de cierre, en la institucionalización del contenido tratado, se sugiere al docente incorporar dos campanas de Gauss, la primera considerando la variable original y la segunda con la variable X, la cual solo varía en su dispersión. ➢ Para corregir el error de tipeo en 7, se debe eliminar "masa total" y el dígito 1 de 180.

		efectivamente permite determinar y observar el comportamiento de las medias muestrales. ➢ La pregunta 7 presenta un error de tipeo o planteo, en caso de que su intencionalidad fuese que los estudiantes calculen la probabilidad. En caso contrario, se considera que la pregunta induce una reflexión. ➢ Para la pregunta 5.a se asume que habla de la media poblacional.	
	General	➢ En general la actividad no presenta errores de notación.	➢ Para complementar y sacar provecho al uso de la herramienta de datos de Excel, se sugiere incorporar una pregunta que incluya la simulación de una muestra que no provenga de una Distribución Normal y evidenciar que, de todas formas, Distribuye Normal.
3	1	➢ Se clasifica como subactividad N°1 a la agrupación de cinco preguntas asociadas a la "*aproximación de una Distribución Binomial mediante una Distribución Normal*". ➢ Para la pregunta 1, se asumió que la finalidad en 1.a y 1.b es trabajar con las estimaciones. En particular la estimación de p, de la Distribución Binomial. ➢ En la pregunta 2, falta mencionar la independencia de las variables. ➢ En la pregunta 3 faltan indicaciones para comprender y mejorar la gráfica de la Figura 1. ➢ La pregunta 4 es bastante inductiva, pues por medio de su desarrollo permite al estudiante visualizar como la Distribución Binomial se	➢ En la pregunta 3 agregar la fórmula de la Distribución Binomial, se sugiere explicar porque el recorrido de la gráfica de la Figura 1 tiene seis elementos, además se debe especificar que el eje corresponde a la cantidad de personas felices. ➢ En la pregunta 3.a agregar los títulos a los ejes del gráfico. ➢ Añadir más detalles de las indicaciones del uso GeoGebra en la pregunta 4.a.

		acampana a medida que n crece.	
	2	➢ Se clasifica como subactividad N°2 a la agrupación de cinco preguntas asociadas al "*Cálculo de probabilidades mediante la aproximación Normal de una Distribución Binomial*". ➢ En la pregunta 3. resulta complicado que los estudiantes identifiquen por sí solos el ajuste de los intervalos que usa la corrección de continuidad. ➢ Se destaca que en la pregunta 4 se deduzca la regla de la corrección de continuidad en vez de ser dada inmediatamente en la fórmula.	➢ En ayuda a la comprensión de la pregunta 1.c se sugiere utilizar los gráficos construidos en 4.c de la primera parte de la actividad, luego preguntar dónde está el centro en las gráficas acampanadas, identificando que este valor corresponde a la media, también se puede preguntar por el valor esperado de la v.a. original, que corresponde a la Binomial, asociando que $\mu = np$. En consecuencia, se sugiere incorporar preguntas desde 1.b, cuya finalidad permita determinar la relación entre los parámetros de la Distribución Binomial y Normal a partir de la visualización de gráficas en el software GeoGebra. ➢ En la pregunta 3 es necesario que el docente dirija el aprendizaje para comprender la corrección de continuidad, mostrando gráficamente el cálculo de la probabilidad con la Distribución Binomial, luego mostrar que si se calcula con la Distribución Normal resulta cero, luego mostrar en un gráfico, la Distribución Normal superpuesta al gráfico de la Distribución Binomial indicando que no se trabajará específicamente en el 75, sino más bien, en un intervalo
	General	➢ A partir de la descripción del propósito de esta actividad, se asume que aproximar la Distribución Binomial a una Normal es un contenido que ya fue tratado con anterioridad. ➢ Las preguntas 2.b, 4.a y 5 de la segunda actividad tienen error de notación, pues utilizan x en vez de	➢ Se sugiere al docente utilizar una herramienta digital que simula el aparato Galton.

		X para la variable aleatoria continua. ➢ La actividad presenta varios ítems donde se deben usar herramientas digitales, no obstante, faltó mostrar el aparato de Galton en dicha aproximación.	
4			➢ En el ítem III, permutar las preguntas 3.b y 3.c, con la finalidad de tener mayor coherencia la actividad.

Anexo 3. Instrumento 2: Cuestionario para Validación de las Actividades

Revisión de actividades y preguntas complementarias a la propuesta ministerial del Programa de Estudio de Probabilidades y Estadísticas Descriptiva e Inferencial

Estimado(a) Profesor(a):

Estamos realizando una investigación relacionada con el Conocimiento Matemático para la enseñanza de la Distribución Normal a partir de una revisión crítica al Programa de Estudio del plan diferenciado de Probabilidades y Estadísticas Descriptiva e Inferencial. Para ello, hemos construido el presente instrumento que tiene como objetivo recabar información que nos ayude a reflexionar y tomar decisiones importantes en relación a la elaboración de preguntas y subactividades planteadas que complementen las ya propuestas por el Ministerio de Educación. Es necesario que sepa que sus respuestas únicamente serán utilizadas para el objetivo mencionado.
Por favor, lea detenidamente el siguiente instrumento. Primeramente, revise el documento donde se encuentran las actividades y luego responda las preguntas que se adjuntan en el formulario. Por último, dispone de tres preguntas abiertas para que exprese su opinión sobre las actividades de aprendizaje elaborada por los investigadores de este estudio y su apreciación sobre las actividades propuestas por el Ministerio de Educación.

Anticipadamente, agradecemos su participación, pues estamos seguros que su información nos será de gran utilidad.

En el siguiente enlace podrá descargar el documento:
https://drive.google.com/file/d/1mEcQBPbb_Em_GT00-IMnjPuHDhylb1Wi/view?usp=sharing

¿Ha realizado docencia para el electivo en estudio? *

- ◯ Sí
- ◯ No

Si su respuesta anterior es afirmativa, ¿utilizó las actividades propuestas en el programa de estudios?

- ◯ Sí
- ◯ No

Con respecto a las subactividades complementarias sugeridas por los investigadores seleccione una alternativa atendiendo al grado de acuerdo que tenga, con la afirmación que se presenta. Para ello considere que, 3: De acuerdo, 2: Ni en acuerdo ni es desacuerdo, 1: En desacuerdo. *

	3	2	1
¿Considera que las preguntas y subactividades complementarias sugeridas se incorporan coherentemente con las actividades de la propuesta ministerial?	○	○	○
¿Considera que fue un aporte incorporar estas preguntas y subactividades para el aprendizaje de los contenidos asociados a la Distribución Normal?	○	○	○
¿Considera que las preguntas y subactividades complementarias sugeridas permiten el logro de los objetivos señalados al inicio de cada una de ellas?	○	○	○
¿Utilizaría las subactividades complementarias sugeridas para la enseñanza del contenido asociado a la Distribución Normal?	○	○	○

A continuación señale los aspectos positivos sobre su apreciación de las preguntas y subactividades elaboradas para esta investigación. *

Tu respuesta

A continuación señale los aspectos negativos sobre su apreciación de las subactividades elaboradas para esta investigación, con el objetivo de considerar futuras mejoras en ellas. *

Tu respuesta

A continuación indique su apreciación crítica sobre las actividades de la propuesta ministerial. *

Tu respuesta

Anexo 4. Respuestas de los expertos

Anexo 4.1. Respuestas experto 1.

¿Ha realizado docencia para el electivo en estudio? *

○ Sí

◉ No

Si su respuesta anterior es afirmativa, ¿utilizó las actividades propuestas en el programa de estudios?

○ Sí

◉ No

Con respecto a las subactividades complementarias sugeridas por los investigadores seleccione una alternativa atendiendo al grado de acuerdo que tenga, con la afirmación que se presenta. Para ello considere que, 3: De acuerdo, 2: Ni en acuerdo ni es desacuerdo, 1: En desacuerdo. *

	3	2	1
¿Considera que las preguntas y subactividades complementarias sugeridas se incorporan coherentemente con las actividades de la propuesta ministerial?	◉	○	○
¿Considera que fue un aporte incorporar estas preguntas y subactividades para el aprendizaje de los contenidos asociados a la Distribución Normal?	◉	○	○
¿Considera que las preguntas y subactividades complementarias sugeridas permiten el logro de los objetivos señalados al inicio de cada una de ellas?	◉	○	○
¿Utilizaría las subactividades complementarias sugeridas para la enseñanza del contenido asociado a la Distribución Normal?	◉	○	○

A continuación señale los aspectos positivos sobre su apreciación de las preguntas y subactividades elaboradas para esta investigación. *

Las preguntas de aplicación tienen un enfoque contextualizado y cercano a lo cotidiano, permitiendo darle un uso real a los conocimientos adquiridos sobre cada parámetro estudiado, facilitando así la comprensión y aplicación de estos.

Considero que el uso de recursos digitales como Geogebra en las actividades, es un aspecto positivo ya que contribuye a procesos indagatorios en el proceso de enseñanza-aprendizaje a través de la comparativa. Por otro lado, ayuda a que los estudiantes valoren el uso de tecnologías en estos procesos probabilísticos permitiendo también dar más enfoque a la interpretación de los resultados que se obtienen a través de estos.

Por último, las orientaciones al docente son clarificadoras y acordes a lo que las actividades permiten lograr.

A continuación señale los aspectos negativos sobre su apreciación de las subactividades elaboradas para esta investigación, con el objetivo de considerar futuras mejoras en ellas. *

Comprendiendo la abstracción que la unidad a trabajar implica, me llama la atención la ausencia de una introducción de carácter motivante y atractiva para el estudiante. Como profesor de segundo ciclo y enseñanza media considero que el inicio de una unidad es clave para una correcta acogida del contenido por parte de los estudiantes, implicando también una entrega e interés.

A continuación indique su apreciación crítica sobre las actividades de la propuesta ministerial. *

Considero que como propuesta curricular es bastante sólida, ya que abarca puntualidades del contenido de probabilidad y estadística descriptiva e inferencial de forma efectiva, haciendo uso de ejemplos, contextualizaciones, recursos digitales en cada concepto a trabajar, de manera que el objetivo de aprendizaje a lograr se va construyendo clase a clase de manera coherente entre éstas, dando lugar a una linealidad conceptual bastante clara.

La conexión interdisciplinaria con ciencias para la ciudadanía en las actividades contribuye enormemente al logro de los objetivos de la unidad, ya que permite a los estudiantes valorar el alcance de lo que están aprendiendo.

Por último, el instrumento de validación apunta a factores demasiado generales sobre la propuesta, impidiendo una valoración mas técnica sobre otros factores importantes tales como; tiempos destinados a cada actividad calzando con un tiempo realista de las clases, también factores como estrategias didácticas especificas a utilizar en alguna clase o actividad.

Anexo 4.2. Respuestas experto 2.

¿Ha realizado docencia para el electivo en estudio? *

◉ Sí

○ No

Si su respuesta anterior es afirmativa, ¿utilizó las actividades propuestas en el programa de estudios?

◉ Sí

○ No

Con respecto a las subactividades complementarias sugeridas por los investigadores seleccione una alternativa atendiendo al grado de acuerdo que tenga, con la afirmación que se presenta. Para ello considere que, 3: De acuerdo, 2: Ni en acuerdo ni es desacuerdo, 1: En desacuerdo. *

	3	2	1
¿Considera que las preguntas y subactividades complementarias sugeridas se incorporan coherentemente con las actividades de la propuesta ministerial?	○	◉	○
¿Considera que fue un aporte incorporar estas preguntas y subactividades para el aprendizaje de los contenidos asociados a la Distribución Normal?	○	○	◉
¿Considera que las preguntas y subactividades complementarias sugeridas permiten el logro de los objetivos señalados al inicio de cada una de ellas?	○	◉	○
¿Utilizaría las subactividades complementarias sugeridas para la enseñanza del contenido asociado a la Distribución Normal?	○	◉	○

A continuación señale los aspectos positivos sobre su apreciación de las preguntas y subactividades elaboradas para esta investigación. *

Es importante mencionar que producto de la pandemia, no alcanzamos a ver el programa completo del electivo, y específicamente las actividades trabajadas corresponden a las unidades 1 y 2 (los estudiantes no desarrollaron actividades de distribución normal presentes en el programa ministerial).

*Las actividades y preguntas tenían por objetivo que el estudiante pudiese interpretar los resultados obtenidos y/o presentados, más que simplemente calcular alguna medida en particular.
* las problemáticas planteadas y bases de datos utilizadas en su mayoría eran situaciones reales e interesantes de estudiar.
*El uso de geogebra para estudiar el comportamiento de la media en distribuciones simétricas y asimétricas.

A continuación señale los aspectos negativos sobre su apreciación de las subactividades elaboradas para esta investigación, con el objetivo de considerar futuras mejoras en ellas. *

*El planteamiento de algunas preguntas era poco claro y carecía de coherencia.
*Algunos links propuestos en el programa estaban caducados (recuerdo el link para construir diagramas de cajas)
*Las orientaciones al docente eran muy generales.

A continuación indique su apreciación crítica sobre las actividades de la propuesta ministerial. *

Nos ayuda a orientar los objetivos y metodología del electivo, con preguntas que buscan el análisis, la reflexión e interpretación de resultados obtenidos en contextos reales. Vincula de buena forma los contenidos con el uso de herramientas digitales.
A mejorar, la claridad en la formulación de algunas preguntas y que las orientaciones al docente sean más específicas en base a lo que se busca en cada pregunta.

Anexo 4.3. Respuestas experto 3.

Primero que todo, tanto en el formulario como en las actividades, hay un uso excesivo de adverbios de modo. (detenidamente, primeramente, formativamente, constantemente). Por lo general, se recomienda no abusar de ellos, ya que atentan contra la economía del lenguaje y entorpecen la comprensión.

A

Página 4. En ejemplo del recuadro hay un error. Lo correcto es que sigma 1 < sigma 2.

Me parece una actividad complementaria que aporta a la comprensión del contenido. En realidad, sin duda que todas las actividades y subactividades que proponen ayudan a la comprensión, así como también enriquecen el contenido y funcionan como bisagra entre las actividades que plantea el Ministerio, las cuales suelen ser algo escuetas, por lo que su trabajo es ir rellenando estos vacíos que detectaron.

Ahora bien, como apreciación personal, me sucede que la mayoría de las actividades que plantean, de alguna forma, mecanizan el aprendizaje. Me explico, en su mayoría son ejercicios que tratan sobre resolver problemas rutinarios, similares unos de otros, donde con unos pocos de ellos bastaría, sin necesidad de revisar todos los casos. Por lo demás, se quedan en una habilidad básica, sin avanzar dentro de la misma habilidad con problemas de mayor complejidad.

Por ejemplo, para A es suficiente con 3 o 4 ejercicios de reconocimiento de los parámetros dados dos campanas. Aquí se desaprovechan otros tipos de ítems. Por ejemplo, se pueden dar dos poblaciones, una de salmones que habitan una región A y otra que vive en una región B. Si se sabe la longitud X de los salmones de la región A se distribuye según X ~ N(50, 9), mientras que la longitud de los salmones de la región B se modela mediante X ~ N(48, 16) [Estoy inventando los números por si acaso], entonces se puede solicitar esbozar las campanas, o bien, reconocer las gráficas a partir de 3 opciones de respuesta posibles. Y lo mismo, se puede consultar cuales salmones, en promedio, son de mayor longitud o cual población tiene una menor desviación, etc, de modo que la actividad no sea tan mecánica ni repetitiva. Presentar diversas situaciones que se vinculen al contenido permite a los estudiantes enfrentarse a distintos escenarios donde no es tan inmediato aplicar el contenido, de modo que estos tendrán que reflexionar un poco para llegar a la respuesta.

En lo personal, me gustan esta incorporación de ejercicios, ya que permite pasar de un sistema de representación a otro, pero por lo mismo, si se tiene una actividad donde se les entregan los parámetros y les piden a ellos que esbocen la gráfica, entonces se profundiza en esta habilidad.

Por otro lado, ustedes hablan de “punto de inflexión”. No estoy seguro si este concepto se define en los planes y programas. Si no se hace, habría que definir qué es. (Aquí se van a topar con otro problema, que es definir concavidad, para que se entienda lo que es un cambio de concavidad. No es el objetivo, pero hay que hacerse cargo).

B

a) ¿A qué se refieren con “tener el mismo centro”? (¿Se refieren al eje de simetría?). Se requiere precisión conceptual.

Paso 5.

En d) y f) se habla de “hallar”, se recomienda utilizar “encontrar” o “determinar”, las cuales sí son habilidades del pensamiento.

Para el resto de las actividades mis comentarios son similares. Sin duda que son un aporte, pero la idea es plantear diversas situaciones, ojalá relacionadas con otra ciencia o situación cotidiana, de formar que se trabaje la habilidad de resolución de problemas (aunque el problema del teleférico trabaja esta habilidad).

Por último, está muy presente la habilidad de argumentar en todas las actividades, lo que sin duda es un aporte.

Anexo 5. Propuesta "La Distribución Normal en el Curso de Probabilidades y Estadísticas Descriptiva e Inferencial"

LA DISTRIBUCIÓN NORMAL EN EL CURSO DE PROBABILIDADES Y ESTADÍSTICAS DESCRIPTIVA E INFERENCIAL

Resolución de las actividades propuestas por el Ministerio de Educación en el Programa de Estudios del curso, junto con subactividades elaboradas por los investigadores

Francesca Fuentealba Aucapán

Juan Gallardo Valenzuela

Santiago de Chile, 2022

A continuación, se incluye un apartado que contiene los contenidos que trata la Distribución Normal en el Programa de Estudio, estos permiten al docente reactivar el saber sabio necesario para abordar la Unidad 3.

Además, se presenta un segundo apartado, que incluye la resolución de actividades del Programa de Estudio del curso Probabilidades y Estadísticas Descriptiva e Inferencial propuesta del Ministerio de Educación, junto con las subactividades y preguntas elaboradas por los investigadores, todas ellas abordan la Distribución Normal y algunas de sus aplicaciones, (como Teorema Central del Límite y Aproximación de la Distribución Binomial a la Normal). A lo largo de la propuesta de los investigadores, se señala el objetivo o finalidad de cada actividad y algunas orientaciones o notas docentes que aportan ciertas indicaciones que van en su ayuda.

Para facilitar la comprensión del desarrollo de los problemas, durante toda la resolución se puede distinguir que existen extractos de la tabla Z que facilitan la visualización de esta. Así, las imágenes presentadas en color rosado indican que se utiliza para encontrar la probabilidad, por el contrario, las que presentan color celeste son utilizadas para determinar cuantiles. También se presentan las campanas con la puntuación Z y la puntuación de la variable original.

Es preciso señalar que para los casos cuando la centésima o milésima es mayor a 5 se redondea al dígito siguiente. Por ejemplo, $-1{,}165$ se redondea a $-1{,}17$.

Asimismo, se incluyen los algoritmos para el uso de las herramientas digitales, como Excel o GeoGebra. Por último, en anexos podrá encontrar, la tabla Z utilizada para las probabilidades o cuantiles, y las bases de datos de las simulaciones realizadas.

CONTENIDO PARA EL DOCENTE

DEFINICIÓN DE VARIABLE ALEATORIA

Una variable aleatoria corresponde a una función que le asigna un número real a cada elemento de un espacio muestral asociado a un experimento aleatorio, así, el dominio de una variable aleatoria es Ω y su recorrido es $\mathbb{R}$ o un subconjunto de este. El recorrido corresponde a todos los valores que puede tomar la variable aleatoria y se denota $Rec(X)$, siendo X la variable en cuestión.

Hay dos tipos de variables aleatorias, las discretas y las continuas. Una variable aleatoria X será discreta si su recorrido es un conjunto de cardinalidad finita o infinita numerable. Por ejemplo, el número de asistencias a las clases de matemática en un año, número de veces que sale cara al lanzar dos monedas, entre otros. Por el contrario, será continua si su conjunto posible de valores es todo un intervalo de números reales. Por ejemplo, el tiempo de vida de un artículo electrónico, la altura o masa de las personas, entre otros (Rincón, 2013). A cada valor dentro del recorrido de una variable aleatoria se le asocia una probabilidad, esta relación es conocida como función de probabilidad, de esta forma, se define:

FUNCIÓN DE PROBABILIDAD PARA UNA VARIABLE ALEATORIA DISCRETA

La función de probabilidad asociada a una variable aleatoria discreta, denominada f.d.p. se denota por $P_X(x)$ o $P(x)$, y su dominio corresponde al recorrido de la variable aleatoria. Según Rincón (2013), para que efectivamente sea f.d.p debe cumplir con dos condiciones:

a. $P_X(x) \geq 0 \; \forall x \in Rec(X)$, es decir, la función de probabilidad es no negativa para todos los valores del recorrido de la variable aleatoria discreta.

b. $\sum P(X = x) = 1, x \in Rec(X)$, es decir, la suma de las imágenes de una función de probabilidad siempre debe ser 1. En otras palabras, todas las probabilidades de la variable sumadas deben resultar 1.

De esta forma, se define la variable aleatoria discreta $P_X(x)$ como sigue:

$$P_X(x) = \begin{cases} P(X = x) & con\ x \in Rec(X) \\ 0 & en\ otro\ caso \end{cases}$$

La función de distribución corresponde a una función de probabilidad acumulada de la variable aleatoria, es decir, la suma de las imágenes de la función de probabilidad, desde la primera, hasta la indicada. Se anota $F_X(x) = P(X \leq x)$ y matemáticamente corresponde a:

$$F_X(x) = P(X \leq x) = \sum_{t \leq x} P_X(t) \ \forall x \in Rec(X)$$

FUNCIÓN DE PROBABILIDAD PARA UNA VARIABLE ALEATORIA CONTINUA (FUNCIÓN DE DENSIDAD)

La función de densidad se asocia a una variable aleatoria continua, cuya función es integrable y no negativa, para cualquier intervalo $[a, b]$, la denotamos por $f_X(x)$, y se define matemáticamente como:

$$P(X \in [a, b]) = \int_a^b f_X(x)dx$$

De esta forma, la probabilidad de que la variable aleatoria continua X tome un valor dentro del intervalo $[a, b]$ se puede calcular como el área bajo la curva de $f_X(x)$ en dicho intervalo (Rincón, 2013).

Al igual que en una f.d.p. para variables aleatorias discreta, la función de densidad debe cumplir con dos condiciones para ser considerada como tal:

a. $f(x) \geq 0 \ \forall x \in \mathbb{R}$, es decir, la función de probabilidad es no negativa para todos los valores del recorrido de la variable aleatoria continua.
b. $\int_{-\infty}^{\infty} f_X(x)dx = 1$, es decir, el área bajo la función de densidad siempre debe ser 1. En otras palabras, todas las probabilidades de la variable sumadas deben resultar 1.

La gráfica de una función de densidad corresponde a rectas o curvas continuas y como se mencionó en a., bajo ellas se encuentra el 100% de los elementos pertenecientes al espacio muestral.

Como menciona Rincón (2013), a diferencia de una función de probabilidad asociada a una variable aleatoria discreta, en una función de densidad no se puede determinar la probabilidad de que la variable aleatoria tome un determinado valor, ya que bajo un punto no es posible determinar el área, por lo que esta será igual a cero, matemáticamente esto es:

$$P(X = x) = \int_x^x f_X(t)dt = 0$$

Las variables aleatorias, sin importar su naturaleza, cuentan con algunas propiedades que ayudan al cálculo de probabilidades, estas son:

1. $P(a < X \leq b) = P(X \leq b) - P(X \leq a)$
2. $P(a \leq X \leq b) = P(X \leq b) - P(X < a)$
3. $P(a < X < b) = P(X < b) - P(X \leq a)$
4. $P(a \leq X < b) = P(X < b) - P(X < a)$
5. $P(X > a) = 1 - P(X \leq a)$
6. $P(X \geq a) = 1 - P(X < a)$

A las propiedades 5. y 6., se les denomina *propiedad del complemento*.

DEFINICIÓN DE DISTRIBUCIÓN NORMAL Y SUS PROPIEDADES

Si X es una variable aleatoria continua, se dice que sigue una Distribución Normal de parámetros μ correspondiente a la media y σ^2 correspondiente a la varianza, que se anota como $X \sim N(\mu, \sigma^2)$, si su función de densidad de probabilidad está dada por

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(x-\mu)^2}{2\cdot\sigma^2}}$$

cuyo dominio son los números reales, es decir $x \in \mathbb{R}$ (Meyer, 1992).

Los dos parámetros de los que depende la función de densidad tienen un significado dentro de la gráfica, presentada en la Figura 1.

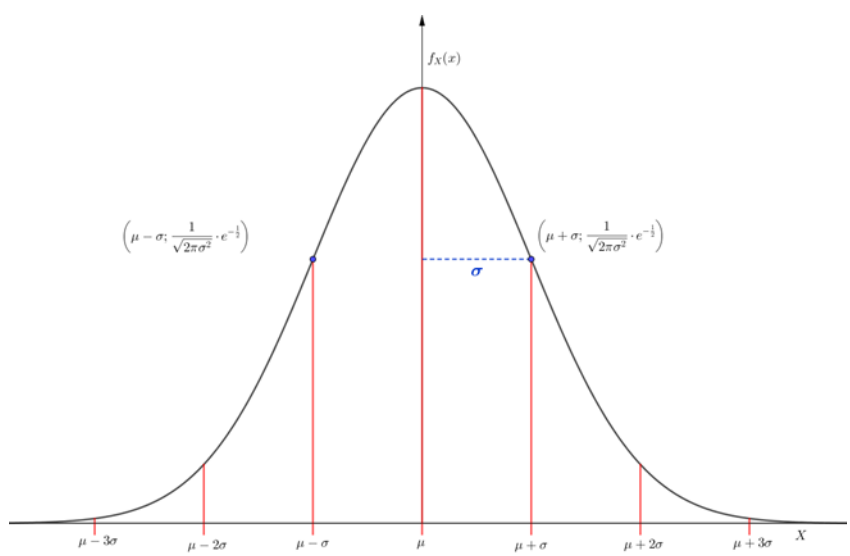


Figura 1: Campana de Gauss, asociada a la Distribución Normal. Se pueden apreciar los cambios de curvatura, el máximo y la desviación estándar.

Por un lado, la media μ es el valor central en la campana, de ella depende la ubicación que tenga en el eje de las abscisas, mientras mayor sea este valor, más hacia la derecha en el eje X se ubicará la gráfica. Este valor μ puede ser negativo, positivo o cero y es el eje de simetría de la campana acumulando probabilidad máxima, es decir, también coincide con la moda y la mediana, pues hasta μ se acumula un 50% de información. Por otro lado. La raíz de la varianza, es decir, la desviación estándar, denotada por σ, informa sobre la dispersión que poseen los datos y se puede visualizarse en la campana, desde el eje de simetría μ, hasta cualquiera de los dos cambios de curvatura que la gráfica tiene en $\left(\mu-\sigma; \frac{1}{\sqrt{2\pi\sigma^2}}\cdot e^{-\frac{1}{2}}\right)$ y $\left(\mu+\sigma; \frac{1}{\sqrt{2\pi\sigma^2}}\cdot e^{-\frac{1}{2}}\right)$, esto es posible verlo en la Figura 1. Finalmente, hay que mencionar que $\sigma^2 > 0$ y que, a mayor desviación estándar, el ancho de la curva aumenta, mientras que su altura disminuye, ya que existe mayor dispersión de los datos, por el contrario, a menor σ, la altura aumenta y el ancho de la curva disminuye (Rincón, 2013; Meyer, 1992).

Para simplificar el cálculo de probabilidades, puesto que la integral de la función de densidad de la Normal no tiene primitiva, se hace uso de una tabla Z, para obtener las probabilidades acumuladas. Como la media y la desviación estándar pueden tomar una infinidad de valores, es evidente que para cada número asignado a μ y σ no se podrá tener una tabla de probabilidades acumuladas, por lo que se tiene que estandarizar mediante un cambio de variable, por medio de la siguiente expresión.

$$Z = \frac{x-\mu}{\sigma} (1)$$

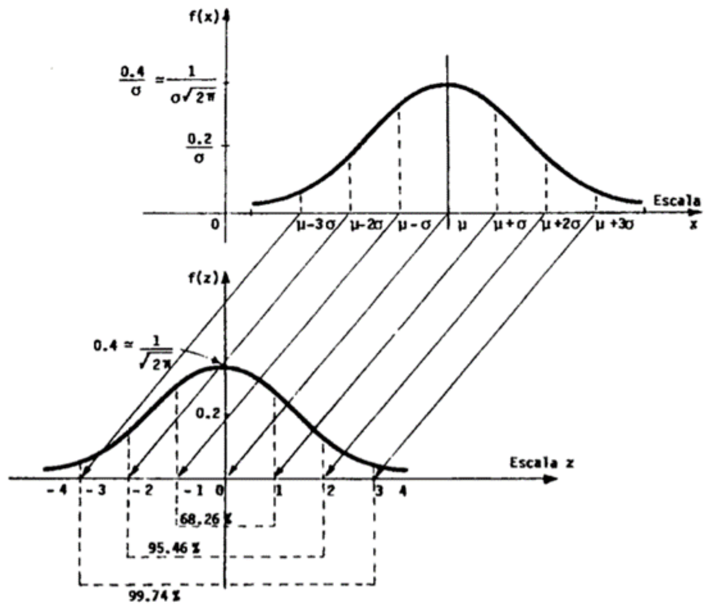


Figura 2: Estandarización de una variable aleatoria $X \sim N(\mu, \sigma^2)$ a la variable aleatoria $Z \sim N(0,1)$. Se puede apreciar el cambio que realiza la estandarización.

En este caso, la estandarización nos conduce a una Normal de parámetros $\mu = 0$ y $\sigma^2 = 1$, que anotamos como $N(0,1)$. Es importante destacar que se divide por la desviación estándar σ y no por la varianza σ^2. Al restar μ, hace la traslación para ubicar a la campana en $X = 0$, y al dividir por σ, se obtiene un ajuste en la escala, acorde al desplazamiento de la campana hasta el origen del plano cartesiano. Una vez realizada la estandarización o tipificación, se obtiene la distribución Normal Estándar, cuya función de densidad de probabilidad está dada por:

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{z^2}{2}}$$

Con $z \in \mathbb{R}$, (Rincón, 2013; Meyer, 1992).

La función de distribución acumulada de $f_Z(z)$, anotada como $P(Z \leq z)$, es tan importante dentro del cálculo de probabilidades, que posee una denominación propia, dada por la letra griega *phi* en otras palabras, se tiene que

$$\phi(z) = P(Z \leq z)$$

En la Figura 3, se puede apreciar el significado de la función de distribución acumulada.

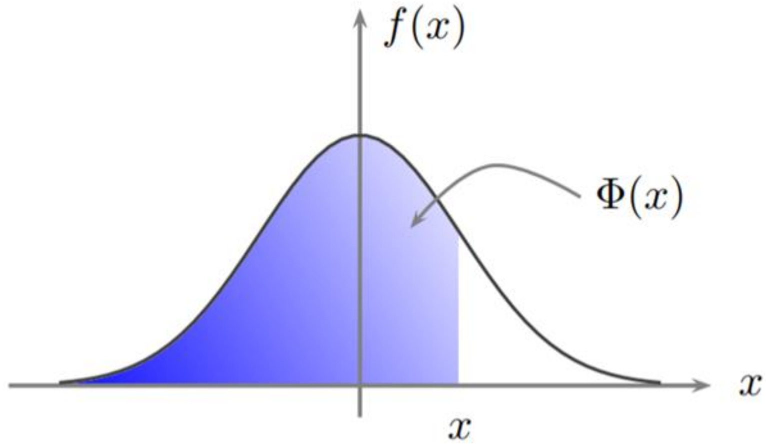


Figura 3: Significado geométrico de la función de distribución acumulada o función ϕ.

Como la tabla Z, da las probabilidades acumuladas menores o iguales a la buscada, entonces se puede abreviar la función de distribución acumulada, por ejemplo $P(Z \leq 1) = \phi(1)$. Muchos ejercicios utilizan la probabilidad acumulada, y se requiere conocer el cuantil que acumula cierta probabilidad, para aquellos ejercicios, se denotará la inversa de *phi,* denotada por $\phi^{-1}(z)$, y corresponde a tener la expresión $P(Z \leq z_\alpha) = \alpha$, con $0 < \alpha < 1$, en este caso, z_α será la incógnita por determinar (Rincón, 2013). Coloquialmente se entiende que $\phi^{-1}(z)$ corresponde a *ver la tabla al revés*, o en forma inversa.

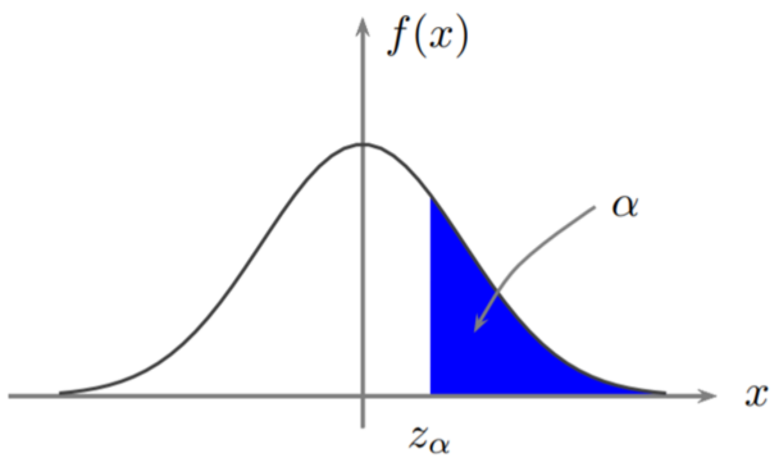


Figura 4: Significado geométrico del número z_α.

En síntesis, para poder encontrar las probabilidades de la distribución Normal de parámetros cualesquiera, se debe estandarizar la variable, mediante la ecuación (1), y buscar la la puntuación z obtenida en la tabla Z. Ya sea con la función de distribución Normal, o Normal Estandarizada, el área total que se acumula bajo la campana es 1.

Según Rincón (2013) y Meyer (1992), la Distribución Normal cuenta con varias propiedades que se presentan a continuación.

1. Dentro de las más importantes está la simetría de la campana respecto al eje $x = \mu$, esta propiedad posee vital importancia, ya que permite que el cálculo de áreas sea más sencillo. Así, por ser simétrica, se tiene que bajo y sobre μ hay un 50% de información respectivamente, es decir, existe la misma cantidad de casos. De esta forma, se puede escribir que $P(X < \mu) = 0{,}5$. Además, de la simetría se deduce la siguiente fórmula,

$$\phi(-z) = 1 - \phi(z)$$

   con ella, se pueden encontrar cuantiles que acumulan la probabilidad α y $1 - \alpha$, solo con cambiar el signo del número z.

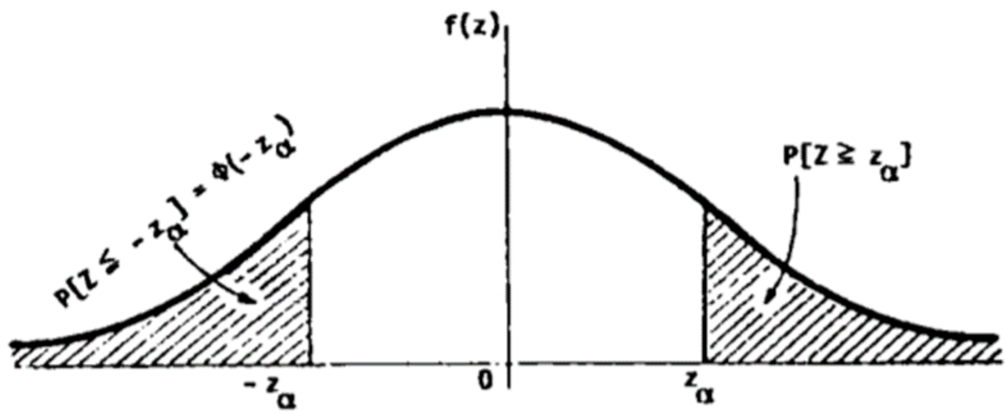


Figura 5: Propiedad de simetría de la Distribución Normal.

2. Respecto a propiedades relacionadas con la media y la desviación estándar se tiene que:

   a. Dos curvas con la misma desviación estándar, que tienen medias distintas, son idénticas en la forma de campana, sin embargo, poseen el eje de simetría en distintas posiciones a lo largo de las abscisas, es decir se desfasan a lo largo del eje horizontal. Esto se muestra en la Figura 6.

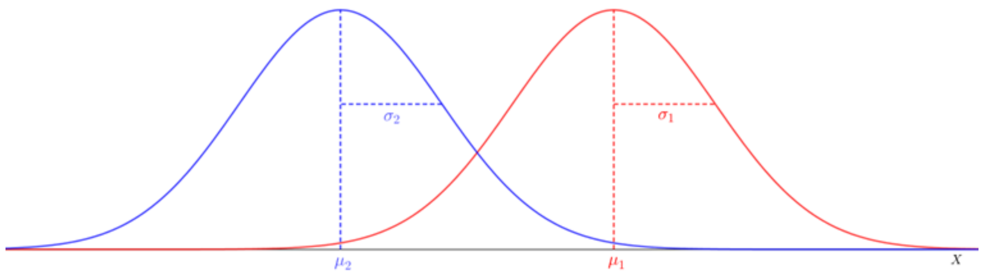


Figura 6: Campanas con igual σ y distinto μ. En la imagen, se muestra que $\mu_2 < \mu_1$, ya que μ_2 está a la izquierda que μ_1.

b. Si dos distribuciones tienen la misma media, pero distintas desviaciones estándar, las curvas poseen el mismo eje de simetría μ, pero la campana que posee desviación estándar es más ancha y baja que la otra, pues los datos se encuentran más dispersos. Se ilustra en la Figura 7.

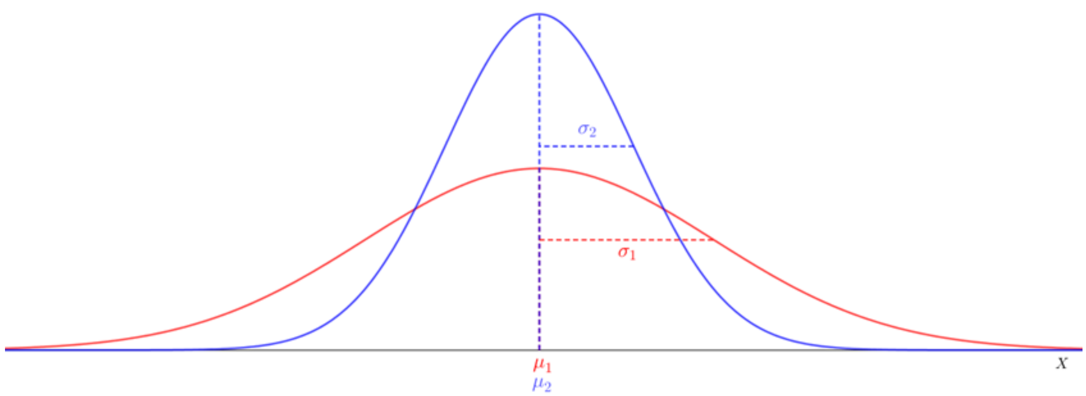


Figura 7: Campanas con igual μ y distinto σ. En la imagen, se muestra que $\sigma_2 < \sigma_1$, ya que la campana azul es menos dispersa, que la campana roja, es decir, tiene más concentración de datos alrededor de la media.

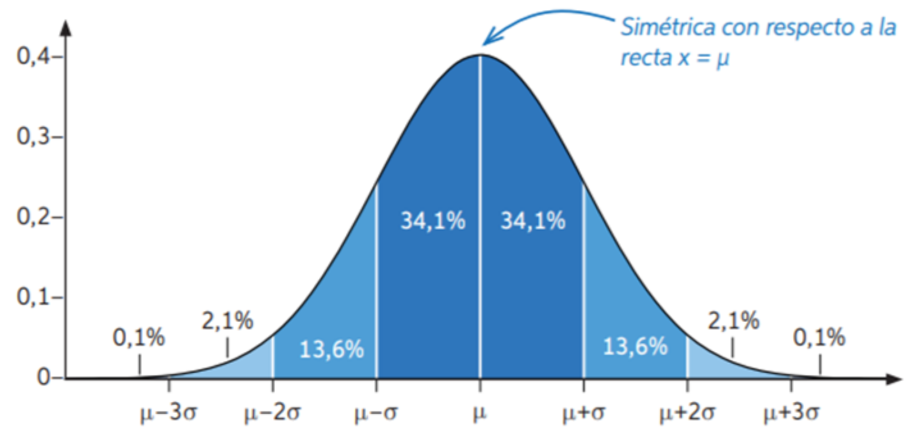
c. Relación entre las probabilidades y la media y desviación estándar: Hay ciertos valores clásicos dentro de la curva Normal, que siempre acumulan la misma probabilidad, sin importar los valores que tomen μ y σ. Aproximadamente se tiene:

$$P(\mu - \sigma \leq X \leq \mu + \sigma) = 0{,}683$$

$$P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) = 0{,}954$$

$$P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) = 0{,}997$$

En la Figura 8, obtenida del *Texto del estudiante Matemática 3° y 4° medio* se aprecian los intervalos clásicos. (Batanero y Godino, 2001; Osorio et al., 2019)



Como se muestra en la figura, los valores de la función de probabilidad son:

- $P(x \leq \mu) = 0{,}5$ Hay un 50 % de los valores hasta la media de la distribución.
- $P(\mu - \sigma \leq x \leq \mu + \sigma) = 0{,}683$ Hay un 68,3 % de los valores en este intervalo.
- $P(\mu - 2\sigma \leq x \leq \mu + 2\sigma) = 0{,}954$
- $P(\mu - 3\sigma \leq x \leq \mu + 3\sigma) = 0{,}997$ Aproximadamente el 0,3 % de los valores quedan fuera de este intervalo.

Figura 8: Intervalos clásicos de la Distribución Normal

La distribución Normal posee diversas aplicaciones dentro de la estadística, sin embargo, las que correspondientes a la Unidad 3 del electivo son:

- Distribución de las medias muestrales, o distribución de $\bar{X}$, también llamado Teorema Central del Límite.
- Aproximación de la distribución Binomial por la Normal.

DISTRIBUCIÓN DE LAS MEDIAS MUESTRALES, O DISTRIBUCIÓN DE $\bar{X}$

Según Walpole et al. (2007), dada una población que no sea necesariamente distribuya Normal, al tomar muestras de igual tamaño n, con $n > 30$, desde la población de origen, la media de los promedios muestrales tiende a tener un comportamiento parecido a una Distribución Normal.

En este caso, la media de los promedios muestrales, denotada $\bar{X}$, corresponderá a la media de la población μ_X, mientras que la desviación estándar de la distribución $\sigma_{\bar{X}}$ será igual al cociente entre la desviación estándar de la población, denotada por σ_X, y la raíz cuadrada del tamaño de cada muestra. Matemáticamente se anota que:

La esperanza o media de $\bar{X}$ es $\mu_{\bar{X}} = \mu_X$

La desviación estándar de $\bar{X}$ es $\sigma_{\bar{X}} = \frac{\sigma_X}{\sqrt{n}}$

Así, se tiene que

$$\bar{X} \xrightarrow{d} N\left(\mu_X; \frac{\sigma_X}{\sqrt{n}}\right)$$

Si la expresión anterior se estandariza, se obtiene lo siguiente:

$$\left(\frac{\bar{X} - \mu_X}{\sigma_X} \cdot \sqrt{n}\right) \xrightarrow{d} N(0,1)$$

En rigor, lo que viene a mencionar el Teorema Central del Límite, es que, sin importar la distribución de la variable original, si $n > 30$, entonces, en media se tiene una convergencia a la Normal (Walpole et al., 2007). Es importante destacar, que, a partir de esta distribución de las medias muestrales, se obtienen los intervalos de confianza para la media con varianza poblacional conocida.

APROXIMACIÓN DE LA DISTRIBUCIÓN BINOMIAL A LA NORMAL.

Antes de hablar de la aproximación, se ofrece una definición de la Distribución Binomial.

DEFINICIÓN DE DISTRIBUCIÓN BINOMIAL

Si un experimento aleatorio Bernoulli o dicotómico tiene las siguientes características:

- Se puede realizar n veces.
- Cada repetición es independiente de las otras, es decir, la repetición no se ve afectada por experimentos anteriores.
- La probabilidad de éxito p y fracaso q son constantes en cada repetición, es decir, no varían de una prueba a otra.

Se dice que la variable aleatoria discreta X sigue una Distribución Binomial, donde la variable expresa el número de éxitos obtenidos en cada repetición, se anota $X \sim Bin(n, p)$ y su función de probabilidad está dada por

$$P(X = x) = \begin{cases} \binom{n}{x} \cdot p^x \cdot q^{n-x} & si\ x = 0,1,2, \ldots, n \\ 0 & en\ otro\ caso \end{cases}$$

Si $X \sim Bin(n, p)$, entonces:

$$E(X) = \mu = n \cdot p$$

$$Var(X) = \sigma^2 = n \cdot p \cdot q$$

(Rincón, 2013)

APROXIMACIÓN Y CORRECCIÓN DE CONTINUIDAD

Si X es una variable aleatoria discreta, que distribuye Binomial de parámetros n y p, es decir $X \sim Bin(n, p)$ y cumple con las siguientes condiciones:

$$n > 10$$

$$n \cdot p \geq 5$$

$$n \cdot q \geq 5$$

Entonces la variable aleatoria X, puede aproximarse a una Distribución Normal, de parámetros $\mu = n \cdot p$ y $\sigma^2 = n \cdot p \cdot q$, es decir:

$$X \longrightarrow N(\mu = n \cdot p; \sigma^2 = n \cdot p \cdot q)$$

Lo que estandarizando equivale a

$$\frac{X - np}{\sqrt{npq}} \xrightarrow{d} N(0,1)$$

(Meyer, 1992 y Walpole, 2007)

Es importante señalar que, al aproximar una variable discreta a una continua, es necesario hacer un ajuste, para mejorar la aproximación. Dicho ajuste recibe el nombre de corrección de continuidad, teniéndose que:

$$P(X = a) \approx P(a - 0{,}5 \leq X \leq a + 0{,}5)$$

$$P(a \leq X \leq b) \approx P(a - 0{,}5 \leq X \leq b + 0{,}5)$$

Una vez realizada la corrección, se estandariza de la forma conocida. Además, de esta última fórmula, se pueden extrapolar los casos para probabilidades solo mayores (o iguales) o menores (o iguales) que un cierto número. (Meyer, 1992)

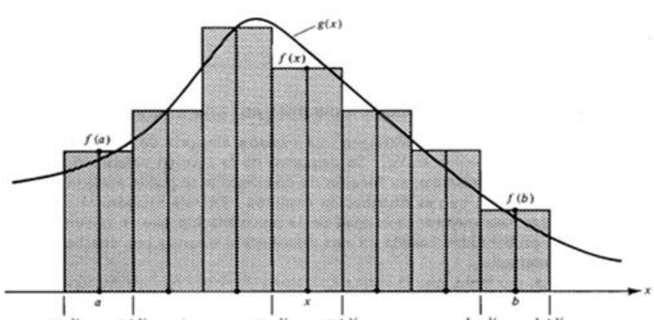


Figura 9: Aproximación de una distribución Binomial a una Normal.

ACTIVIDAD 1: COMPRENDER EL MODELO NORMAL DE PROBABILIDADES

PROPÓSITO

Se espera que los estudiantes profundicen en los conceptos clave de la distribución Normal y la manera de llegar a ellos, y que argumenten sobre cómo gráficos de barras que reflejan distribuciones Binomiales se "parecen" cada vez más a "distribuciones Normales" al aumentar el número de repeticiones. Se pretende también que dialoguen acerca de la necesidad de incorporar la distribución Normal estándar para calcular determinadas probabilidades en contexto. Del mismo modo, pueden reforzar la aproximación de una distribución Binomial por una distribución Normal en contextos determinados.

OBJETIVOS DE APRENDIZAJE

OA 3. Modelar fenómenos o situaciones cotidianas del ámbito científico y del ámbito social, que requieran el cálculo de probabilidades y la aplicación de las distribuciones Binomial y Normal.

- OA b. Resolver problemas que impliquen variar algunos parámetros en el modelo utilizado y observar cómo eso influye en los resultados obtenidos.
- OA e. Construir modelos, realizando conexiones entre variables para predecir posibles escenarios de solución a un problema, y tomar decisiones fundamentadas.
- OA i. Buscar, seleccionar, manejar y producir información matemática/cuantitativa confiable a través de la web.

ACTITUDES

- Trabajar colaborativamente en la generación, desarrollo y gestión de proyectos y la resolución de problemas, integrando las diferentes ideas y puntos de vista.
- Aprovechar las herramientas disponibles para aprender y resolver problemas.

DURACIÓN

18 horas pedagógicas

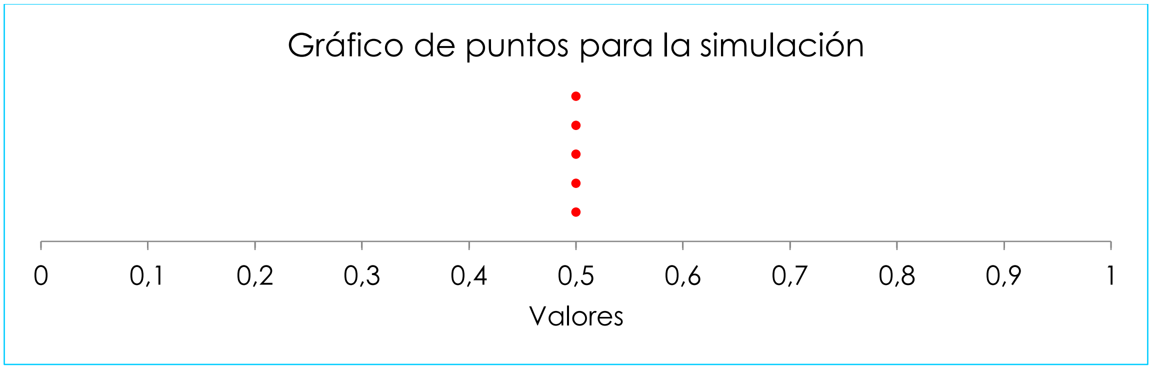
DESARROLLO

Se sugiere que trabajen colaborativamente en las siguientes actividades.

DESARROLLO Y SIGNIFICADO DE LA DISTRIBUCIÓN NORMAL

PASO 1. EL CONCEPTO DE UNA VARIABLE ALEATORIA CONTINUA

El recuadro siguiente muestra una recta numérica con subdivisión en décimas. Se marcó el número racional de $0{,}5$. Se sugiere realizar el siguiente experimento aleatorio, utilizando un generador digital de números al azar disponible en internet.



a. **Primer experimento:** Se elige números naturales de 1 a 10. Entre los números generados, se registra los números "5" –que se interpretan como "$0{,}5$"– y se pone un punto rojo en la recta numérica sobre la posición del número "$0{,}5$". Después de $n = 100$ o 200 números generados, se determina la frecuencia relativa de los "5". Contrasten la frecuencia relativa obtenida en el experimento con la probabilidad teórica correspondiente.

SOLUCIÓN:

En el apartado *Recursos y sitios web* presentes al final de esta actividad se ofrecen múltiples enlaces para generar números aleatorios utilizando plataformas en línea, sin embargo, a continuación, se recurre al programa Excel para realizar la simulación.

Abra una nueva hoja de cálculo en Excel. Luego, utilice la función

$$= ALEATORIO.ENTRE(mínimo, máximo)$$

observe que el ejercicio pide generar valores aleatorios entre 1 y 10, entonces estos corresponden al valor mínimo y máximo respectivamente, de forma que se tiene:

=ALEATORIO.ENTRE(1;10)

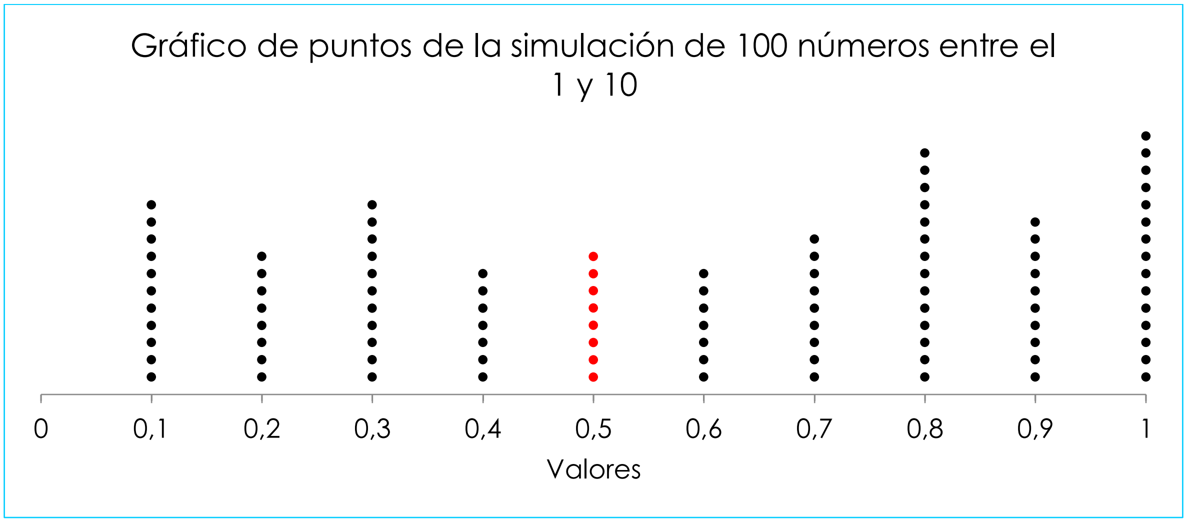
Valores obtenidos en la simulación

Nº	Valor	Nº	Valor	Nº	Valor	Nº	Valor	Nº	Valor
1	9	21	10	41	7	61	1	81	8
2	1	22	2	42	7	62	5	82	10
3	2	23	7	43	6	63	9	83	5
4	3	24	5	44	4	64	5	84	1
5	6	25	10	45	9	65	8	85	8
6	7	26	6	46	4	66	1	86	9
7	10	27	1	47	2	67	7	87	1
8	10	28	5	48	1	68	10	88	6
9	10	29	3	49	2	69	8	89	6
10	8	30	3	50	8	70	10	90	8
11	8	31	4	51	9	71	1	91	3
12	2	32	6	52	5	72	3	92	8
13	7	33	3	53	3	73	2	93	8
14	7	34	1	54	5	74	4	94	6
15	5	35	10	55	10	75	8	95	7
16	4	36	8	56	2	76	3	96	7
17	3	37	10	57	4	77	9	97	10
18	4	38	9	58	9	78	10	98	10
19	1	39	3	59	10	79	8	99	1
20	2	40	3	60	9	80	9	100	8

Luego de simular los 100 valores, se obtiene la siguiente tabla y gráfica:

Resultados de la simulación

Número	Frecuencia	Frecuencia Relativa
1	11	0,11
2	8	0,08
3	11	0,11
4	7	0,07
5	8	0,08
6	7	0,07
7	9	0,09
8	14	0,14
9	10	0,1
10	15	0,15
Total	100	1



Como es posible ver en el gráfico, el número "$0{,}5$" se obtuvo 8 veces. Al calcular la frecuencia relativa de dicho valor, se tiene:

$$f_r = \frac{8}{100} = \frac{2}{25} = 0{,}08$$

Por otro lado, la probabilidad teórica se calcula mediante la Regla de Laplace, donde el caso favorable es solo 1 (que corresponde a obtener el número 5), y los casos totales son 10 (que corresponden a los posibles valores a obtener, es decir, los números naturales del 1 al 10).

Luego, sea $A\text{: }Obtener\ un\ número\ 5$, se tiene que:

$$P(A) = \frac{1}{10} = 0{,}1$$

También se puede escribir:

$$P(Obtener\ un\ número\ 5) = \frac{1}{10} = 0{,}1$$

De esta forma, se puede concluir que los valores obtenidos mediante la simulación, y la probabilidad teórica difieren en un valor muy pequeño.

b. **Segundo experimento:** Se elige números naturales de 1 al 100. Entre los números generados, se registra los números "50" –que se interpretan como "$0{,}50$"– y se pone un punto rojo en la recta numérica sobre la posición del número "$0{,}50$", que es la misma que la de "$0{,}5$". La subdivisión de la recta numérica entre 0 y 1 se cambió en centésimas. Después de $n = 100$ o 200 números generados, se determina la frecuencia relativa de los "50". Contrasten la frecuencia relativa obtenida en el experimento con la probabilidad teórica correspondiente.

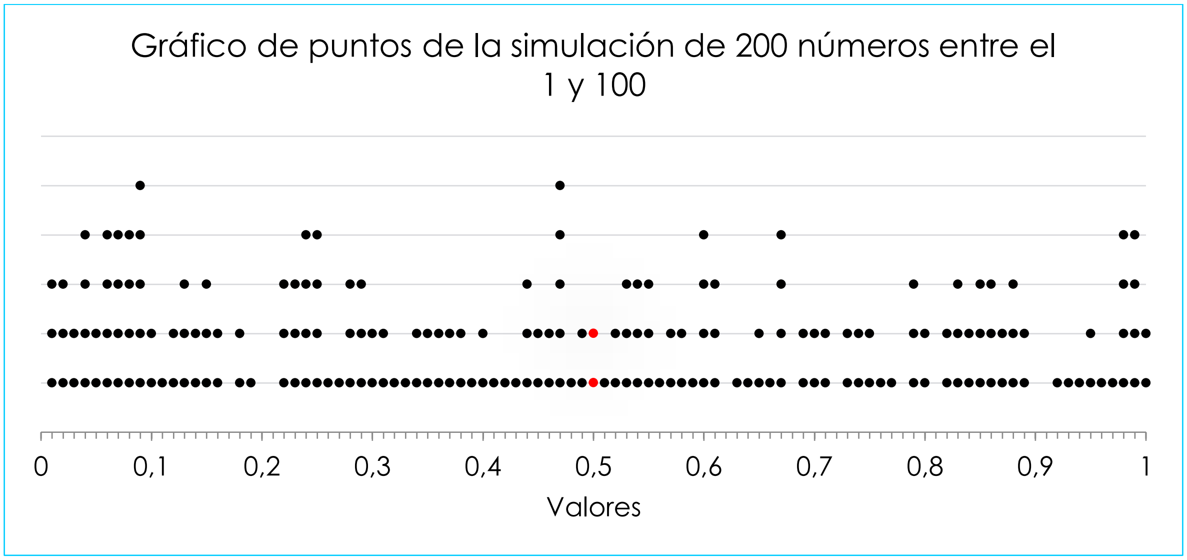
SOLUCIÓN:

Como el ejercicio pide valores entre 1 y 100, entonces ahora esos serán los valores mínimo y máximo respectivamente, de forma que se tiene los resultados que se muestran a continuación (por distribución de espacio el total de datos no se adjunta):

Valores obtenidos en la simulación

Nº	Valor	Nº	Valor	Nº	Valor	Nº	Valor	Nº	Valor	Nº	Valor	Nº	Valor	Nº	Valor	Nº	Valor	Nº	Valor
1	5	21	47	41	99	61	8	81	88	101	7	121	65	141	80	161	77	181	63
2	60	22	35	42	54	62	40	82	25	102	73	122	86	142	100	162	24	182	44
3	48	23	23	43	31	63	99	83	83	103	34	123	67	143	8	163	49	183	18
4	70	24	14	44	66	64	4	84	84	104	15	124	51	144	6	164	82	184	11
5	88	25	60	45	13	65	33	85	49	105	79	125	99	145	37	165	28	185	59
6	61	26	73	46	92	66	86	86	79	106	8	126	3	146	4	166	93	186	7
7	70	27	7	47	98	67	12	87	84	107	69	127	50	147	1	167	75	187	44
8	7	28	18	48	10	68	99	88	46	108	98	128	9	148	76	168	75	188	71
9	89	29	47	49	29	69	89	89	24	109	10	129	53	149	24	169	60	189	52
10	14	30	13	50	47	70	57	90	26	110	52	130	37	150	46	170	38	190	60
11	35	31	83	51	2	71	6	91	1	111	56	131	55	151	53	171	85	191	3
12	2	32	9	52	24	72	98	92	85	112	34	132	95	152	16	172	64	192	32
13	29	33	53	53	45	73	15	93	42	113	6	133	95	153	47	173	31	193	50
14	28	34	39	54	96	74	85	94	5	114	22	134	58	154	87	174	80	194	83
15	4	35	74	55	58	75	1	95	9	115	43	135	36	155	30	175	79	195	87
16	100	36	41	56	61	76	82	96	55	116	98	136	40	156	9	176	65	196	69
17	47	37	9	57	13	77	27	97	12	117	22	137	2	157	22	177	67	197	38
18	16	38	67	58	23	78	74	98	25	118	25	138	36	158	30	178	67	198	97
19	57	39	86	59	28	79	45	99	6	119	8	139	19	159	71	179	15	199	54
20	61	40	29	60	23	80	55	100	44	120	4	140	25	160	54	180	88	200	94

Luego de simular los 200 valores, se obtiene la siguiente gráfica:



Como es posible ver en el gráfico, el número "$0{,}5$" se obtuvo 2 veces. Al calcular la frecuencia relativa de dicho valor, se tiene:

$$f_r = \frac{2}{200} = \frac{1}{100} = 0{,}01$$

Por otro lado, la probabilidad teórica se calcula mediante la Regla de Laplace, donde el caso favorable es solo 1 (que corresponde a obtener el número 50), y los casos totales es 100 (ya que son los posibles valores por obtener, es decir, los números naturales del 1 al 100).

Luego, sea A: $Obtener\ un\ número\ 50$, se tiene que:

$$P(A) = \frac{1}{100} = 0{,}01$$

También se puede escribir:

$$P(Obtener\ un\ número\ 50) = \frac{1}{100} = 0{,}01$$

De esta forma, se puede concluir que los valores obtenidos mediante la simulación, y la probabilidad teórica son iguales.

Nota al docente: Considere que, para la simulación, los números que se generan son aleatorios, por lo que puede suceder, que el valor de f_r difiera de la probabilidad teórica. Se aconseja que esta actividad los estudiantes la realicen en parejas o grupos para comparar distintos resultados.

c. Se piensa en cambiar la graduación en $\frac{1}{1\ 000}$, $\frac{1}{10\ 000}$, $\frac{1}{100\ 000}$, $\frac{1}{1\ 000\ 000}$, $\frac{1}{10\ 000\ 000}$, $\frac{1}{100\ 000\ 000}$, $\frac{1}{1\ 000\ 000\ 000}$, ... siguiendo con cambiar la subdivisión según el mismo patrón. El generador de números al azar generaría números de 1 a $1\,000\,000\,000$, continuando con aumentar los números generados con el factor 10. ¿Cuál sería la tendencia de la probabilidad de obtener "$0{,}500$", "$0{,}5\ 000$", "$0{,}50\ 000$", "$0{,}500\ 000$", "$0{,}5\ 000\ 000$", "$0{,}50\ 000\ 000$", "$0{,}500\ 000\ 000$" etc.? Argumenten su respuesta.

SOLUCIÓN:

La tendencia de forma teórica es 0, y se puede evidenciar con los primeros casos:

$$P(Obtener\ un\ número\ 500) = \frac{1}{1000} = 0{,}001$$

$$P(Obtener\ un\ número\ 5000) = \frac{1}{10\ 000} = 0{,}0001$$

$$P(Obtener\ un\ número\ 50000) = \frac{1}{100\ 000} = 0{,}00001$$

$$P(Obtener\ un\ número\ 500000) = \frac{1}{1\ 000\ 000} = 0{,}000001$$

Esto sucede porque cada vez son más casos totales, sin embargo, la búsqueda del valor central permanece constante, siendo solo un número.

d. Lee la siguiente situación y luego responde:

El arroz es uno de los principales cereales usados en la alimentación humana en el mundo. Las distintas características que poseen los granos de arroz se asocian al hecho de que los consumidores de distintos países demanden tipos de arroz diferentes en términos de variedad, calidad y aspecto. Es así, como estas industrias deben considerar ciertos estándares de calidad, uno de ellos se relaciona con la clasificación del arroz de acuerdo con sus proporciones, ya sea según su longitud o ancho, como también por su masa. Por una parte, la masa de cada granito de arroz está entre $26{,}1\ mg$ y $30\ mg$. Además, en Chile el arroz de grano largo es aquel que posee una longitud media superior a $6{,}5mm$.

Considerando la información anterior, supongan que tienen un kilo de arroz de grano largo y seleccionan al azar un granito de arroz del paquete.

i. ¿Qué variables aleatorias podrías definir a partir del enunciado anterior? ¿Qué valores pueden tomar estas variables aleatorias?

SOLUCIÓN:

Se espera que los estudiantes puedan definir dos variables aleatorias, con relación a la masa y longitud del granito de arroz en las unidades de medida que corresponda. Luego, deben identificar que estas variables aleatorias toman valores en un subconjunto de $\mathbb{R}$.

ii. ¿Cuál es la probabilidad de encontrar un granito de arroz con una masa exacta de $28,\overline{3}\ mg$? Argumenten su respuesta.

SOLUCIÓN:

La finalidad de esta pregunta induce que la probabilidad pedida es 0, porque al tomar una bolsa de arroz de grano largo de $1\ kg$ y sacar un granito con la masa pedida, es muy poco probable y prácticamente imposible que aquello suceda, sobre todo con esos decimales requeridos.

iii. ¿Cuál es la probabilidad de encontrar un granito de arroz de longitud $6{,}777778\ mm$? Argumenten su respuesta.

SOLUCIÓN

Análogo a la respuesta anterior, ya que desde el punto de vista teórico corresponde a la probabilidad en un valor puntual del recorrido. La finalidad de esta actividad es complementar lo que se pregunta anteriormente, y concluir que la probabilidad de una variable aleatoria continua en un valor puntual del recorrido de esta será siempre cero.

Nota al docente: La actividad está diseñada para que los estudiantes la desarrollen en pareja o grupos. Además, se sugiere el uso de material didáctico para deducir desde lo práctico, lo teórico. Para ello utilice, un kilo de arroz grano largo, regla y una balanza. Solicite que seleccionen un granito al azar para que midan su masa y longitud, posteriormente pida que respondan las preguntas. Finalmente haga una puesta en común para contrastar los resultados obtenidos por los estudiantes. Luego institucionalice la probabilidad en un valor puntual del recorrido de una variable aleatoria continua.

e. Considerando que las variables aleatorias continuas toman valores en un subconjunto de los reales o en todos ellos, ¿cuál es la probabilidad $P(r)$, si r es un número real? Argumenten su respuesta.

SOLUCIÓN:

La probabilidad $P(r)$ es cero para cualquier $r \in \mathbb{R}$, ya que se pide la probabilidad en exactamente un valor puntual del recorrido.

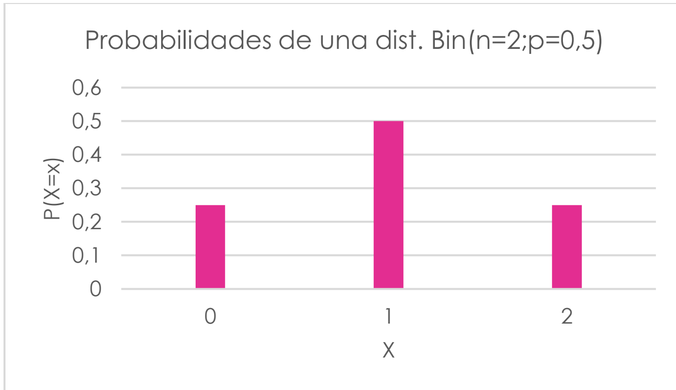
Una interpretación intuitiva corresponde a la del área de un punto, ya que las probabilidades en variables aleatorias continuas corresponden a áreas bajo la curva, por lo que el área en un valor puntual del recorrido corresponde a 0, ya que no hay distancia en el eje de las abscisas que pueda otorgarle una "base" al punto.

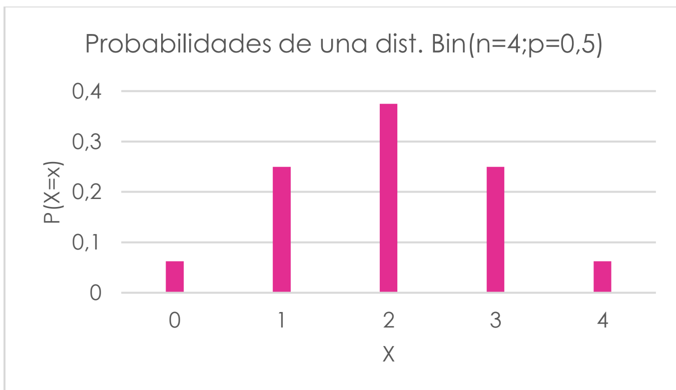
Nota al docente: La explicación matemática de esto corresponde a la integral de un punto del recorrido, de esta forma, se tiene que:

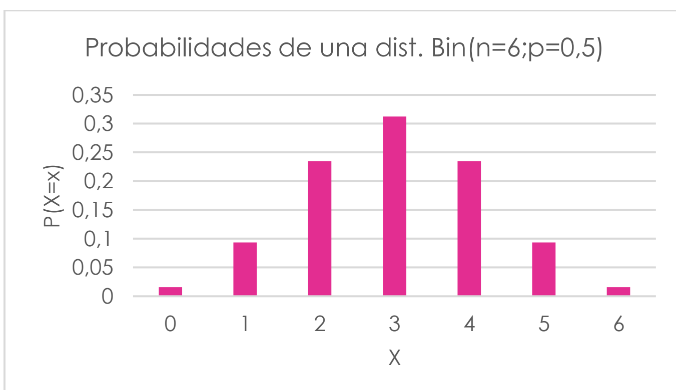
$$P(R = r) = \int_r^r f_R(t)dt = 0.$$

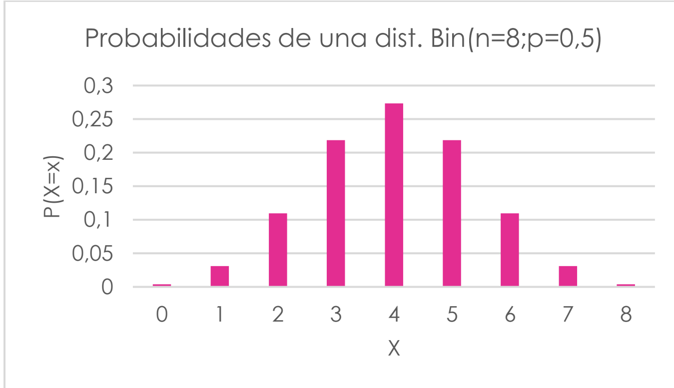
PASO 2: APROXIMÁNDONOS A LA DISTRIBUCIÓN NORMAL

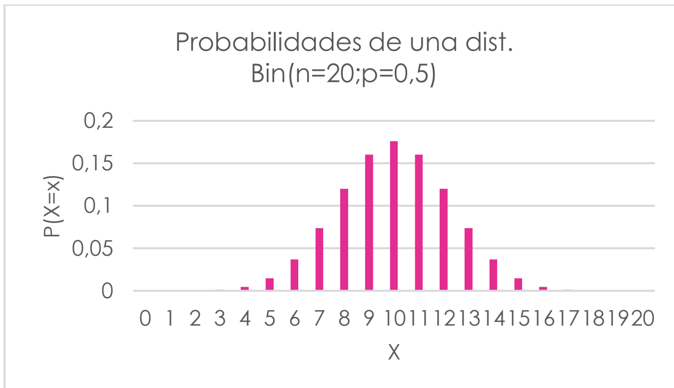
Se muestra los siguientes gráficos de barras de distribuciones Binomiales de la probabilidad de éxito de $p = 0{,}5$, incrementando cada vez el valor de "n".











a. Al aumentar el valor del número n, ¿qué cambio experimentan los gráficos de barras en cuanto a la forma, la ubicación del valor esperado y la probabilidad que corresponde al valor esperado μ? Argumenten su respuesta.

SOLUCIÓN:

El valor esperado de una variable que sigue una distribución Binomial es:

$$E(X) = \mu = n \cdot p$$

Como $p = \frac{1}{2}$, entonces se tiene que:

Cuando $n = 2$, entonces $\mu = 2 \cdot \frac{1}{2} = 1$.

Cuando $n = 4$, entonces $\mu = 4 \cdot \frac{1}{2} = 2$.

Cuando $n = 6$, entonces $\mu = 6 \cdot \frac{1}{2} = 3$.

Cuando $n = 8$, entonces $\mu = 8 \cdot \frac{1}{2} = 4$.

Cuando $n = 20$, entonces $\mu = 20 \cdot \frac{1}{2} = 10$.

Así, cuando se aumenta el valor de n, la curva va adquiriendo una forma más acampanada y las probabilidades asociadas a los valores de los extremos comienzan a tender a cero. Respecto a la ubicación del valor esperado, este se mueve de posición, encontrándose, en el centro de la figura acampanada que se forma. En este caso, la probabilidad pedida corresponde a la del valor esperado, es decir, $P(X = n \cdot p)$ que además coincide con la moda, quien es el valor que tiene mayor probabilidad. Esto sucede porque la probabilidad de éxito usada es $\frac{1}{2}$.

Nota al docente: La curva es simétrica desde un inicio, pues al considerar $p = \frac{1}{2}$, se tiene que la probabilidad más alta se obtiene en $X = \mu$. Por ello, se recomienda que el docente trabaje con distintos valores de p, porque cambiará la probabilidad más alta y forma de la gráfica, de esta forma, los estudiantes comprenderán que la Distribución Binomial no siempre es simétrica o tiene probabilidad máxima en el valor correspondiente a la mitad del recorrido.

b. Con herramientas tecnológicas digitales como Excel, GeoGebra u otros, elaboren los gráficos de barra para $n = 40$ y $n = 60$. ¿A qué forma se acercan los gráficos de barra?

SOLUCIÓN:

- Para hacerlo con Excel:

  Se debe hacer una tabla con los valores del recorrido de la Binomial, y luego utilizar la función:

$$= DISTR.BINOM.N(x; n; p; FALSO\ O\ VERDADERO)$$

  donde:

  x: número de éxitos.

  n: número de repeticiones del experimento.

  p: probabilidad de éxito asociada al experimento.

  $FALSO$: el resultado es la probabilidad de cada valor, sin acumular (función de probabilidad).

  $VERDADERO$: el resultado es la probabilidad acumulada (función de distribución acumulada).

Para poder realizar el gráfico de barras, se necesita ingresar el valor lógico $FALSO$, ya que graficará los valores de la función de probabilidad y no la función de distribución acumulada.

Así la función a ingresar es:

$$= DISTR.BINOM.N(x; 40; 0{,}5; FALSO)$$

Para esta parte del ejercicio, el único valor que varía es x. Antes de graficar, es necesario poner todos estos valores en una tabla, como la siguiente:

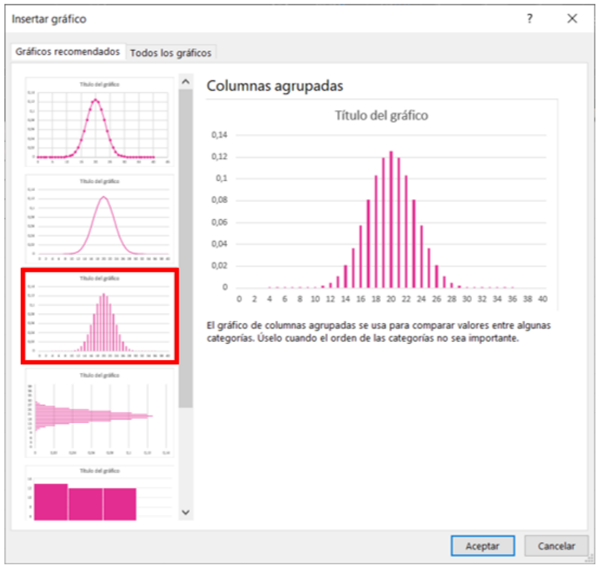
=DISTR.BINOM.N(0;40;0,5;FALSO)

Tabla de probabilidades de
$\boldsymbol{Bin(n = 40; p = 0,5)}$

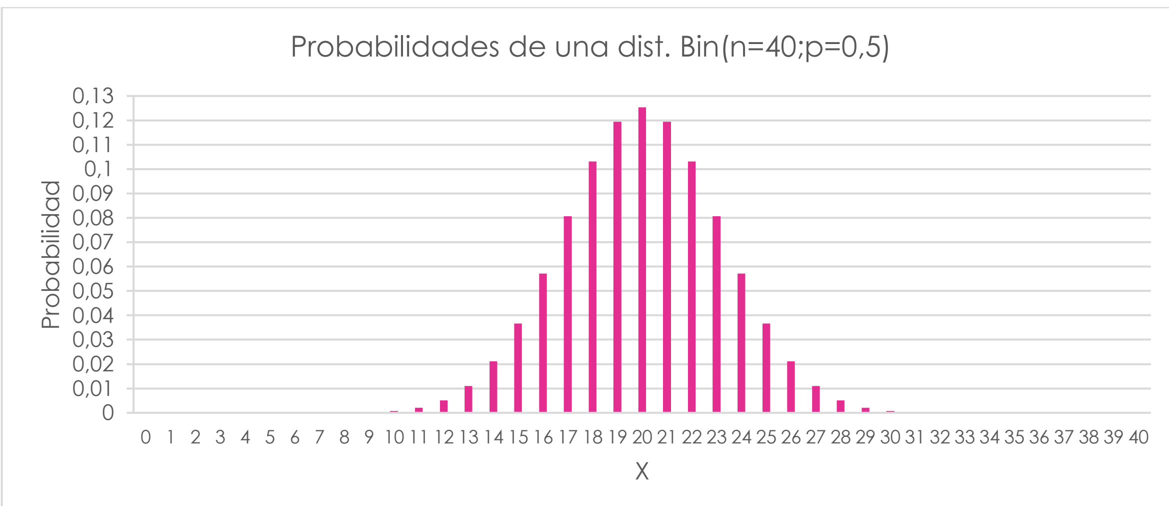
X	$P(X = x)$	X	$P(X = x)$
0	9,0949E-13	21	0,11940065
1	3,638E-11	22	0,10311875
2	7,0941E-10	23	0,08070163
3	8,9858E-09	24	0,05716365
4	8,3119E-08	25	0,03658474
5	5,9845E-07	26	0,02110658
6	3,491E-06	27	0,01094415
7	1,6956E-05	28	0,00508121
8	6,9944E-05	29	0,00210257
9	0,00024869	30	0,00077094
10	0,00077094	31	0,00024869
11	0,00210257	32	6,9944E-05
12	0,00508121	33	1,6956E-05
13	0,01094415	34	3,491E-06
14	0,02110658	35	5,9845E-07
15	0,03658474	36	8,3119E-08
16	0,05716365	37	8,9858E-09
17	0,08070163	38	7,0941E-10
18	0,10311875	39	3,638E-11
19	0,11940065	40	9,0949E-13
20	0,12537069		

Para realizar el gráfico, se selecciona la columna que tiene los éxitos buscados (X) y la de probabilidad $\big(P(X = x)\big)$, posteriormente, se hace clic en la pestaña *Insertar → Gráficos recomendados*, como se muestra a continuación:

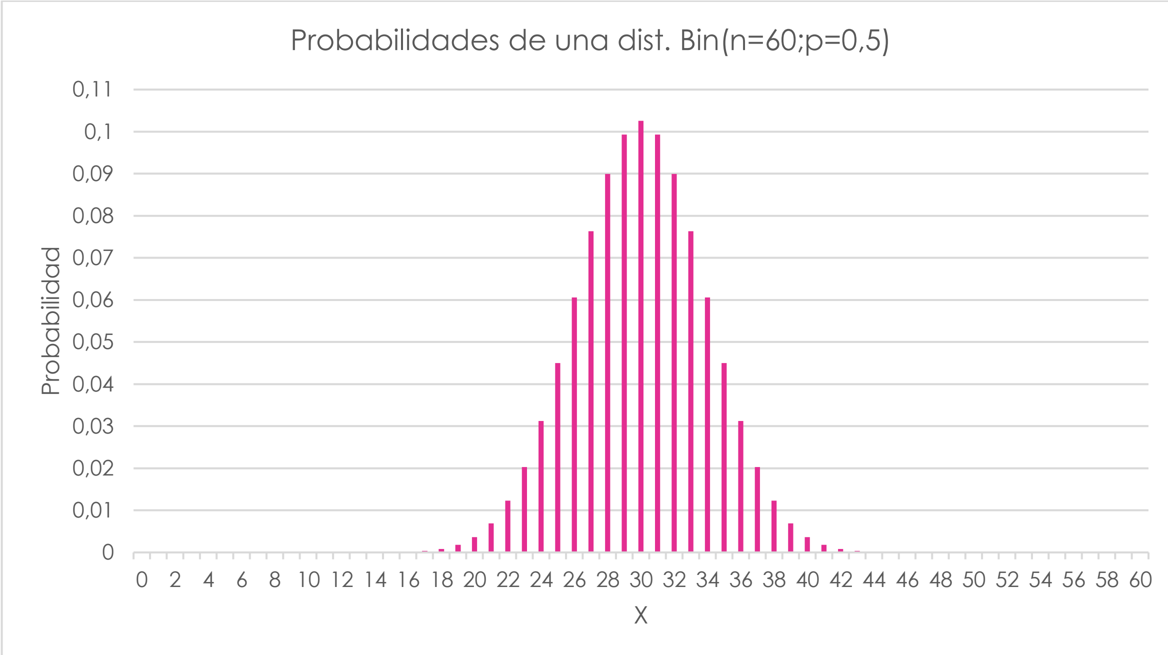
Al presionar *Gráficos recomendados* se desplegará una ventana como la siguiente, donde se debe presionar el gráfico de *Columnas agrupadas*.



Luego de seleccionar ese tipo de gráfico, se obtiene lo siguiente:

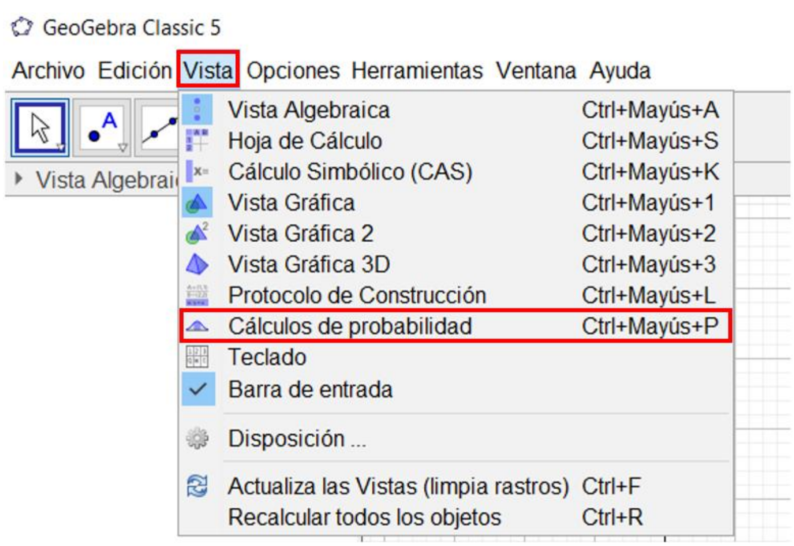


De la misma forma, se generan los datos para $n = 60$, obteniendo el siguiente gráfico.



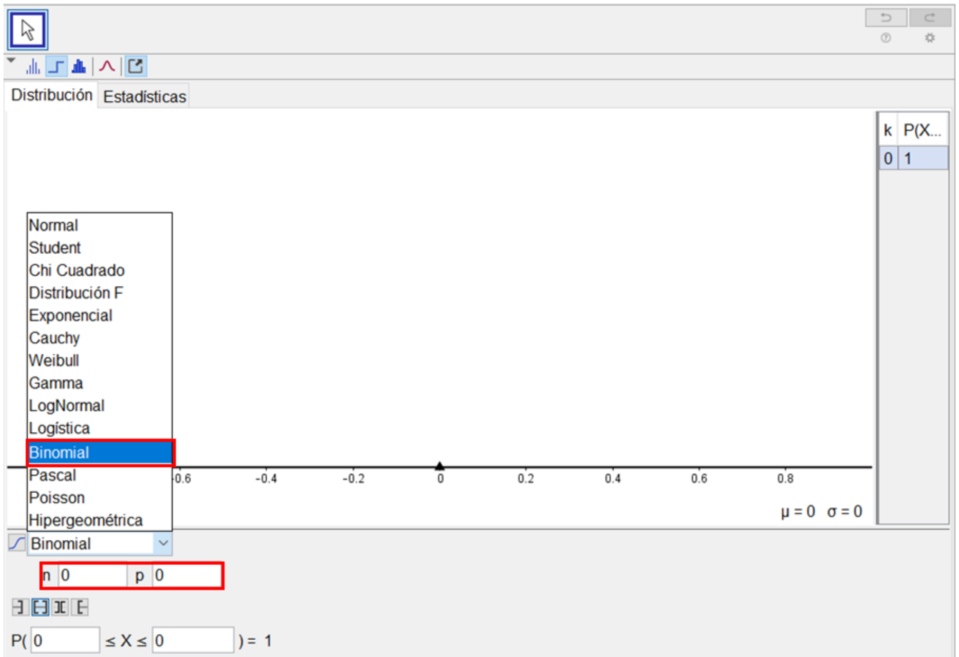
- Por otra parte, en GeoGebra:

  Para poder configurar en GeoGebra la vista de cálculo de probabilidades, se presiona la pestaña *Vista*, y se selecciona la opción *Cálculos de probabilidad*, como se muestra en la siguiente figura.

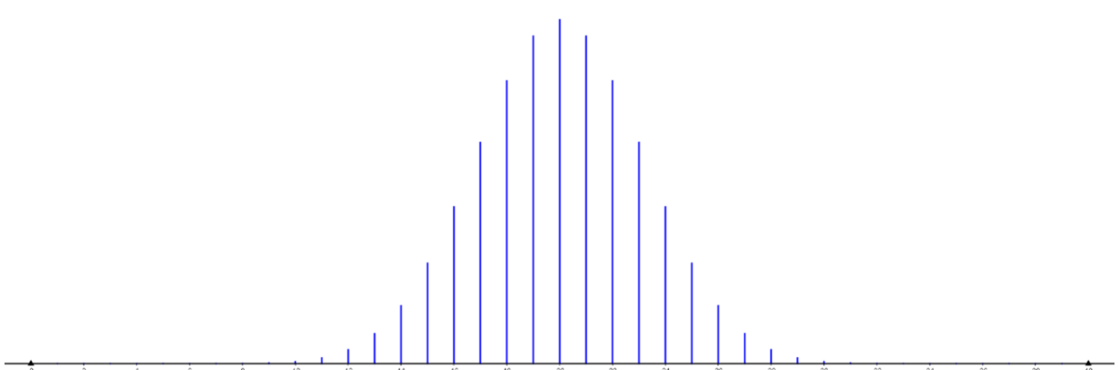


  Una vez en la sección de *Cálculos de probabilidad*, se encontrará con una ventana donde aparece el eje de las abscisas, donde se observarán los posibles valores del recorrido de la variable aleatoria. En este lugar, aparecerá la gráfica de la función de probabilidad.

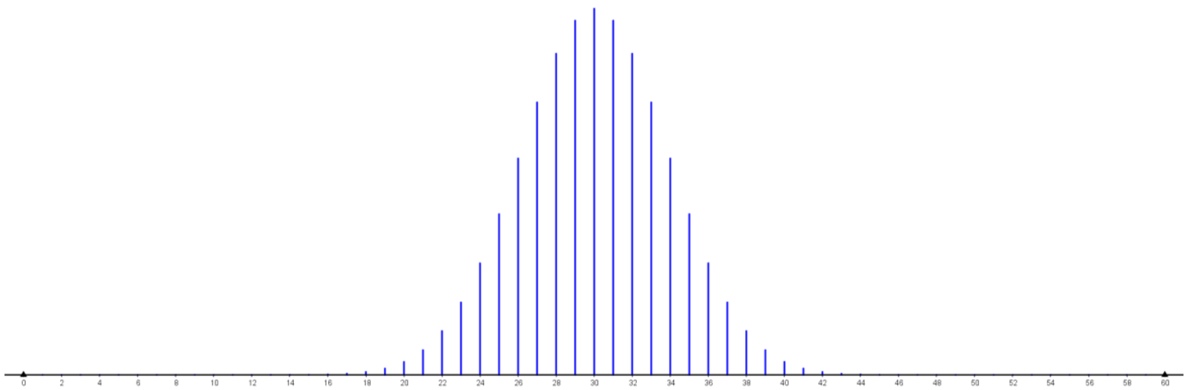
En la ventana derecha, aparecen los valores que toma el recorrido de la función, y su respectiva probabilidad. En la pestaña, se despliegan ciertos tipos de distribuciones, para nuestro caso se selecciona la *Binomial*, y bajo esta pestaña desplegable, se escriben los valores de los parámetros de la distribución, y el intervalo que se requiera.



En el ejemplo, $n = 40$ y $p = 0{,}5$, de forma que el gráfico con Geogebra queda de la siguiente forma:



De la misma forma se generan los datos para $n = 60$, obteniéndose:


0
2
4
6
8
10
12
14
16
18
20
22
24
26
28
30
32
34
36
38
40
42
44
46
48
50
52
54
56
58
60

PASO 3. IDENTIFICANDO LOS PARÁMETROS DE UNA DISTRIBUCIÓN NORMAL

DESCRIPCIÓN DE LA SUBACTIVIDAD

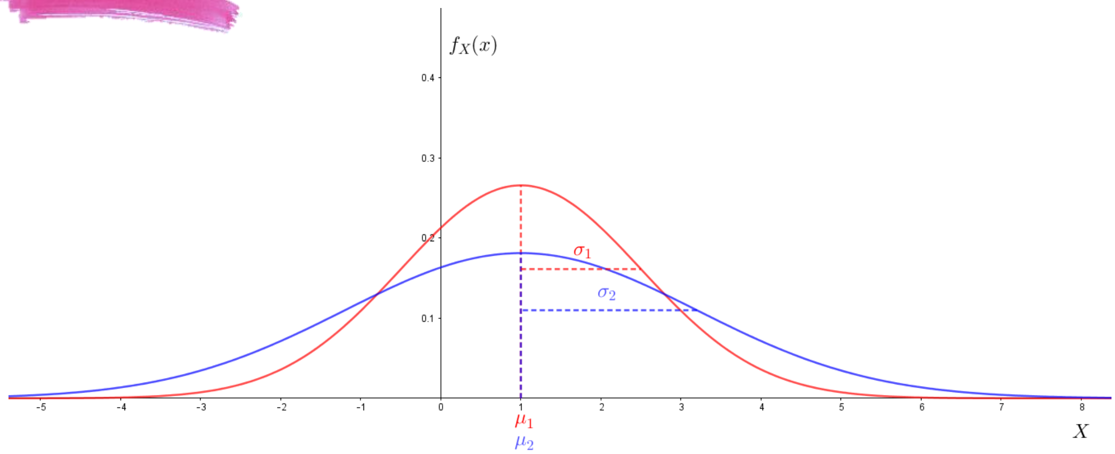
- Los ítems A y B están diseñadas a partir de problemas bajo contextos determinados. Aquí, los estudiantes identifican los parámetros de la Distribución Normal y comparan poblaciones a partir de ellas. El ítem C muestra que sucede gráficamente cuando $\mu = 0$ y $\sigma = 1$.
- El ítem A permite que los estudiantes identifiquen cuándo un parámetro es mayor, menor o igual que otro en una gráfica, a partir de los colores señalados para cada campana. Además, comprenden de qué forma influye el valor del parámetro en la gráfica de la Distribución Normal.
- El ítem B permite que los estudiantes apliquen la identificación de parámetros a problemas puestos en contexto, donde no solo deben comparar los valores que dan forma a las campanas, sino que también deben extraer información del enunciado para plantear su resolución y tomar decisiones a partir de su respuesta.
- El ítem C permite que los estudiantes descubran qué pasa al variar el valor de los parámetros en la campana de Gauss, y como el valor de estos va a afectar en la forma y escala que adopta la campana. Estas preguntas sirven para que los estudiantes puedan reflexionar y visualizar que sucede en los casos pedidos, siempre explicando con sus palabras lo que comprenden de ello.

DESAROLLO

La media μ y varianza σ^2 son los dos parámetros de los que depende una distribución Normal, y estos condicionan la ubicación y forma de la campana de Gauss.

A. Analiza las siguientes gráficas, y reconoce la relación ($<$, $>$ o $=$) que existe entre la media y desviación estándar de ambas. Para ello considera que μ_1 y σ_1 es la media y desviación estándar de la campana de color azul y que μ_2 y σ_2 son la media y desviación estándar de la campana de color rojo.

Ejemplo

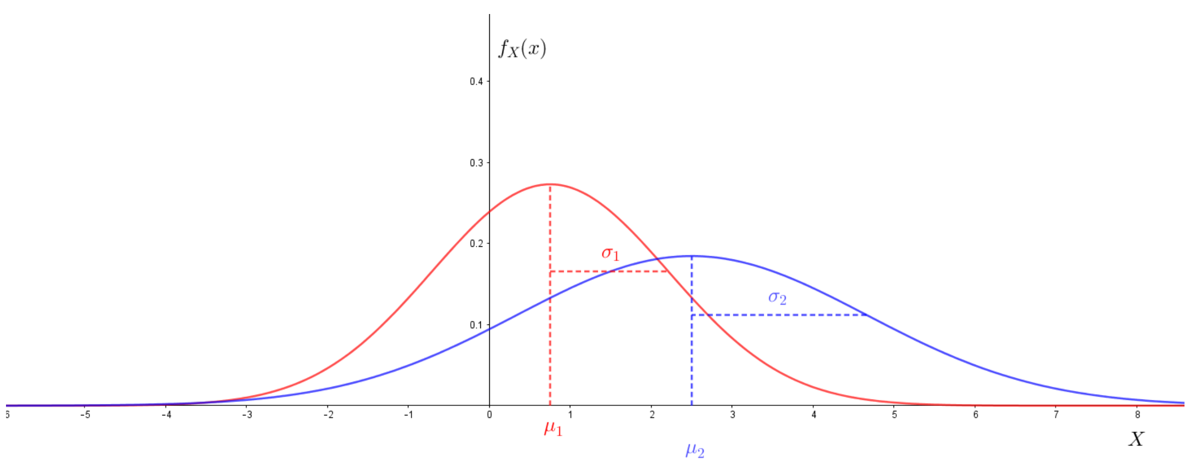


En este caso, como μ_1 y μ_2 coinciden en el mismo valor sobre el eje X, diremos que son iguales. Respecto a la desviación estándar, en la campana de color rojo, se visualiza que su abertura desde el eje $X = \mu$, hasta el cambio de curvatura es mayor que la que existe en la campana azul, de esta forma podemos concluir que σ_1 es mayor que σ_2. Por lo tanto:

$$\mu_1 \underline{=} \mu_2$$

$$\sigma_1 \underline{<} \sigma_2$$

a.



$$\mu_1 \; \_\_\_ \; \mu_2$$

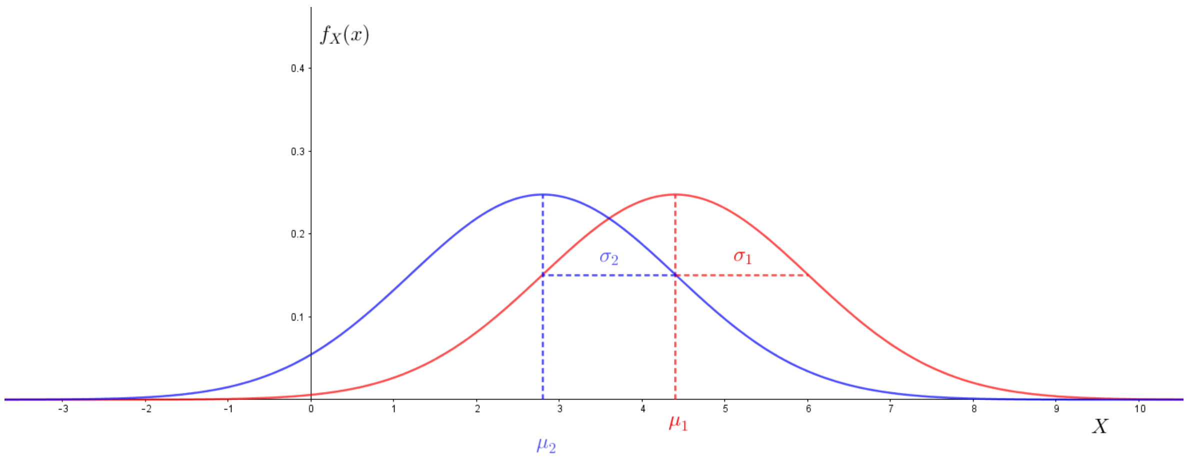
$$\sigma_1^2 \; \_\_\_ \; \sigma_2^2$$

SOLUCIÓN:

$$\mu_1 \underline{\ <\ } \mu_2$$

$$\sigma_1^2 \underline{\ <\ } \sigma_2^2$$

b.



$$\mu_1 \underline{\qquad} \mu_2$$

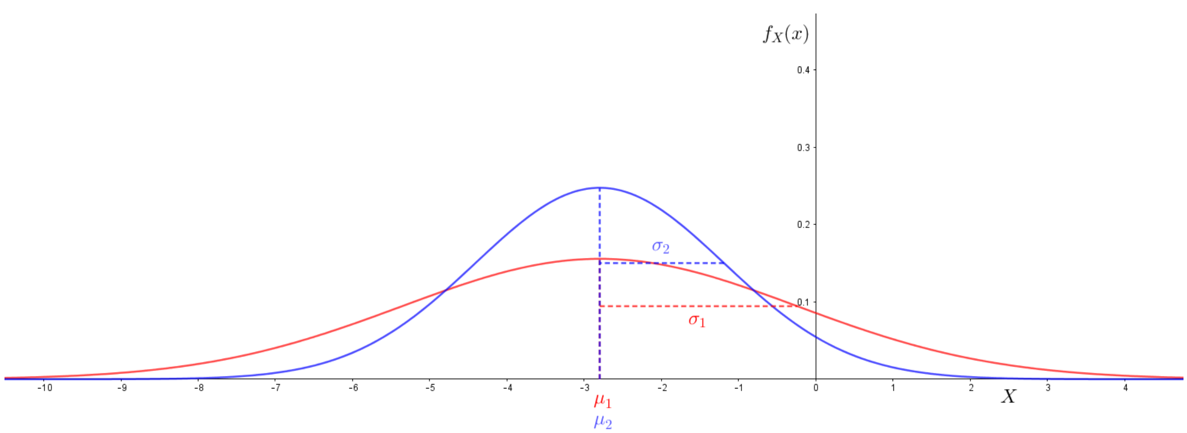
$$\sigma_1^2 \underline{\qquad} \sigma_2^2$$

SOLUCIÓN:

$$\mu_1 \underline{\ >\ } \mu_2$$

$$\sigma_1^2 \underline{\ =\ } \sigma_2^2$$

c.



$$\mu_1 \underline{\qquad} \mu_2$$

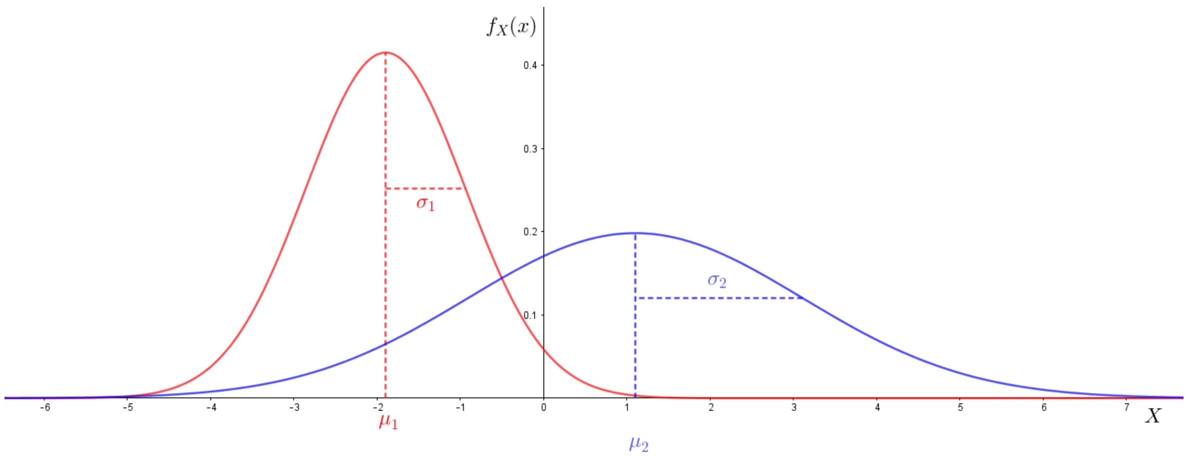
$$\sigma_1^2 \underline{\qquad} \sigma_2^2$$

SOLUCIÓN:

$$\mu_1 \underline{\ = \ } \mu_2$$

$$\sigma_1^2 \underline{\ > \ } \sigma_2^2$$

d.



$$\mu_1 \_\_\_ \mu_2$$

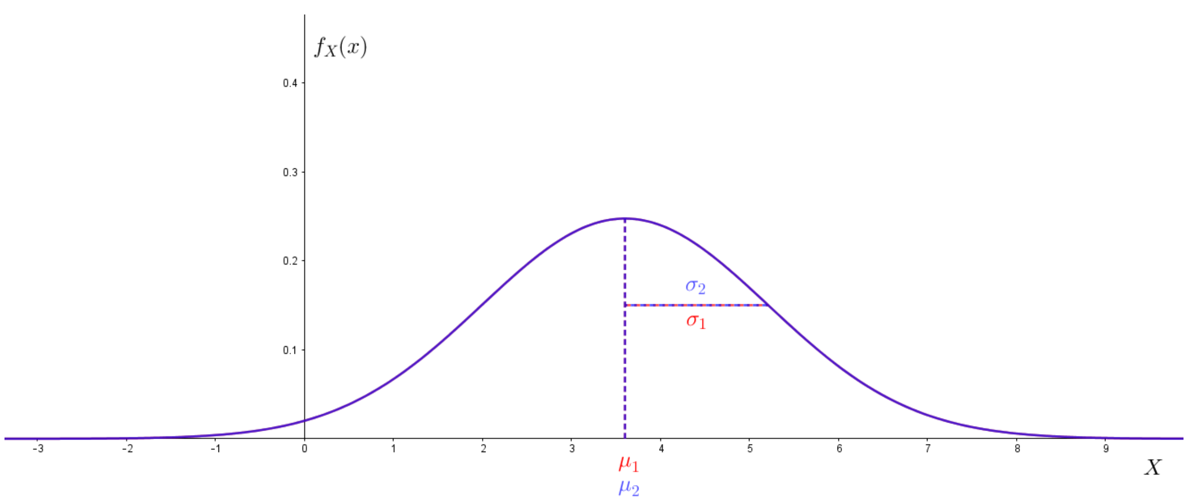
$$\sigma_1^2 \_\_\_ \sigma_2^2$$

SOLUCIÓN:

$$\mu_1 \underline{\ < \ } \mu_2$$

$$\sigma_1^2 \underline{\ < \ } \sigma_2^2$$

e.



$$\mu_1 \_\_\_ \mu_2$$

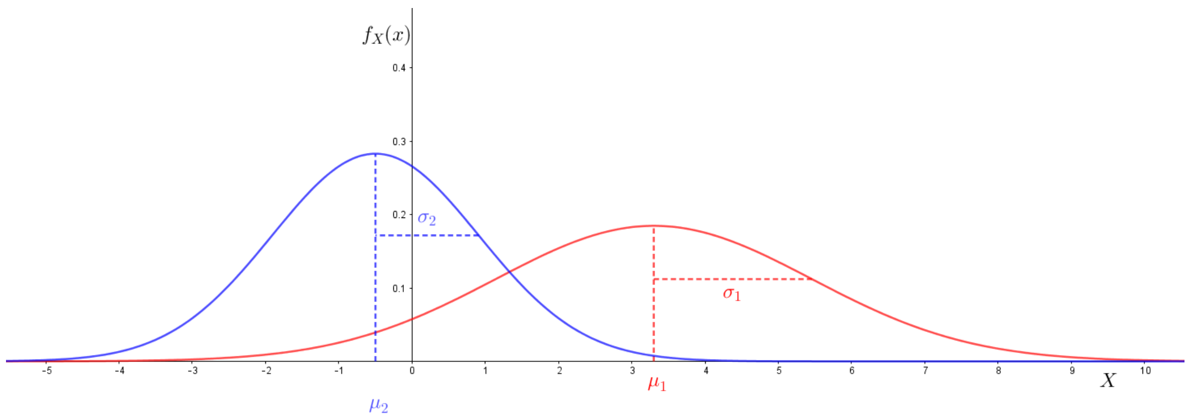
$$\sigma_1^2 \;\_\_\_\; \sigma_2^2$$

<u>SOLUCIÓN:</u>

$$\mu_1 \;\underline{\;=\;}\; \mu_2$$

$$\sigma_1^2 \;\underline{\;=\;}\; \sigma_2^2$$

f.



$$\mu_1 \;\_\_\_\; \mu_2$$

$$\sigma_1^2 \;\_\_\_\; \sigma_2^2$$

<u>SOLUCIÓN:</u>

$$\mu_1 \;\underline{\;>\;}\; \mu_2$$

$$\sigma_1^2 \;\underline{\;>\;}\; \sigma_2^2$$

<u>Nota</u>

En algunos enunciados se entrega la desviación estándar y varianza en conjunto. Se sugiere poner atención a los parámetros, para que no haya confusiones o errores.

B. Resuelve los siguientes problemas.

a. El animal más alto del mundo es la jirafa, vive en sabanas, pastizales y bosques abiertos repartidos en casi todo el continente africano y su dieta alimenticia se compone principalmente de las hojas de la acacia. La subespecie denominada *jirafa de Sudáfrica* habita principalmente en esa región, también vive en Botsuana, Zimbabue y Mozambique. Su altura sigue una Distribución Normal de parámetros $\mu =$

$5{,}22\ m$ y $\sigma = 0{,}2\ m$. Por otro lado, la *jirafa Masai* o *jirafa del Kilimanjaro* habita solamente en Kenia y Tanzania, su altura igualmente distribuye Normal, con media $\mu = 5{,}12\ m$, y desviación estándar $\sigma = 0{,}25\ m$.

i. En tu cuaderno, esboza las campanas de ambas poblaciones en un mismo eje.

SOLUCIÓN:

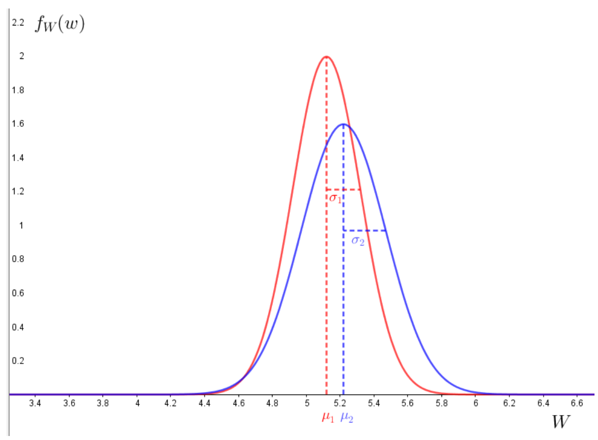
Sean:

$$W_1: Altura\ de\ la\ jirafa\ Masai\ o\ del\ Kilimanjaro, en\ m.$$

$$W_2: Altura\ de\ la\ jirafa\ de\ Sudáfrica, en\ m.$$

Luego $W_1 \sim N(\mu_1 = 5{,}12; \sigma_1^2 = 0{,}25^2)$ y $W_2 \sim N(\mu_2 = 5{,}22; \sigma_2^2 = 0{,}2^2)$.

Al realizar la gráfica, se tiene que la campana roja representa a la población W_1, mientras que la azul corresponde a W_2.



ii. ¿Qué población de jirafas posee mayor variabilidad en su altura, respecto de la media?

SOLUCIÓN:

Al comparar las desviaciones estándar de ambos grupos de jirafas se concluye que la población con mayor variabilidad de altura es la subespecie Masai o del Kilimanjaro, pues $\sigma = 0{,}25\ m$ es más alto en contraste con la desviación de la jirafa de Sudáfrica que corresponde a $\sigma = 0{,}2\ m$.

iii. En media, ¿qué población de jirafas tiene la menor altura?

SOLUCIÓN:

Al contrastar las medias de ambas subespecies de jirafas se concluye que la población con menor altura es la *Masai* o *del Kilimanjaro*, pues su media es $\mu = 5{,}12\ m$ que resulta ser inferior comparada con la de la *jirafa de Sudáfrica* correspondiente a $\mu = 5{,}22\ m$.

b. Josefa y sus amigos desean comer pizza, quieren que sea de entrega rápida pues tienen bastante hambre. No tienen claro a qué pizzería pedir, sin embargo, cerca de su casa hay dos locales que realizan delivery el primero es "La pizzería de la esquina", tiene un tiempo de producción y entrega que sigue una Distribución Normal de parámetros $\mu = 57\ min$ y varianza de $\sigma^2 = 25\ min^2$, mientras que el tiempo del otro local, "Pizzamanía", también sigue una distribución Normal de parámetros $\mu = 54\ min$ y desviación estándar $\sigma = 5{,}92\ min$.

i. En tu cuaderno, esboza ambas campanas en un mismo eje.

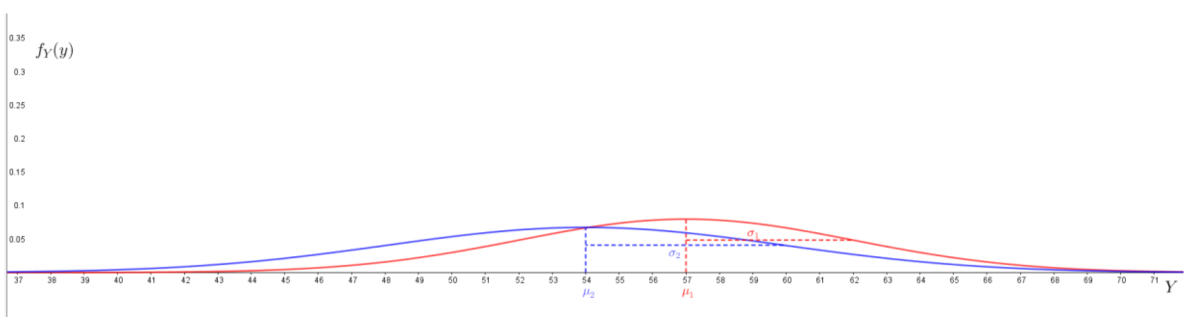
SOLUCIÓN:

Sean:

$$Y_1 = Tiempo\ en\ minutos\ que\ demora\ "La\ pizzería\ de\ la\ esquina"$$
$$en\ producir\ y\ entregar\ la\ pizza.$$

$$Y_2 = Tiempo\ en\ minutos\ que\ demora\ "Pizzamanía"$$
$$en\ producir\ y\ entregar\ la\ pizza.$$

De modo que $Y_1 \sim N(\mu_1 = 57; \sigma_1^2 = 25)$ está representada por la campana roja y $Y_2 \sim N(\mu_2 = 54; \sigma_2^2 = 5{,}92^2)$ está descrita por la curva de color azul. En virtud de aquello, la gráfica es la siguiente:



ii. En media, ¿qué pizzería demora menos tiempo en realizar el delivery? ¿Qué elementos te permitieron identificarlo? Argumenta tu respuesta.

SOLUCIÓN:

El local que demora menos tiempo es "Pizzamanía", pues tiene menor tiempo medio de delivery en comparación a "La pizzería de la esquina". También es posible evidenciarlo en la gráfica del apartado anterior, ya que el eje de simetría de la campana azul está a la izquierda de la roja.

Para la segunda pregunta, se espera que los estudiantes identifiquen lo pedido de cualquiera de las dos maneras.

iii. ¿Qué pizzería tiene menor variabilidad respecto al tiempo medio de producción y entrega? ¿Cómo identificaste que era ese local? Argumenta tu respuesta.

SOLUCIÓN:

Primero, se calcula la desviación estándar del local "La pizzería de la esquina", que corresponde a $\sigma = \sqrt{25\,min^2} = 5\,min$. Así, el local que tiene menos variabilidad de tiempo es "La pizzería de la esquina", pues posee menor desviación estándar comparado con "Pizzamanía".

Su explicación también puede justificarse usando las campanas tal como la construida en i., donde se observa que la campana roja es más alta que la azul, lo que indica que posee menor dispersión de datos. Es decir, la "pizzería de la esquina" es la que tiene menor variabilidad respecto al tiempo medio de producción y entrega.

iv. Considerando que la pizzería favorita de Josefa es la que demora más tiempo, porque elabora pizzas más ricas. ¿A qué local encargarías tú, si fueses amiga o amigo de Josefa? Argumenta tu respuesta.

SOLUCIÓN:

Luego de resolver las preguntas anteriores, se espera que los estudiantes tomen una decisión sustentada estadísticamente.

c. Tres grupos de distintos géneros musicales realizan un concierto el mismo día en Santiago de Chile. Los tiempos de duración del evento de las tres bandas siguen una distribución Normal, de forma que, el tiempo de duración del grupo

de rock tiene parámetros $\mu = 150\, min$ y desviación estándar de $\sigma = 10\, min$. Por su parte, el evento del grupo de pop posee una media $\mu = 146\, min$ y varianza de $\sigma^2 = 64\, min^2$. Por último, el tiempo de duración del grupo de cumbia distribuye Normal con parámetros $\mu = 152\, min$ y $\sigma^2 = 36\, min^2$.

i. Esboza las tres campanas en un mismo eje.

SOLUCIÓN:

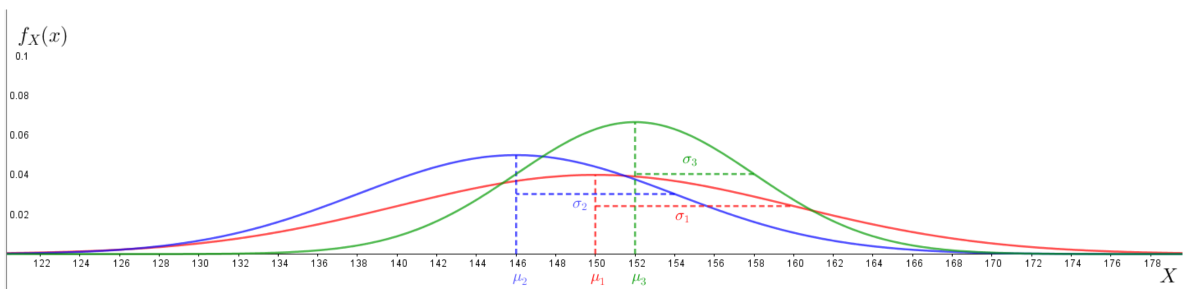
Sean:

$$X_1: Tiempo\ de\ duración\ del\ concierto\ del\ grupo\ de\ rock, en\ minutos.$$

$$X_2: Tiempo\ de\ duración\ del\ concierto\ del\ grupo\ de\ pop, en\ minutos.$$

$$X_3: Tiempo\ de\ duración\ del\ concierto\ del\ grupo\ de\ cumbia, en\ minutos.$$

De modo que, $X_1 \sim N(\mu_1 = 150; \sigma_1^2 = 10^2)$, $X_2 \sim N(\mu_2 = 146; \sigma_2^2 = 64)$ y $X_3 \sim N(\mu_3 = 152; \sigma_3^2 = 36)$. Las campanas se representan por el color rojo, azul y verde respectivamente.



ii. Identifica qué grupo tiene el menor tiempo medio de duración del concierto.

SOLUCIÓN:

El grupo que tiene menor tiempo medio de duración es el que canta pop, ya que es el menor valor de los tres entregados.

Adicionalmente es posible evidenciar en el gráfico de la pregunta anterior que la menor media corresponde al valor que se encuentra más a la izquierda que las medias de las otras poblaciones, en este caso, la duración del concierto del grupo de pop.

iii. ¿Qué grupo tiene mayor variabilidad respecto al tiempo medio de duración del concierto?

SOLUCIÓN:

Antes de comparar, se obtienen las desviaciones estándar. Para el grupo de pop, $\sigma = \sqrt{64\,min^2} = 8\,min$. Por otro lado, para el grupo de cumbia, $\sigma = \sqrt{36\,min^2} = 6\,min$. Así, la mayor desviación estándar es la del grupo de rock, por lo tanto, es el que tiene mayor variabilidad en el tiempo de duración del concierto respecto a la media.

También es posible evidenciarlo en la gráfica de la pregunta i., ya que la campana roja es más plana y baja que las otras.

iv. Supón que a Matías le gustan los tres grupos que se presentan y no tiene claro a cuál evento asistir, solo sabe que quiere ir al concierto que sea de mayor duración media. Indica a qué concierto le aconsejarías ir. Argumenta tu respuesta.

SOLUCIÓN:

Se espera que los estudiantes aconsejen ir al concierto de cumbia, ya que es el que más duración media tiene.

C. A partir de la información analizada en los ejercicios anteriores, y utilizando el enlace https://www.geogebra.org/m/fdvyufzw responde:

a. Si $\mu_1 = \mu_2$ ¿se puede concluir que las campanas tendrán el mismo eje de simetría? Argumenta tu respuesta.

SOLUCIÓN:

Efectivamente tendrán el mismo eje de simetría si las medias son iguales. La forma de visualizarlo puede ser con el ejercicio c. o e. del ítem A, o moviendo los deslizadores en el enlace puesto en el enunciado.

b. Al variar el valor de μ, en cualquiera de las dos campanas, se produce un cambio en la gráfica. ¿En qué afecta el valor de la media a la gráfica de la distribución?

SOLUCIÓN:

El valor de μ afectará en la ubicación de la campana en el eje X, si el valor es positivo, el centro de la campana estará en el lado derecho del eje, si el valor es negativo, el centro estará en el lado izquierdo del eje. Por último, si el valor es cero, la campana estará centrada en $X = 0$. Esto es posible evidenciarlo comparando las desigualdades obtenidas en la subactividad anterior, o en el enlace adjunto.

c. Al variar el valor de σ en cualquiera de las dos campanas produce un cambio en la gráfica. ¿En qué afecta el valor de la desviación estándar a la gráfica de la distribución?

SOLUCIÓN:

El valor de σ produce cambios en la abertura de la campana y su forma, siendo más puntiaguda o aplanada. A medida que $\sigma > 1$, la campana se hace más aplanada y abierta, lo que se traduce en una dispersión mayor de los datos. Por el contrario, si $\sigma < 1$, la campana se hace más puntiaguda, y menos ancha, lo que explica en un nivel de dispersión de datos menor. Esto es posible evidenciarlo comparando las desigualdades obtenidas en la subactividad anterior, o en el enlace adjunto.

d. ¿Qué sucede con la distancia entre el eje de simetría $X = \mu$ hasta la primera desviación estándar de las campanas cuando $\sigma_1 > \sigma_2$?

SOLUCIÓN:

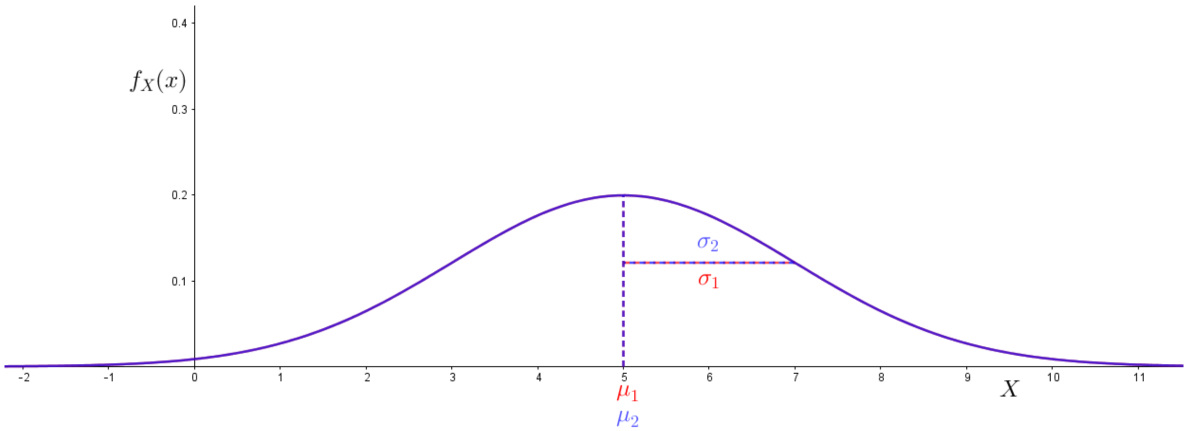
Si $\sigma_1 > \sigma_2$, entonces se puede decir que la campana roja será más plana respecto a la campana azul, ya que al ser mayor el valor de la desviación estándar, la dispersión de datos será mayor. Esto es posible visualizarlo en los ejercicios c. o f. del ítem A, o en el enlace adjunto.

e. Posiciona los deslizadores μ_1 y μ_2 en el mismo valor, que sea distinto de 0. Realiza el mismo procedimiento para σ_1 y σ_2, sin que el valor escogido sea 1. Una vez se tengan las campanas superpuestas, y considerando que ahora las campanas son una sola, varía el valor de μ_2 hacia el $X = 0$, ¿qué cambio observas en la campana? Argumenta tu respuesta.

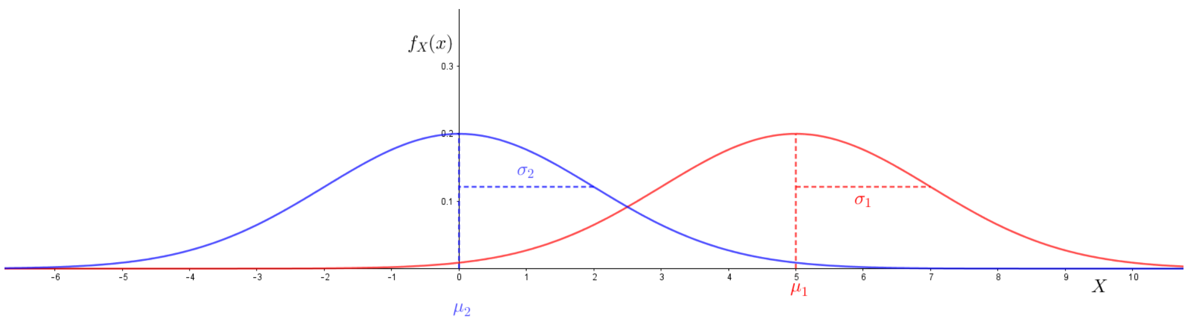
SOLUCIÓN:

Para esta actividad es imprescindible el uso del enlace, para que el software ayude a los estudiantes a visualizar de forma inmediata lo que sucede cuando se varía el valor de μ. En este caso, se ubicará $\mu_1 = \mu_2 = 5$ y $\sigma_1 = \sigma_2 = 2$.

Al inicio, las campanas deben lucir así:



Ahora, deslizando μ_2 (media de la campana azul) hasta 0, se tiene

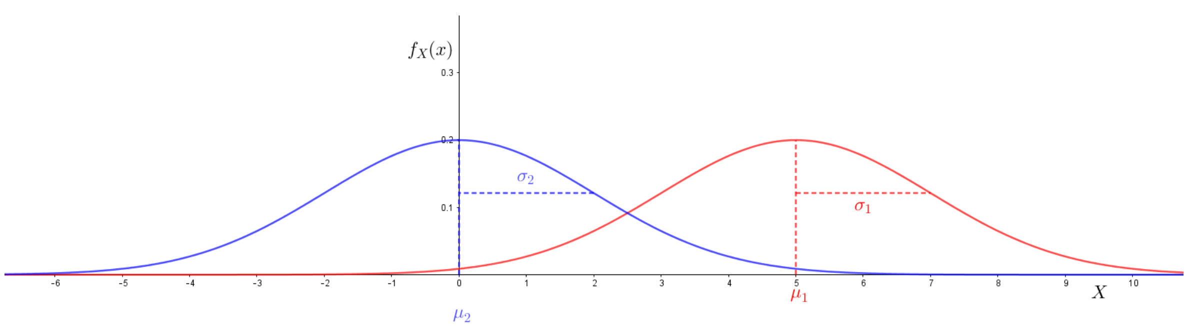


Como es posible ver, se produce una traslación del eje de simetría de la campana desde el valor $X = 5$ hasta $X = 0$.

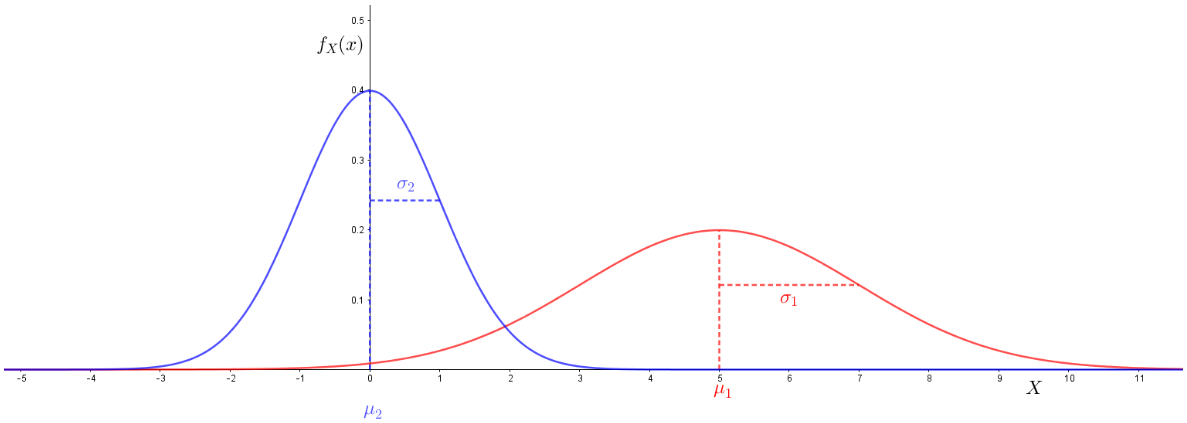
f. Ahora, sin mover los deslizadores para la media, varía el deslizador σ_2 hasta el valor 1, ¿Qué cambio observas en la campana? Argumenta tu respuesta.

SOLUCIÓN:

Una vez realizada la pregunta e., se tiene la siguiente gráfica:



Al deslizar σ_2 hasta el valor 1, se tiene



Como el valor de σ se redujo, la dispersión de datos también disminuyó, ajustándose a un cambio muy conocido en la Distribución Normal.

g. A partir de lo anterior, ¿qué transformación algebraica se debe realizar para que el valor esperado μ cambie a la posición $X = 0$?

SOLUCIÓN:

Se debiese restar μ, como se pudo ver en el ejercicio e. de este ítem.

PASO 4. UN TEOREMA BASTANTE UTILIZADO

DESCRIPCIÓN DE LA SUBACTIVIDAD:

- Esta subactividad permite que los estudiantes ejerciten y apliquen el cálculo de probabilidades con la tabla Z. Se enfatiza en la necesidad de resolver diferentes expresiones de probabilidad, considerando que los estudiantes de tercero medio (y algunos de cuarto) probablemente no conozcan la Distribución Normal, ni su estandarización.

DESARROLLO

Estandariza las siguientes variables aleatorias, expresándolas en puntuación Z, y luego calcula la probabilidad indicada en cada caso. Además, realiza el registro gráfico de lo que se está calculando. asociada a ese valor.

Ejemplo

La fórmula para poder estandarizar una variable aleatoria X con distribución $N(\mu,\sigma^2)$ a una variable aleatoria Z con distribución $N(0,1)$ es:

$$Z = \frac{X-\mu}{\sigma}$$

Ejemplo:

$P(X \leq 52)$ si $X \sim N(\mu = 50; \sigma^2 = 64)$

SOLUCIÓN:

Si $\sigma^2 = 64$, entonces $\sigma = 8$, así

$$P(X \leq 52) = P\left(\frac{X-50}{8} \leq \frac{52-50}{8}\right)$$

$$\Rightarrow P(X \leq 52) = P(Z \leq 0{,}25)$$

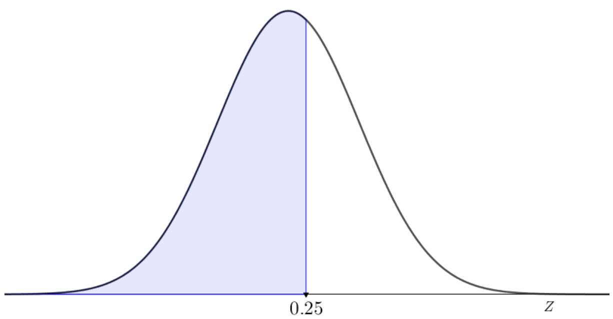
Ahora, para calcular $P(Z \leq 0{,}25)$, solo se utiliza la tabla Z, considerando que en las filas están las décimas, y en las columnas las centésimas del número que se está buscando, luego $P(Z \leq 0{,}25) = 0{,}5987$, como se muestra en la imagen adjunta.

Décimas

Centésimas

z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389

La campana asociada a la probabilidad buscada corresponde a la siguiente.



Nota: Se hace importante recordar que la tabla indica el área bajo la curva Normal estándar, que corresponde a $\phi(z) = P(Z \leq z)$, para valores de z que varían entre $-3{,}99$ y $3{,}99$. Bajo el $-3{,}99$ la probabilidad será 0, y sobre el $3{,}99$ la probabilidad será 1.

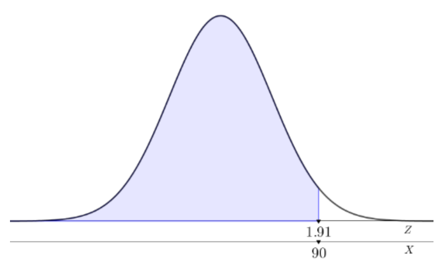
a. $P(X \leq 90)$ si $X \sim N(\mu = 82; \sigma^2 = 17{,}5)$

SOLUCIÓN:

Si $\sigma^2 = 17{,}5$, entonces $\sigma = \sqrt{17{,}5}$, así

$P(X \leq 90) = P\left(\frac{X-82}{\sqrt{17{,}5}} \leq \frac{90-82}{\sqrt{17{,}5}}\right)$

$= P(Z \leq 1{,}91)$



z	0	0.01	0.02
1.8	0.9641	0.9649	0.9656
1.9	0.9713	0.9719	0.9726
2.0	0.9772	0.9778	0.9783

Ahora, para calcular $P(Z \leq 1{,}91)$, solo se utiliza la tabla Z, de forma que:

$P(Z \leq 1{,}91) = \phi(1{,}91)$

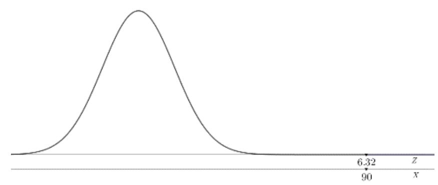
$= 0{,}9719$

Por lo tanto, $P(X \leq 90) = 0{,}9719$

b. $P(X > 22)$ si $X \sim N(\mu = 12; \sigma^2 = 2{,}5)$

SOLUCIÓN:

Si $\sigma^2 = 2{,}5$, entonces $\sigma = \sqrt{2{,}5}$, así



$P(X > 22) = 1 - P(X \leq 22)$

$= 1 - P\left(\frac{X-12}{\sqrt{2{,}5}} \leq \frac{22-12}{\sqrt{2{,}5}}\right)$

$= 1 - P(Z \leq 6{,}32)$

Ahora, para calcular $1 - P(Z \leq 6{,}32)$, se tiene que:

$1 - P(Z \leq 6{,}32) = 1 - \phi(6{,}32)$

$= 1 - 1$

$= 0$

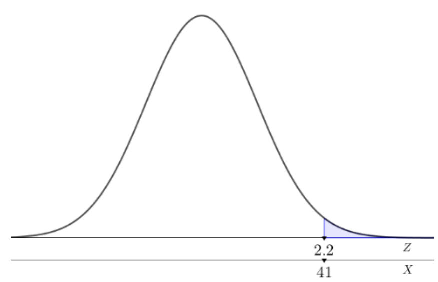
Por lo tanto, $P(X > 22) = 0$

Nota al docente: Para este ejercicio, el estudiante debe recordar la propiedad del complemento, es decir, $P(X > a) = 1 - P(X \leq a)$. Además, deben identificar que cualquier probabilidad mayor a 3,99 siempre es 1, por lo que no es necesario utilizar la tabla Z.

c. $P(X \geq 41)$ si $X \sim N(\mu = 30; \sigma^2 = 25)$

SOLUCIÓN:

Si $\sigma^2 = 25$, entonces $\sigma = 5$, así



z	0	0.01	0.02
2.1	0.9821	0.9826	0.9830
2.2	0.9861	0.9864	0.9868
2.3	0.9893	0.9896	0.9898

$P(X \geq 41) = 1 - P(X < 41)$

$= 1 - P\left(\frac{X-30}{5} < \frac{41-30}{5}\right)$

$= 1 - P(Z < 2{,}2)$

Ahora, para calcular $1 - P(Z < 2{,}2)$, solo se utiliza la tabla Z, de forma que:

$1 - P(Z < 2{,}2) = 1 - \phi(2{,}2)$

$= 1 - 0{,}9861$

$= 0{,}0139$

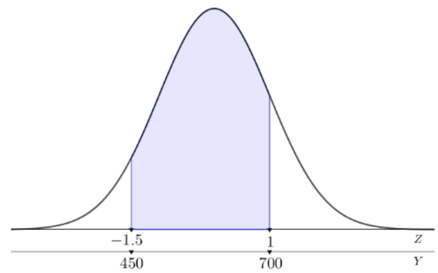
Por lo tanto, $P(X \geq 41) = 0{,}0139$

Nota al docente: Para este ejercicio, los estudiantes deben identificar que el cuantil está expresado solo en un decimal, por lo que la centésima en la tabla Z corresponde a 0.

d. $P(450 \leq Y \leq 700)$ si $Y \sim N(\mu = 600; \sigma^2 = 10\,000)$

SOLUCIÓN:

Si $\sigma^2 = 10\,000$, entonces $\sigma = 100$, así



$P(450 \leq Y \leq 700)$

$= P\left(\frac{450-600}{100} \leq \frac{Y-600}{100} \leq \frac{700-600}{100}\right)$

$= P(-1{,}5 \leq Z \leq 1)$

$= P(Z \leq 1) - P(Z < -1{,}5)$

z	0	0.01	0.02
-1.6	0.0548	0.0537	0.0526
-1.5	0.0668	0.0655	0.0643
-1.4	0.0808	0.0793	0.0778

Ahora, para calcular $P(Z \leq 1) - P(Z < -1{,}5)$ solo se utiliza la tabla Z, de forma que:

z	0	0.01	0.02
0.9	0.8159	0.8186	0.8212
1.0	0.8413	0.8438	0.8461
1.1	0.8643	0.8665	0.8686

$P(Z \leq 1) - P(Z < -1{,}5) = \phi(1) - \phi(-1{,}5)$

$= 0{,}8413 - 0{,}0668$

$= 0{,}7745$

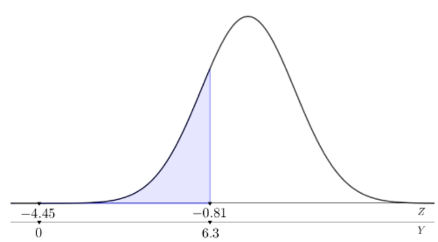
Por lo tanto, $P(450 \leq Y \leq 700) = 0{,}7745$

Nota al docente: En este caso, los estudiantes deben identificar que la probabilidad pedida está en un intervalo cerrado, por lo tanto, aplican la propiedad correspondiente, es decir, $P(a \leq X \leq b) = P(X \leq b) - P(X < a)$.

e. $P(0 < Y < 6{,}3)$ si $Y \sim N(\mu = 7{,}7; \sigma^2 = 3)$

SOLUCIÓN:

Si $\sigma^2 = 3$, entonces $\sigma = \sqrt{3}$, así



$P(0 < Y < 6{,}3)$

$= P\left(\frac{0-7{,}7}{\sqrt{3}} < \frac{Y-7{,}7}{\sqrt{3}} < \frac{6{,}3-7{,}7}{\sqrt{3}}\right)$

$= P(-4{,}45 < Z < -0{,}81)$

$= P(Z < -0{,}81) - P(Z \leq -4{,}45)$

z	0	0.01	0.02
-0.9	0.1841	0.1814	0.1788
-0.8	0.2119	0.2090	0.2061
-0.7	0.2420	0.2389	0.2358

Ahora, para calcular $P(Z < -0{,}81) - P(Z \leq -4{,}45)$, solo se utiliza la tabla Z, de forma que:

$P(Z < -0{,}81) - P(Z \leq -4{,}45) = \phi(-0{,}81) - \phi(-4{,}45)$

$= 0{,}209 - 0$

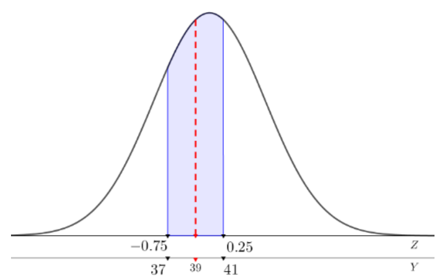
$= 0{,}209$

Por lo tanto, $P(0 < Y < 6{,}3) = 0{,}209$

Nota al docente: En este caso, los estudiantes deben identificar que la probabilidad pedida está en un intervalo abierto, por lo tanto, aplican la propiedad correspondiente, es decir, $P(a < X < b) = P(X < b) - P(X \leq a)$. Además, deben identificar que cualquier probabilidad menor a $-3{,}99$ siempre es 0, por lo que no es necesario utilizar la tabla Z.

f. $P(|X - 39| < 2)$ si $X \sim N(\mu = 40; \sigma^2 = 16)$

SOLUCIÓN:



z	0.04	0.05	0.06
-0.8	0.2005	0.1977	0.1949
-0.7	0.2296	0.2266	0.2236
-0.6	0.2611	0.2578	0.2546

z	0.04	0.05	0.06
0.1	0.5557	0.5596	0.5636
0.2	0.5948	0.5987	0.6026
0.3	0.6331	0.6368	0.6406

Si $\sigma^2 = 16$, entonces $\sigma = 4$, así

$P(|X - 39| < 2)$

$= P(-2 < X - 39 < 2)$

$= P(37 < X < 41)$

$= P\left(\frac{37-40}{4} < \frac{X-40}{4} < \frac{41-40}{4}\right)$

$= P(-0{,}75 < Z < 0{,}25)$

$= P(Z < 0{,}25) - P(Z \leq -0{,}75)$

$= P(Z < 0{,}25) - P(Z \leq -0{,}75)$

Ahora, para calcular $P(Z < 0{,}25) - P(Z < -0{,}25)$, solo se utiliza la tabla Z, de forma que:

$P(Z < 0{,}25) - P(Z \leq -0{,}25) = \phi(0{,}25) - \phi(-0{,}25)$

$= 0{,}5987 - 0{,}2266$

$= 0{,}3721$

Por lo tanto, $P(|X - 39| < 2) = 0{,}3721$

Nota al docente: Se sugiere recordar la propiedad de desigualdades con valor absoluto, es decir, $|b| < a \Longrightarrow -a < b < a$, además de la propiedad $P(a < X < b) = P(X < b) - P(X \leq a)$.

PASO 5. DESCUBRIENDO UNA NUEVA PROPIEDAD

DESCRIPCIÓN DE LA SUBACTIVIDAD:

- Esta subactividad permite inducir la propiedad simétrica de la Distribución Normal, a partir de la observación de los valores de la tabla Z y de las gráficas. Se justifica porque es una propiedad necesaria para el cálculo de probabilidades cuando se tiene solo una parte de la tabla.

DESARROLLO

Considera los valores de las siguientes tablas, y responde las preguntas.

$\boldsymbol{Z}$	$\mathbf{0}$	$\mathbf{0.01}$	$\mathbf{0.02}$
$\mathbf{-1.7}$	0.0446	0.0436	0.0427
$\mathbf{-1.6}$	0.0548	0.0537	0.0526
$\mathbf{-1.5}$	0.0668	0.0655	0.0643
$\mathbf{-1.4}$	0.0808	0.0793	0.0778

$\boldsymbol{Z}$	$\mathbf{0}$	$\mathbf{0.01}$	$\mathbf{0.02}$
$\mathbf{1.3}$	0.9032	0.9049	0.9066
$\mathbf{1.4}$	0.9192	0.9207	0.9222
$\mathbf{1.5}$	0.9332	0.9345	0.9357
$\mathbf{1.6}$	0.9452	0.9463	0.9474

a. Si $P(Z \leq -1{,}5)$ ¿Cuál es la probabilidad pedida?

SOLUCIÓN:

Mirando la tabla anterior, la probabilidad pedida es:

$$P(Z \leq -1{,}5) = 0{,}0668$$

b. Si $P(Z \leq 1{,}5)$ ¿Cuál es la probabilidad pedida?

SOLUCIÓN:

Mirando la tabla anterior, la probabilidad pedida es:

$$P(Z \leq 1{,}5) = 0{,}9332$$

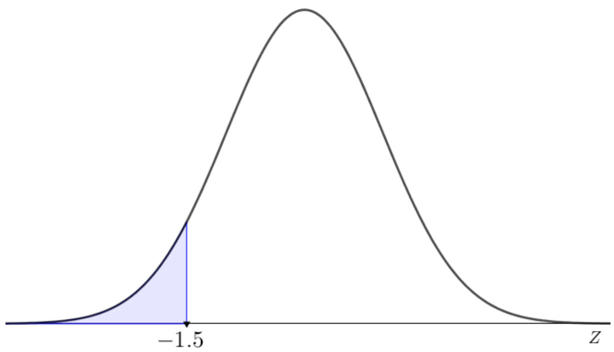
c. Dibuja las campanas asociadas a las siguientes probabilidades:

- $P(Z \leq -1{,}5)$

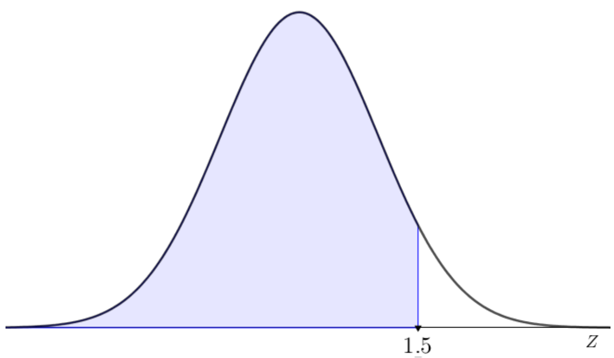
- $P(Z \leq 1{,}5)$
- $P(Z \geq 1{,}5)$

SOLUCIÓN:

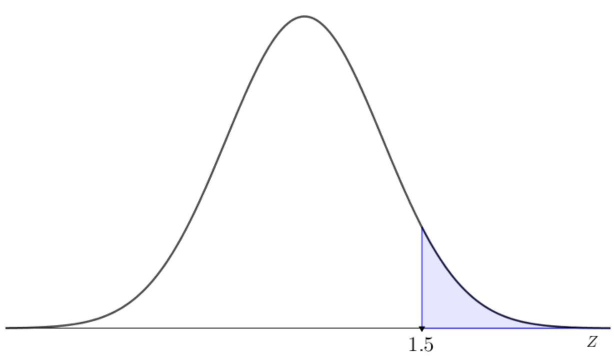
- Para $P(Z \leq -1{,}5)$:



- Para $P(Z \leq 1{,}5)$:



- Para $P(Z > 1{,}5)$:



d. A partir de las gráficas encuentra alguna relación entre ellas y conjetura una regla para determinar la probabilidad de $P(Z > 1{,}5)$ en función de la $P(Z \leq 1{,}5)$

SOLUCIÓN:

La relación que hay entre ellas es la siguiente:

$$P(Z > 1{,}5) = 1 - P(Z \leq 1{,}5) = 1 - 0{,}9332 = 0{,}0668$$

e. ¿Encuentras alguna relación entre el resultado anterior y $P(Z \leq -1{,}5)$?

SOLUCIÓN:

La relación que se debe encontrar es que los valores son iguales, entonces

$$P(Z > 1{,}5) = 1 - P(Z \leq 1{,}5) = 1 - 0{,}9332 = 0{,}0668 = P(Z \leq -1{,}5)$$

f. A partir de los resultados anteriores, conjetura una regla general para encontrar la probabilidad $P(Z \leq -z)$.

SOLUCIÓN:

La regla corresponde a:

$$P(Z \leq -z) = 1 - P(Z \leq z)$$

Lo que expresado en notación ϕ corresponde a:

$$\phi(-z) = 1 - \phi(z)$$

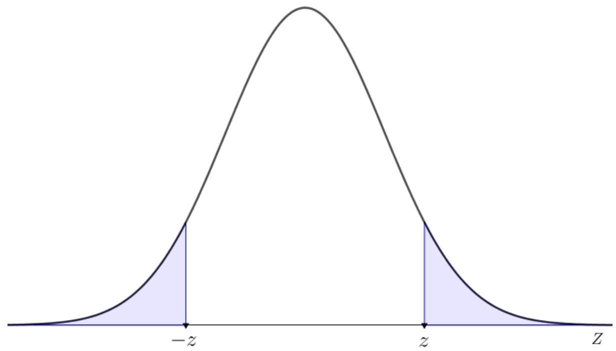
g. Expresa con tus palabras lo que significa la regla encontrada en la pregunta f. Además, grafícala en una campana.

SOLUCIÓN:

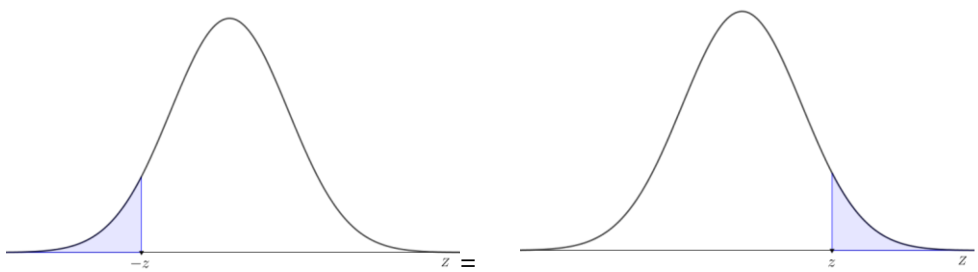
La propiedad expresa que la probabilidad que se encierra hasta el valor $-z$, es decir $P(Z \leq -z)$ es lo mismo que calcular la probabilidad del complemento de su opuesto aditivo, en otras palabras $1 - P(Z \leq z)$.

También puede expresarse que, la probabilidad que se encierra hasta el valor $-z$ es la misma que se encierra sobre el valor z, es decir, $P(Z > z)$.

En cualquiera de los dos casos, la información que representa la campana es la misma, y corresponde gráficamente a:



o



h. Utiliza las siguientes probabilidades para aplicarlas en la regla descubierta en la pregunta f. y verifica que el valor obtenido coincide con el de la tabla Z.

- $P(Z \leq 2{,}25) = 0{,}9878$
- $P(Z \leq 0{,}97) = 0{,}834$

<u>SOLUCIÓN:</u>

- $P(Z \leq 2{,}25) = 0{,}9878$

  Aplicando la conjetura, se tiene que:

  $P(Z \leq -2{,}25) = 1 - P(Z \leq 2{,}25)$

  $= 1 - 0{,}9878$

  $= 0{,}0122$

  Finalmente, se verifica que el valor obtenido con la propiedad coincide con la tabla Z.

- $P(Z \leq 0{,}97) = 0{,}834$

  Aplicando la conjetura, se tiene que:

  $P(Z \leq -0{,}97) = 1 - P(Z \leq 0{,}97)$

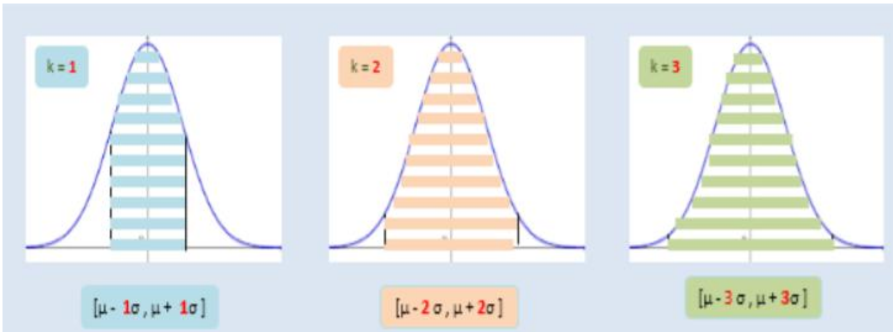
$= 1 - 0{,}834$

$= 0{,}166$

Finalmente, se verifica que el valor obtenido con la propiedad coincide con la tabla Z.

PASO 6. SIGNIFICADO DE LA DISTRIBUCIÓN NORMAL

La imagen siguiente muestra los gráficos de una distribución Normal de $N(\mu, \sigma^2)$ con los intervalos $[\mu - k \cdot \sigma, \mu + k \cdot \sigma]$ para $k = 1, 2, 3$.



a. La probabilidad de un intervalo simétrico al valor esperado μ de $P(|X - \mu| \leq c)$ se estandariza mediante la función ϕ en

$$\phi\left(\frac{c}{\sigma}\right) - \phi\left(-\frac{c}{\sigma}\right) = 2 \cdot \phi\left(\frac{c}{\sigma}\right) - 1$$

Con $c = k \cdot \sigma$ se obtiene $2 \cdot \phi(k) - 1$. Determinen las probabilidades que corresponden a los intervalos $[\mu - 1 \cdot \sigma, \mu + 1 \cdot \sigma]$, $[\mu - 2 \cdot \sigma, \mu + 2 \cdot \sigma]$ y $[\mu - 3 \cdot \sigma, \mu + 3 \cdot \sigma]$.

SOLUCIÓN:

Considerando la fórmula que se entrega, tenemos que:

Para $k = 1$:

$2 \cdot \phi(k) - 1$

$= 2 \cdot \phi(1) - 1$

$= 2 \cdot (0{,}8413) - 1$

$= 1{,}6826 - 1$

$= 0{,}6826$

Por lo tanto, la probabilidad que la variable aleatoria se aleje de la media en a lo más 1 desviación estándar es aproximadamente $68{,}26\%$.

Para $k = 2$,

Importante

Lo que la expresión

$$P(|X - \mu| \leq c)$$

señala es que, la distancia entre la variable y su media sea a lo más o como máximo c.

La deducción de esta fórmula se presenta a continuación:

$P(|X - \mu| \leq c)$

$= P(-c \leq X - \mu \leq c)$

$= P(\mu - c \leq X \leq \mu + c)$

$= P\left(\frac{(\mu - c) - \mu}{\sigma} \leq Z \leq \frac{(\mu + c) - \mu}{\sigma}\right)$

$= P\left(-\frac{c}{\sigma} \leq Z \leq \frac{c}{\sigma}\right)$

$= P\left(Z \leq \frac{c}{\sigma}\right) - P\left(Z \leq -\frac{c}{\sigma}\right)$

$= P\left(Z \leq \frac{c}{\sigma}\right) - \left[1 - P\left(Z \leq \frac{c}{\sigma}\right)\right]$

$= P\left(Z \leq \frac{c}{\sigma}\right) - 1 + P\left(Z \leq \frac{c}{\sigma}\right)$

$= 2 \cdot P\left(Z \leq \frac{c}{\sigma}\right) - 1$

$= 2 \cdot \phi\left(\frac{c}{\sigma}\right) - 1$

$2 \cdot \phi(k) - 1$

$= 2 \cdot \phi(2) - 1$

$= 2 \cdot (0{,}9772) - 1$

$= 1{,}9544 - 1$

$= 0{,}9544$

Por lo tanto, la probabilidad que la variable aleatoria se aleje de la media en a lo más 2 desviaciones estándar es aproximadamente 95,44%.

Para $k = 3$,

$2 \cdot \phi(k) - 1$

$= 2 \cdot \phi(3) - 1$

$= 2 \cdot (0{,}9987) - 1$

$= 1{,}9974 - 1$

$= 0{,}9974$

Por lo tanto, la probabilidad que la variable aleatoria se aleje de la media en a lo más 3 desviaciones estándar es aproximadamente 99,74%.

Nota al docente: Los valores obtenidos también pueden redondearse, de forma que $P(\mu - \sigma \leq X \leq \mu + \sigma) = 0{,}683$; $P(\mu - 2 \cdot \sigma \leq X \leq \mu + 2 \cdot \sigma) = 0{,}954$; $P(\mu - 3 \cdot \sigma \leq X \leq \mu + 3 \cdot \sigma) = 0{,}997$.

b. ¿Por qué estas probabilidades son válidas para todas las distribuciones Normales estandarizadas? Argumenten su respuesta.

SOLUCIÓN:

Porque la forma de estandarizar una variable aleatoria Normal cualquiera es solo una y corresponde al teorema de Estandarización $Z = \frac{X-\mu}{\sigma}$. Como es posible apreciar, esta expresión (para cualquier variable Normal) solo requiere conocer la media y la desviación estándar, sin importar sus valores. Por ello las áreas serán equivalentes para cualquier Normal. Esto es posible visualizarlo en la expresión

$$\phi\left(\frac{c}{\sigma}\right) - \phi\left(-\frac{c}{\sigma}\right) = 2 \cdot \phi\left(\frac{c}{\sigma}\right) - 1$$

Cuando se considera $c = k \cdot \sigma, k \in \mathbb{R}$, ya que siempre se terminará calculando $\phi(k)$, que corresponde a un valor constante que se debe ubicar la tabla Z, pues la desviación estándar se simplifica.

c. ¿Qué ventaja tienen las distribuciones Normales Estandarizadas? Argumenten su respuesta.

SOLUCIÓN

Considerando que la Distribución Normal Estandarizada es solo una, que tiene como parámetros $\mu = 0$ y $\sigma = 1$, la ventaja es que facilita el cálculo de probabilidades, ya que solo se requiere de la tabla Z para determinar probabilidades.

OBJETIVO DE LAS PREGUNTAS COMPLEMENTARIAS

- Resolver problemas que involucren el uso de los intervalos icónicos y estandarización para el cálculo de probabilidades.

DESARROLLO

d. Una máquina que rellena botellas de bebidas sufrió un desperfecto técnico con la variabilidad del líquido respecto al volumen medio a vaciar en el envase, por ello la empresa ha decidido realizar un estudio para diagnosticar el nuevo valor de la desviación estándar. Los técnicos solo conocen que el volumen de líquido depositado sigue una distribución Normal de parámetros $\mu = 591\ ml$ con una desviación estándar desconocida.

i. Si se sabe que aproximadamente 3415 botellas de 5000 tienen un volumen entre $588{,}5\ ml$ y $593{,}5\ ml$ ¿cuál es la desviación estándar σ del volumen del líquido?

SOLUCIÓN:

Sea:

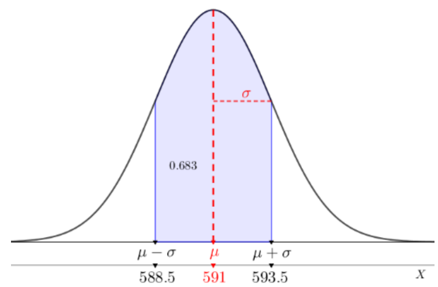
$$X: Volumen\ de\ la\ bebida, en\ ml.$$

Se sabe que 3415 botellas de las 5000 tienen un volumen de líquido entre $588{,}5$ y $593{,}5\ ml$, por lo que, se determinará el porcentaje correspondiente a las 3415 botellas. Así:

$$\begin{array}{ccc} 5000 & \longrightarrow & 100\% \\ 3415 & \longrightarrow & x \end{array}$$

De modo que $x = \frac{3415 \cdot 100}{5000} = 68{,}3$, por lo tanto, 68,3% de las botellas tienen un volumen de líquido entre 588,5 y 593,5 ml. Ahora, hay que recordar que la probabilidad que la variable aleatoria se aleje de su media en a lo más 1 desviación estándar, es 68,3% aproximadamente, en otras palabras

$$P(\mu - \sigma \leq X \leq \mu + \sigma) = 0{,}683$$



Como el volumen del líquido está entre 588,5 y 593,5 ml, se puede anotar

$$P(588{,}5 \leq X \leq 593{,}5) = 0{,}683$$

Así, es posible ver que $\mu - \sigma = 588{,}5$ (también puede ser $\mu + \sigma = 593{,}5$). Como $\mu = 591\ ml$, se tiene que

$$591 - \sigma = 588{,}5 \Longrightarrow \sigma = 591 - 588{,}5 \Longrightarrow \sigma = 2{,}5$$

Por lo tanto, $\sigma = 2{,}5\ ml$, así $X \sim N(\mu = 591; \sigma^2 = 2{,}5^2)$.

ii. ¿Cuál es la probabilidad que el contenido de la botella supere los $600\ ml$ de líquido?

SOLUCIÓN:

Usando la variable definida en 1, se pide $P(X > 600)$. Así:

$$P(X > 600) = 1 - P(X \leq 600)$$

$$= 1 - P\left(\frac{X-591}{2{,}5} \leq \frac{600-591}{2{,}5}\right)$$

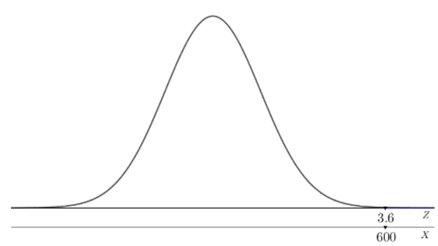
$$= 1 - P\left(Z \leq \frac{18}{5}\right)$$

$$= 1 - P(Z \leq 3{,}6)$$

$$= 1 - \phi(3{,}6)$$

$$= 1 - 0{,}9998$$

$$= 0{,}0002$$



z	0	0.01	0.02
3.5	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999
3.7	0.9999	0.9999	0.9999

En conclusión, es casi imposible que el volumen de líquido de la botella supere los $600\ ml$, ya que la probabilidad tiende a cero.

e. La longitud diametral de las *donuts* que consume Homero Simpson siguen una distribución Normal de parámetros μ desconocido y varianza $\sigma^2 = 2\ cm^2$. Suponiendo que anualmente consume 1000 *donuts* de las cuales aproximadamente 683 tienen un diámetro entre $6{,}79$ y $9{,}61\ cm$. ¿Cuál es el valor de la media de la variable en cuestión?

SOLUCIÓN:

Sea:

$$W: Longitud\ diametral\ de\ las\ donuts\ que\ consume\ Homero\ Simpson, en\ cm.$$

Se sabe que 683 de las 1000 donuts tienen una longitud diametral entre $6{,}79$ y $9{,}61\ cm$ por lo que, el porcentaje de las 683 donuts respecto del total es:

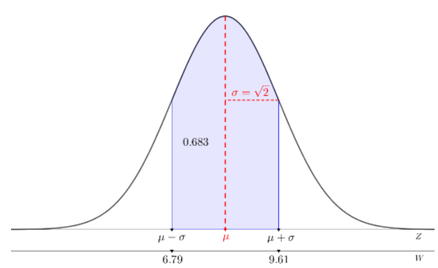
$$\begin{array}{ccc} 1000 & \longrightarrow & 100\% \\ 683 & \longrightarrow & w \end{array}$$

De manera que $w = \frac{683 \cdot 100}{1000} = 68{,}3$. Luego, $68{,}3\%$ de las donuts tienen una longitud diametral entre $6{,}79$ y $9{,}61\ cm$. Ahora, hay que recordar que la probabilidad que la variable aleatoria se aleje de su media en a lo más 1 desviación estándar es $68{,}3\%$, lo que se puede escribir como:

$$P(\mu - \sigma \leq W \leq \mu + \sigma) = 0{,}683$$

Puesto en el contexto del problema, la expresión anterior queda de la siguiente forma:

$$P(6{,}79 \leq W \leq 9{,}61) = 0{,}683$$



Así, se puede observar que $\mu + \sigma = 9{,}61$, con $\sigma = \sqrt{2}\ cm$.

Despejando se tiene:

$$\mu + \sqrt{2} = 9{,}61 \Rightarrow \mu = 9{,}61 - \sqrt{2} \Rightarrow \mu = 8{,}2$$

Por lo tanto, el valor de longitud medio del diámetro de las *donuts* que come Homero Simpson es de $8{,}2\ cm$. Así $W \sim N(\mu = 8{,}2; \sigma^2 = 2)$.

PASO 7. DESESTANDARIZANDO UNA VARIABLE ALEATORIA NORMAL

DESCRIPCIÓN DE LA SUBACTIVIDAD:

- Resolver ejercicios que involucren el cálculo de probabilidades y cuantiles de la Distribución Normal a través del uso de la tabla Z en forma inversa.
- La primera parte de la subactividad permite que los estudiantes comprendan que la tabla Z también puede utilizarse en forma inversa, es decir, dada una probabilidad, buscar el valor del cuantil que acumula dicha probabilidad.
- La segunda parte de la subactividad permite que los estudiantes comprendan y apliquen la desestandarización.

DESARROLLO

A. Calcula los siguientes cuantiles de una variable aleatoria $N(0{,}1)$, esbozando la gráfica donde se identifique la probabilidad dada.

Ejemplo

Para calcular el valor z, tal que hasta ese valor se acumule un $80{,}5\%$ de probabilidad, en otras palabras $P(Z < z) = 0{,}805$, solo se utiliza la tabla Z, observando en su parte interior. Una vez ubicada la probabilidad al interior de la tabla (en este caso $0{,}805$), se busca el cuantil que acumula dicha probabilidad. Este valor corresponde a 0,86. El proceso inverso a la función ϕ, se denota ϕ^{-1}.

Así,

$P(Z < z) = 0{,}805$

$\Rightarrow \phi(z) = 0{,}805$

$\Rightarrow \phi^{-1}(\phi(z)) = \phi^{-1}(0{,}805)$

$\Rightarrow z = \phi^{-1}(0{,}805)$

$\Rightarrow z = 0{,}86$

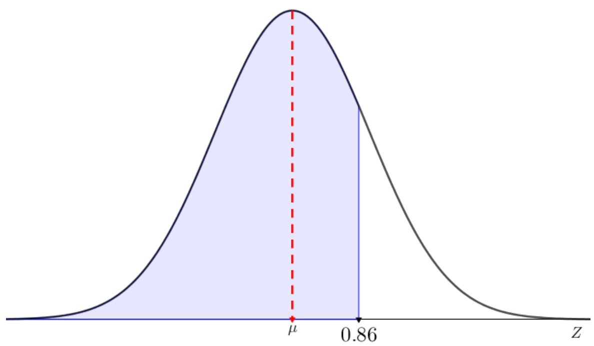
Finalmente, $P(Z < 0{,}86) = 0{,}805$

Décimas

Centésimas

z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389

De forma gráfica, se tiene que



a. $P(Z \leq z) = 0{,}9$

SOLUCIÓN:

$P(Z \leq z) = 0{,}9$

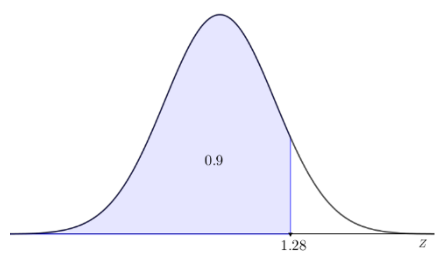
$\Longrightarrow \phi(z) = 0{,}9$

$\Longrightarrow \phi^{-1}\left(\phi(z)\right) = \phi^{-1}(0{,}9)$

$\Longrightarrow z = \phi^{-1}(0{,}9)$

$\Longrightarrow z = 1{,}28$

Luego, $P(Z \leq 1{,}28) = 0{,}9$.



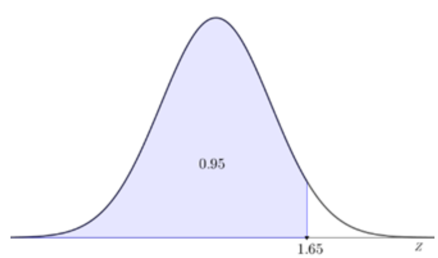
z	0.07	0.08	0.09
1.1	0.8790	0.8810	0.8830
1.2	0.8980	0.8997	0.9015
1.3	0.9147	0.9162	0.9177

b. $P(Z < z) = 0{,}95$

SOLUCIÓN:

$P(Z \leq z) = 0{,}95$

$\Longrightarrow \phi(z) = 0{,}95$



$\Rightarrow z = \phi^{-1}(0{,}95)$

$\Rightarrow z = 1{,}65$

z	0.04	0.05	0.06
1.5	0.9382	0.9394	0.9406
1.6	0.9495	0.9505	0.9515
1.7	0.9591	0.9599	0.9608

En consecuencia, $P(Z \leq 1{,}65) = 0{,}95$.

Nota al docente: Para resolver el ejercicio a. y b. solo se utiliza la tabla en forma inversa, es decir, mirándola desde adentro hacia afuera. Para b., también se puede considerar que $z = 1{,}64$, pues la diferencia absoluta es la misma hacia ambos valores, de forma que aproximando resulta ser $0{,}9495 \approx 0{,}95$.

c. $P(Z \geq z) = 0{,}33$

SOLUCIÓN:

$P(Z \geq z) = 0{,}33$

$\Rightarrow 1 - P(Z < z) = 0{,}33$

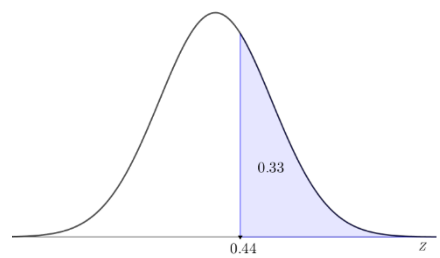
$\Rightarrow P(Z < z) = 1 - 0{,}33$

$\Rightarrow \phi(z) = 0{,}67$

$\Rightarrow \phi^{-1}(\phi(z)) = \phi^{-1}(0{,}67)$

$\Rightarrow z = \phi^{-1}(0{,}67)$

$\Rightarrow z = 0{,}44$



z	0.03	0.04	0.05
0.3	0.6293	0.6331	0.6368
0.4	0.6664	0.6700	0.6736
0.5	0.7019	0.7054	0.7088

Por lo tanto, $P(Z \geq 0{,}44) = 0{,}33$

d. $P(Z > z) = 0{,}9$

SOLUCIÓN:

$P(Z > z) = 0{,}9$

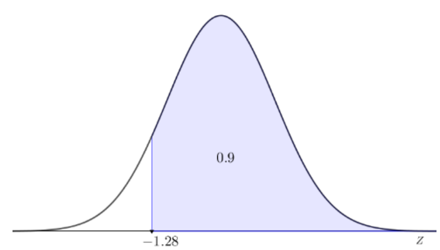
$\Rightarrow 1 - P(Z \leq z) = 0{,}9$

$\Rightarrow P(Z \leq z) = 1 - 0{,}9$

$\Rightarrow \phi(z) = 0{,}1$

$\Rightarrow z = \phi^{-1}(0{,}1)$

$\Rightarrow z = -1{,}28$



z	0.07	0.08	0.09
-1.3	0.0853	0.0838	0.0823
-1.2	0.1020	0.1003	0.0985
-1.1	0.1210	0.1190	0.1170

Luego, $(Z > -1{,}28) = 0{,}9$.

Nota al docente: Para los ejercicios c. y d., los estudiantes antes de utilizar la tabla en forma inversa deben aplicar la propiedad del complemento $P(X \geq a) = 1 - P(X < a)$ y $P(X > a) = 1 - P(Z \leq a)$ respectivamente.

e. $P(z < Z < 1{,}7) = 0{,}9108$

SOLUCIÓN:

$P(z < Z < 1{,}7) = 0{,}9108$

$\Longrightarrow P(Z \leq 1{,}7) - P(Z \leq z) = 0{,}9108$

$\Longrightarrow \phi(1{,}7) - \phi(z) = 0{,}9108$

$\Longrightarrow 0{,}9554 - \phi(z) = 0{,}9108$

$\Longrightarrow \phi(z) = 0{,}9554 - 0{,}9108$

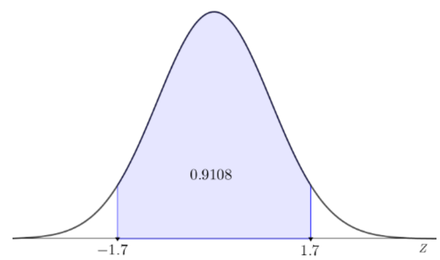
$\Longrightarrow \phi(z) = 0{,}0446$

$\Longrightarrow \phi^{-1}(\phi(z)) = \phi^{-1}(0{,}0446)$

$\Longrightarrow z = \phi^{-1}(0{,}0446)$

$\Longrightarrow z = -1{,}7$

Por lo que, $P(-1{,}7 < Z < 1{,}7) = 0{,}9108$.



z	0	0.01	0.02
1.6	0.9452	0.9463	0.9474
1.7	0.9554	0.9564	0.9573
1.8	0.9641	0.9649	0.9656

z	0	0.01	0.02
-1.8	0.0359	0.0351	0.0344
-1.7	0.0446	0.0436	0.0427
-1.6	0.0548	0.0537	0.0526

f. $P(-0{,}46 < Z \leq z) = 0{,}6142$

SOLUCIÓN:

$P(-0{,}46 < Z \leq z) = 0{,}6142$

$\Longrightarrow P(Z \leq z) - P(Z \leq -0{,}46) = 0{,}6142$

$\Longrightarrow \phi(z) - \phi(-0{,}46) = 0{,}6142$

$\Longrightarrow \phi(z) - 0{,}3228 = 0{,}6142$

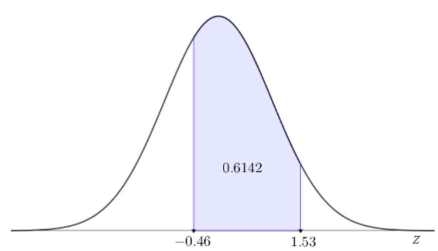
$\Longrightarrow \phi(z) = 0{,}937$

$\Longrightarrow \phi^{-1}(\phi(z)) = \phi^{-1}(0{,}937)$

$\Longrightarrow z = \phi^{-1}(0{,}937)$

$\Longrightarrow z = 1{,}53$

Luego, $P(-0{,}46 < Z \leq 1{,}53) = 0{,}6142$.



z	0.05	0.06	0.07
-0.5	0.2912	0.2877	0.2843
-0.4	0.3264	0.3228	0.3192
-0.3	0.3632	0.3594	0.3557

z	0.02	0.03	0.04
1.4	0.9222	0.9236	0.9251
1.5	0.9357	0.9370	0.9382
1.6	0.9474	0.9484	0.9495

Nota al docente: Los estudiantes deben identificar que para resolver los ejercicios e. y f. se utiliza tanto la tabla Z para hallar la probabilidad como para encontrar el *cuantil. Además, se utilizan las propiedades* $P(a < X < b) = P(X \leq b) - P(X \leq a)$ *y* $P(a < X \leq b) = P(X \leq b) - P(X \leq a)$ *respectivamente.*

B. Encuentra el valor de la variable original que encierra la probabilidad dada.

Ejemplo

$P(X \leq x) = 0{,}99$, si $X \sim N(\mu = 15{,}5; \sigma^2 = 4)$

Para encontrar el puntaje x que encierra dicha probabilidad, es necesario expresar la probabilidad pedida en notación ϕ, e identificar los parámetros de la Distribución Normal, para poder expresar a Z como $\frac{X-\mu}{\sigma}$. De forma que:

$P(X \leq x) = 0{,}99$

$\Rightarrow P\left(Z \leq \frac{x-15{,}5}{2}\right) = 0{,}99$

$\Rightarrow \phi\left(\frac{x-15{,}5}{2}\right) = 0{,}99$

$\Rightarrow \phi^{-1}\left(\phi\left(\frac{x-15{,}5}{2}\right)\right) = \phi^{-1}(0{,}99)$

$\Rightarrow \frac{x-15{,}5}{2} = \phi^{-1}(0{,}99)$

$\Rightarrow \frac{x-15{,}5}{2} = 2{,}33$

$\Rightarrow x = 15{,}5 + 2 \cdot 2{,}33$

$\Rightarrow x = 20{,}16$

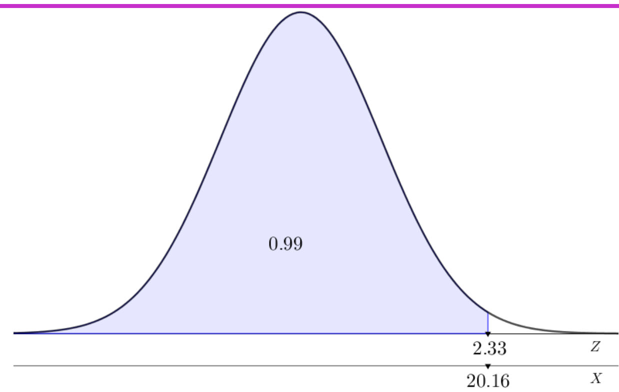
Finalmente, $x = 20{,}16$

Décimas

Centésimas

z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964

De forma gráfica, se tiene que



a. $P(X \leq x) = 0{,}9772$, si $X \sim N(\mu = 30; \sigma^2 = 25)$

SOLUCIÓN:

$P(X \leq x) = 0{,}9772$

$\Longrightarrow P\left(Z \leq \frac{x-30}{5}\right) = 0{,}9772$

$\Longrightarrow \phi\left(\frac{x-30}{5}\right) = 0{,}9772$

$\Longrightarrow \phi^{-1}\left(\phi\left(\frac{x-30}{5}\right)\right) = \phi^{-1}(0{,}9772)$

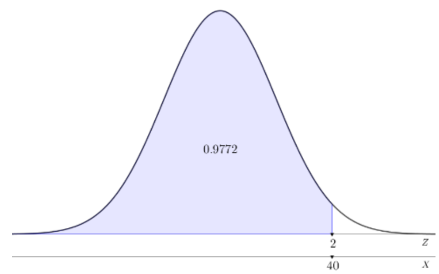
$\Longrightarrow \frac{x-30}{5} = \phi^{-1}(0{,}9772)$

$\Longrightarrow \frac{x-30}{5} = \phi^{-1}(0{,}9772)$

$\Longrightarrow \frac{x-30}{5} = 2$

$\Longrightarrow x = 30 + 5 \cdot 2$

$\Longrightarrow x = 40$



z	0	0.01	0.02
1.9	0.9713	0.9719	0.9726
2.0	0.9772	0.9778	0.9783
2.1	0.9821	0.9826	0.9830

Por lo tanto, $P(X \leq 40) = 0{,}9772$, si $X \sim N(\mu = 30; \sigma^2 = 25)$.

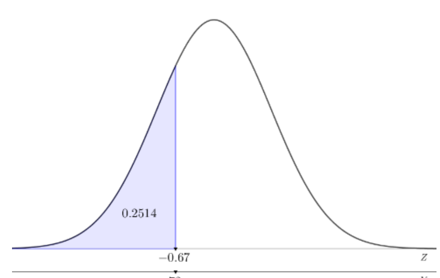
b. $P(Y < y) = 0{,}2514$, si $Y \sim N(\mu = 87; \sigma^2 = 500)$

SOLUCIÓN:

$P(Y < y) = 0{,}2514$

$\Longrightarrow P\left(Z < \frac{y-87}{\sqrt{500}}\right) = 0{,}2514$

$\Longrightarrow \phi\left(\frac{y-87}{\sqrt{500}}\right) = 0{,}2514$



$\Longrightarrow \phi^{-1}\left(\phi\left(\frac{y-87}{\sqrt{500}}\right)\right) = \phi^{-1}(0{,}2514)$

z	0.06	0.07	0.08
-0.7	0.2236	0.2206	0.2177
-0.6	0.2546	0.2514	0.2483
-0.5	0.2877	0.2843	0.2810

$\Longrightarrow \frac{y-87}{\sqrt{500}} = \phi^{-1}(0{,}2514)$

$\Longrightarrow \frac{y-87}{\sqrt{500}} = -0{,}67$

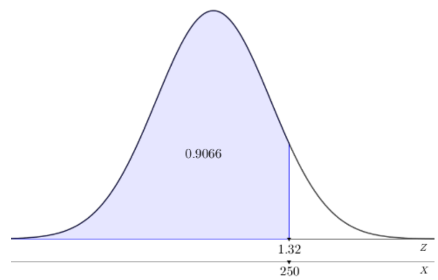
$\Longrightarrow y = 87 + \sqrt{500} \cdot (-0{,}67)$

$\Longrightarrow y = 72$

Luego, $P(Y < 72) = 0{,}2514$, si $Y \sim N(\mu = 87; \sigma^2 = 500)$.

c. $P(X \leq x) = 0{,}9066$, si $X \sim N(\mu = 217; \sigma^2 = 625)$

SOLUCIÓN:



$P(X \leq x) = 0{,}9066$

$P\left(Z \leq \frac{x-217}{25}\right) = 0{,}9066$

$\Longrightarrow \phi\left(\frac{x-217}{25}\right) = 0{,}9066$

$\Longrightarrow \phi^{-1}\left(\phi\left(\frac{x-217}{25}\right)\right) = \phi^{-1}(0{,}9066)$

z	0.01	0.02	0.03
1.2	0.8869	0.8888	0.8907
1.3	0.9049	0.9066	0.9082
1.4	0.9207	0.9222	0.9236

$\Longrightarrow \frac{x-217}{25} = \phi^{-1}(0{,}9066)$

$\Longrightarrow \frac{x-217}{25} = 1{,}32$

$\Longrightarrow x = 217 + 25 \cdot 1{,}32$

$\Longrightarrow x = 250$

Por lo tanto, $P(X \leq 250) = 0{,}9066$, si $X \sim N(\mu = 217; \sigma^2 = 625)$.

d. $P(Y \leq y) = 0{,}1379$, si $Y \sim N(\mu = 90; \sigma^2 = 68)$

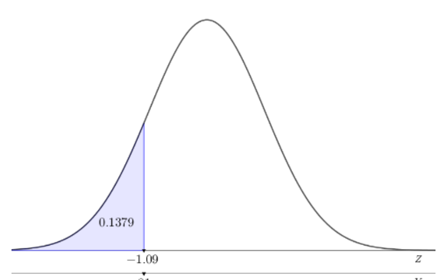
SOLUCIÓN:

$P(Y \leq y) = 0{,}1379$

$\Longrightarrow P\left(Z \leq \frac{y-90}{\sqrt{68}}\right) = 0{,}1379$

$\Longrightarrow \phi\left(\frac{y-90}{\sqrt{68}}\right) = 0{,}1379$

$\Longrightarrow \phi^{-1}\left(\phi\left(\frac{y-90}{\sqrt{68}}\right)\right) = \phi^{-1}(0{,}1379)$



$\Longrightarrow \frac{y-90}{\sqrt{68}} = \phi^{-1}(0{,}1379)$

$\Longrightarrow \frac{y-90}{\sqrt{68}} = -1{,}09$

z	0.07	0.08	0.09
-1.1	0.1210	0.1190	0.1170
-1.0	0.1423	0.1401	0.1379
-0.9	0.1660	0.1635	0.1611

$\Longrightarrow y = 90 + \sqrt{68} \cdot (-1{,}09)$

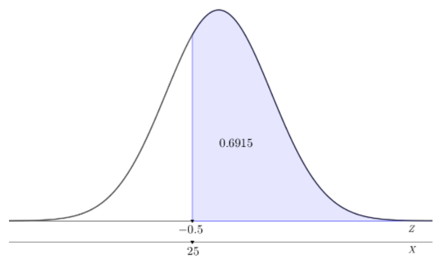
$\Longrightarrow y = 81$

En consecuencia, $P(Y \leq 81) = 0{,}1379$, si $Y \sim N(\mu = 90; \sigma^2 = 68)$

e. $P(X \geq x) = 0{,}6915$, si $X \sim N(\mu = 26; \sigma^2 = 4)$

SOLUCIÓN:

$P(X \geq x) = 0{,}6915$



$\Longrightarrow 1 - P(X < x) = 0{,}6915$

$\Longrightarrow P(X < x) = 1 - 0{,}6915$

$\Longrightarrow P(X < x) = 0{,}3085$

$\Longrightarrow P\left(Z < \frac{x-26}{2}\right) = 0{,}3085$

z	0	0.01	0.02
-0.6	0.2743	0.2709	0.2676
-0.5	0.3085	0.3050	0.3015
-0.4	0.3446	0.3409	0.3372

$\Longrightarrow \phi\left(\frac{x-26}{2}\right) = 0{,}3085$

$\Longrightarrow \phi^{-1}\left(\phi\left(\frac{x-26}{2}\right)\right) = \phi^{-1}(0{,}3085)$

$\Longrightarrow \frac{x-26}{2} = \phi^{-1}(0{,}3085)$

$\Longrightarrow \frac{x-26}{2} = -0{,}5$

$\Longrightarrow x = 26 + 2 \cdot (-0{,}5)$

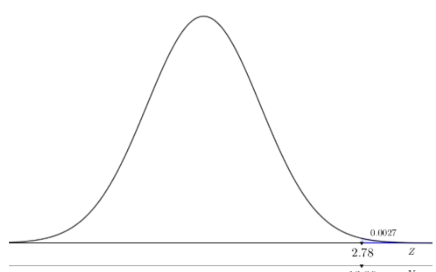
$\Longrightarrow x = 25$

Luego $P(X \geq 25) = 0{,}6915$, si $X \sim N(\mu = 26; \sigma^2 = 4)$

f. $P(Y > y) = 0{,}0027$, si $Y \sim N(\mu = 6; \sigma^2 = 8)$

SOLUCIÓN:

$P(Y > y) = 0{,}0027$



$\Longrightarrow 1 - P(Y \leq y) = 0{,}0027$

$\Longrightarrow P(Y \leq y) = 1 - 0{,}0027$

$\Longrightarrow P(Y \leq y) = 1 - 0{,}9973$

$\Longrightarrow P\left(Z \leq \frac{y-6}{\sqrt{8}}\right) = 0{,}9973$

z	0.07	0.08	0.09
2.6	0.9962	0.9963	0.9964
2.7	0.9972	0.9973	0.9974
2.8	0.9979	0.9980	0.9981

$\Longrightarrow \phi\left(\frac{y-6}{\sqrt{8}}\right) = 0{,}9973$

$\Longrightarrow \phi^{-1}\left(\phi\left(\frac{y-6}{\sqrt{8}}\right)\right) = \phi^{-1}(0{,}9973)$

$\Longrightarrow \frac{y-6}{\sqrt{8}} = \phi^{-1}(0{,}9973)$

$\Longrightarrow \frac{y-6}{\sqrt{8}} = 2{,}78$

$\Longrightarrow y = 6 + \sqrt{8} \cdot 2{,}78$

$\Longrightarrow y = 13{,}86$

Por lo que, $P(Y > y) = 0{,}0027$, si $Y \sim N(\mu = 6; \sigma^2 = 8)$

PASO 8. APROXIMACIÓN DE UNA DISTRIBUCIÓN BINOMIAL POR LA DISTRIBUCIÓN NORMAL ESTANDARIZADA

Conexión Interdisciplinaria:
Educación Ciudadana
OA a, 3° y 4° medio.

Según el registro electoral de una región, la participación en las últimas elecciones era de aproximadamente 80%. Para una encuesta telefónica, se eligió al azar una muestra de 400 personas.

a. Determinen los parámetros de μ y σ de la distribución Binomial y transfórmenla en la distribución Normal Estandarizada correspondiente.

<u>SOLUCIÓN</u>

La pregunta se resuelve con la aproximación de la Binomial a la Normal, también conocido como el Teorema de De Moivre-Laplace. De esta forma:

Sea:

$$X\text{: }número\ de\ personas\ que\ no\ participaron\ en\ las\ últimas\ elecciones,$$

$$de\ un\ total\ de\ 400\ personas.$$

Se tiene que $X \sim Bin(n = 400, p = 0{,}2)$.

Luego nuestra función de probabilidad en el contexto tiene la forma

$$P(X = x) = \begin{cases} \binom{400}{x} \cdot (0{,}2)^x \cdot (0{,}8)^{400-x} & si\ x = 0{,}1{,}2, \dots ,400 \\ 0 & en\ otro\ caso \end{cases}$$

Para la esperanza y la varianza de X, se tiene

$$\mu = n \cdot p = 400 \cdot 0{,}2 = 80\ personas$$

$$\sigma = \sqrt{n \cdot p \cdot q} = \sqrt{400 \cdot 0{,}2 \cdot 0{,}8} = \sqrt{64} = 8\ personas$$

Ahora, para verificar las condiciones de la aproximación se tiene que:

$$n = 400 > 10$$

$$n \cdot p = 400 \cdot 0{,}2 = 80 \geq 5$$

$$n \cdot q = 400 \cdot 0{,}8 = 320 \geq 5$$

Como se cumplen las condiciones, se tiene que:

$$\mu = n \cdot p = 400 \cdot 0{,}2 = 80$$

$$\sigma = \sqrt{n \cdot p \cdot q} = \sqrt{400 \cdot 0{,}2 \cdot 0{,}8} = \sqrt{64} = 8$$

Luego, la fórmula para estandarizar en el contexto del ejercicio es:

$$Z = \frac{X-\mu}{\sigma} \Rightarrow Z = \frac{X-80}{8}$$

Nota al docente: La forma en que se definió la variable X es netamente por conveniencia para los ejercicios que siguen, por lo que se sugiere tener cuidado al escoger el valor de p, que corresponde a $0{,}2$ y no $0{,}8$. Hay que recordar que el valor de p es la probabilidad de "éxito" de la variable aleatoria X, y no siempre se asocia a eventos positivos.

b. ¿Cuál es la probabilidad de tener en la muestra más de 88 personas que suelen no participar en las elecciones? Argumenten su respuesta.

SOLUCIÓN:

Nota al docente: Esta pregunta se resuelve de dos formas, con y sin la corrección de continuidad. La diferencia se tiene cuando no se calculan probabilidades en un valor puntual del recorrido de la variable aleatoria, sino de la forma mayor o igual, esta situación permite realizar el cálculo directo sin necesidad de hacer la corrección. Si bien ambas probabilidades no difieren en un gran valor, se recomienda realizar la corrección, ya que se aproxima más a la probabilidad calculada con la Binomial.

Se sugiere realizar el cálculo de la misma probabilidad con la Distribución Binomial en GeoGebra y hacer la comparación, para ver la importancia de la Corrección de Continuidad.

- Forma 1:

Considerando que la variable aleatoria es la definida en a., en esta pregunta se pide $P(X \geq 88)$. Así, se resolverá sin utilizar la aproximación:

$P(X \geq 88) = 1 - P(X < 88)$

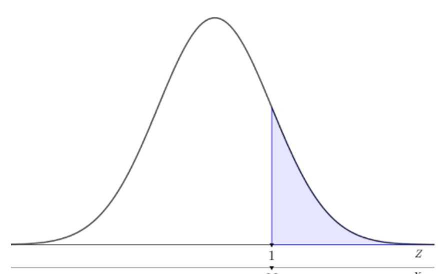
$= 1 - P\left(Z < \frac{88-80}{8}\right)$

$= 1 - P(Z < 1)$

$= 1 - \phi(1)$

$= 1 - 0{,}8413$

$= 0{,}1587$



z	0	0.01	0.02
0.9	0.8159	0.8186	0.8212
1.0	0.8413	0.8438	0.8461
1.1	0.8643	0.8665	0.8686

Por consiguiente, la probabilidad de tener en la muestra más de 88 personas que suelen no participar en las elecciones es de un 15,87%.

- Forma 2:

Considerando que la variable aleatoria es la definida en a., en esta pregunta se pide $P(X \geq 88)$. Así, se resolverá utilizando la aproximación:

$P(X \geq 88) \approx P(X \geq 88 - 0{,}5)$

$= 1 - P(X < 87{,}5)$

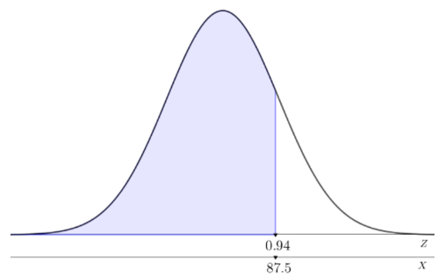
$\approx 1 - P\left(Z < \frac{87{,}5-80}{8}\right)$

$= 1 - P(Z < 0{,}94)$

$= 1 - \phi(0{,}94)$

$= 1 - 0{,}8264$

$= 0{,}1736$



z	0.03	0.04	0.05
0.8	0.7967	0.7995	0.8023
0.9	0.8238	0.8264	0.8289
1.0	0.8485	0.8508	0.8531

Por consiguiente, la probabilidad de tener en la muestra más de 88 personas que suelen no participar en las elecciones es de un 17,36%.

c. ¿Cuál es la probabilidad de tener en la muestra, como máximo 70 personas que suelen no participar en las elecciones? Argumenten su respuesta.

SOLUCIÓN:

- Forma 1:

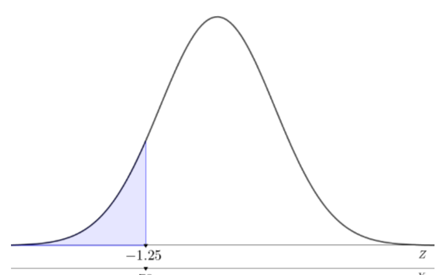
Considerando que la variable aleatoria es la definida en a., en esta pregunta se pide $P(X \leq 70)$. Así, se resolverá sin utilizar la aproximación:

$P(X \leq 70) = P\left(Z \leq \frac{70-80}{8}\right)$

$= P(Z \leq -1{,}25)$

$= \phi(-1{,}25)$

$= 0{,}1056$



z	0.04	0.05	0.06
-1.3	0.0901	0.0885	0.0869
-1.2	0.1075	0.1056	0.1038
-1.1	0.1271	0.1251	0.1230

Por lo tanto, la probabilidad de tener en la muestra como máximo 70 personas que suelen no participar en las elecciones es de un 10,56%.

- Forma 2:

  Considerando que la variable aleatoria es la definida en a., en esta pregunta se pide $P(X \leq 70)$. Así, se resolverá utilizando la aproximación:

  $P(X \leq 70) \approx P(X \leq 70 + 0{,}5)$

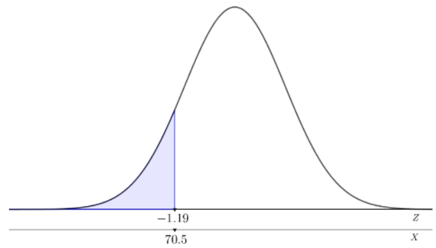
  $= P(X < 70{,}5)$

  $\approx P\left(Z < \frac{70{,}5-80}{8}\right)$

  $= P(Z < -1{,}19)$

  $= \phi(-1{,}19)$

  $= 0{,}117$



z	0.07	0.08	0.09
-1.2	0.1020	0.1003	0.0985
-1.1	0.1210	0.1190	0.1170
-1.0	0.1423	0.1401	0.1379

  Por consiguiente, la probabilidad de tener en la muestra más de 88 personas que suelen no participar en las elecciones es de un 11,7%.

d. ¿Es posible determinar la probabilidad de tener exactamente 88 personas en la muestra que suelen no participar en las elecciones? Argumenten su respuesta.

SOLUCIÓN:

Se considerarán dos casos:

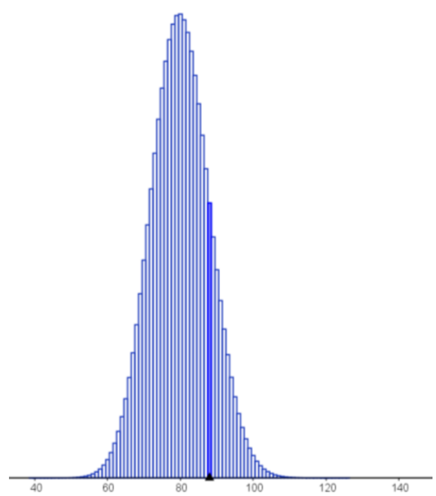
- Forma 1: Se puede resolver usando la Distribución Binomial, ya que el valor pedido pertenece al recorrido, que en este caso es $Rec(X) = \{0,1,2,\ldots,400\}$. De esta forma, se pide $P(X = 88)$, así:

  $P(X = 88) = \binom{400}{88} \cdot (0{,}2)^{88} \cdot (0{,}8)^{400-88}$

  $= \binom{400}{88} \cdot (0{,}2)^{88} \cdot (0{,}8)^{312}$

  $= 0{,}0295 \quad (*)$

  $(*)$: Valor calculado con GeoGebra.



  Por lo tanto, sí es posible y la probabilidad de tener exactamente 88 personas que suelen no participar en las elecciones es de un 2,95%.

- Forma 2: Si se aproxima y trabaja usando la Distribución Normal, no se puede calcular la probabilidad, ya que corresponde al área en un valor puntual del recorrido, y este es cero, es decir:

$$P(X = 88) = 0$$

Sin embargo, es posible realizar la aproximación, de forma que se tiene que

$P(X = 88)$

$\approx P(88 - 0,5 \leq X \leq 88 + 0,5)$

$= P(87,5 \leq X \leq 88,5)$

$= P\left(\frac{87,5-80}{8} \leq \frac{X-80}{8} \leq \frac{88,5-80}{8}\right)$

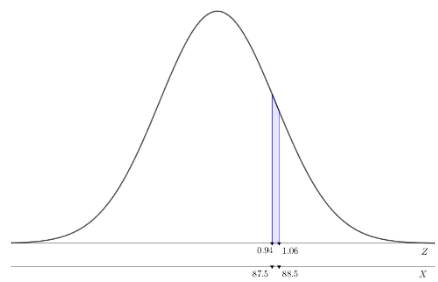
$\approx P\left(\frac{87,5-80}{8} \leq \frac{X-80}{8} \leq \frac{88,5-80}{8}\right)$

$= P(0,94 \leq Z \leq 1,06)$

$= P(Z \leq 1,06) - P(Z \leq 0,94)$

$= 0,8554 - 0,8264$

$= 0,029$

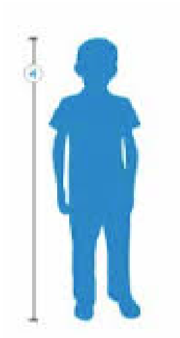


z	0.03	0.04	0.05
0.8	0.7967	0.7995	0.8023
0.9	0.8238	0.8264	0.8289
1.0	0.8485	0.8508	0.8531

z	0.05	0.06	0.07
0.9	0.8289	0.8315	0.8340
1.0	0.8531	0.8554	0.8577
1.1	0.8749	0.8770	0.8790

Por lo tanto, sí es posible (con la aproximación) y la probabilidad de tener exactamente 88 personas que suelen no participar en las elecciones es de un 2,9%.

PASO 9. APLICACIÓN DE LA DISTRIBUCIÓN NORMAL



Conexión Interdisciplinaria: **Ciencias para la Ciudadanía** OA c, 3° y 4° medio.

A. Según el estudio antropométrico en párvulos atendidos por el Sistema Educativo Público Chileno para el diseño de mobiliario, publicado en la revista científica Scielo[3], la estatura media de niños en el grupo de 25 a 36 meses es de $\mu = 91{,}30\ cm$, con una desviación estándar de $\sigma = 4{,}27\ cm$.

a. ¿Cuál es el porcentaje de párvulos que tienen una estatura de $87\ cm$ como máximo? Argumenten su respuesta.

SOLUCIÓN:

Sea

$$X: Estatura\ media\ de\ los\ niños\ en\ el\ grupo\ de\ 25\ a\ 36\ meses, en\ cm.$$

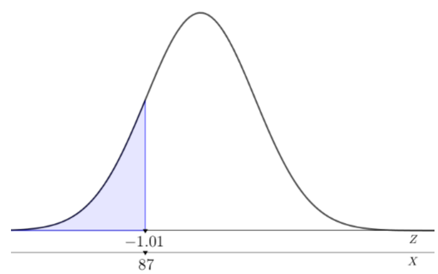
Se tiene que $X \sim N(\mu = 91{,}3; \sigma^2 = 4{,}27^2)$, luego se pide $P(X \leq 87)$, así

$P(X \leq 87) = P\left(Z \leq \frac{87-91{,}3}{4{,}27}\right)$

$= P(Z \leq -1{,}01)$

$= \phi(-1{,}01)$

$= 0{,}1562$



z	0	0.01	0.02
-1.1	0.1357	0.1335	0.1314
-1.0	0.1587	0.1562	0.1539
-0.9	0.1841	0.1814	0.1788

En consecuencia, la probabilidad de encontrar párvulos que tengan una estatura de $87\ cm$ o menos es de $15{,}62\%$.

b. ¿Cuál es el porcentaje de párvulos que tienen una estatura de $87\ cm$ como mínimo y $96\ cm$ como máximo? Argumenten su respuesta.

SOLUCIÓN

Usando la misma variable aleatoria en a., se pide $P(87 \leq X \leq 96)$, así

$P(87 \leq X \leq 96)$

$= P\left(\frac{87-91{,}3}{4{,}27} \leq Z \leq \frac{96-91{,}3}{4{,}27}\right)$

3 https://scielo.conicyt.cl/scielo.php?script=sci_arttext&pid=S0717-95022013000100032

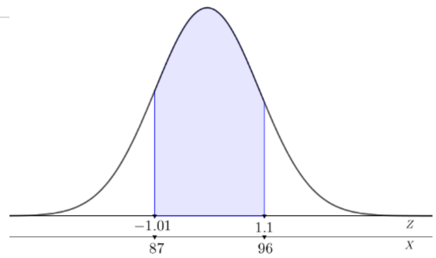
$= P(-1{,}01 \leq Z \leq 1{,}10)$

$= P(Z \leq 1{,}10) - P(Z < -1{,}01)$

$= \phi(1{,}10) - \phi(-1{,}01)$

$= 0{,}8643 - 0{,}1562$

$= 0{,}7081$



z	0	0.01	0.02
-1.1	0.1357	0.1335	0.1314
-1.0	0.1587	0.1562	0.1539
-0.9	0.1841	0.1814	0.1788

z	0	0.01	0.02
1.0	0.8413	0.8438	0.8461
1.1	0.8643	0.8665	0.8686
1.2	0.8849	0.8869	0.8888

Por lo tanto, la probabilidad de que los párvulos tengan una estatura entre $87\ cm$ y $96\ cm$ es de un $70{,}81\%$.

Nota al docente: En este caso, no se necesita hacer diferencias entre $<$ y $\leq$, ya que se está trabajando con una variable aleatoria continua. Además, debe utilizarse la propiedad $P(a \leq X \leq b) = P(X \leq b) - P(X < a)$.

OBJETIVO DE PREGUNTAS COMPLEMENTARIAS

- Las preguntas complementarias permiten la aplicación de la desestandarización, donde los estudiantes deben extraer información de los enunciados para plantear lo que se pide y resolverlo.

DESARROLLO

c. ¿A partir de qué estatura los niños se encuentran sobre el 12% superior de la población?

SOLUCIÓN

Usando la variable definida en a., se pide determinar un puntaje x, es decir, un valor de la estatura tal que un 12% de los niños tengan estaturas mayores a esta. Así:

$P(X > x) = 0{,}12$

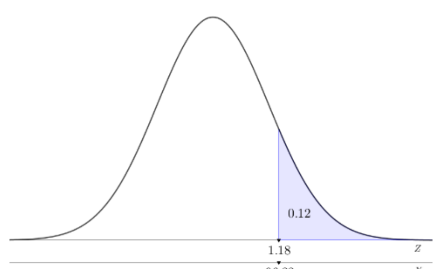
$\Rightarrow 1 - P(X \leq x) = 0{,}12$

$\Rightarrow P(X \leq x) = 1 - 0{,}12$

$\Rightarrow P(X \leq x) = 0{,}88$

$\Rightarrow P\left(Z \leq \frac{x-91{,}3}{4{,}27}\right) = 0{,}88$

$\Rightarrow \phi\left(\frac{x-91{,}3}{4{,}27}\right) = 0{,}88$



z	0.07	0.08	0.09
1.0	0.8577	0.8599	0.8621
1.1	0.8790	0.8810	0.8830
1.2	0.8980	0.8997	0.9015

$$\Rightarrow \phi^{-1}\left(\phi\left(\frac{x-91{,}3}{4{,}27}\right)\right) = \phi^{-1}(0{,}88)$$

$$\Rightarrow \frac{x-91{,}3}{4{,}27} = \phi^{-1}(0{,}88)$$

$$\Rightarrow \frac{x-91{,}3}{4{,}27} = 1{,}18$$

$$\Rightarrow x = 91{,}3 + 4{,}27 \cdot 1{,}18$$

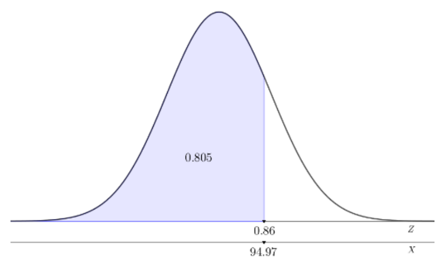
$$\Rightarrow x = 96{,}33$$

Por lo tanto, el 12% de los niños tienen estatura superior a $96{,}33\ cm$.

d. ¿Hasta que estatura, los niños se encuentran en el $80{,}5\%$ de las estaturas?

SOLUCIÓN:

Usando la variable definida en a., se pide determinar un puntaje x, es decir, un valor de la estatura tal que el $80{,}5\%$ de los niños tengan estaturas menores a esta. Así:



z	0.05	0.06	0.07
0.7	0.7734	0.7764	0.7794
0.8	0.8023	0.8051	0.8078
0.9	0.8289	0.8315	0.8340

$$P(X \leq x) = 0{,}805$$

$$\Rightarrow P\left(Z \leq \frac{x-91{,}3}{4{,}27}\right) = 0{,}805$$

$$\Rightarrow \phi\left(\frac{x-91{,}3}{4{,}27}\right) = 0{,}805$$

$$\Rightarrow \phi^{-1}\left(\phi\left(\frac{x-91{,}3}{4{,}27}\right)\right) = \phi^{-1}(0{,}805)$$

$$\Rightarrow \frac{x-91{,}3}{4{,}27} = \phi^{-1}(0{,}805)$$

$$\Rightarrow \frac{x-91{,}3}{4{,}27} = 0{,}86$$

$$\Rightarrow x = 91{,}3 + 4{,}27 \cdot 0{,}86$$

$$\Rightarrow x = 94{,}97$$

Por lo tanto, el $80{,}5\%$ de los niños tienen estatura menor a $94{,}97\ cm$.

B. La vida útil de los motores de un automóvil de cierta marca se distribuye según una distribución Normal con el valor esperado de $\mu = 105\,000\ km$ y una desviación estándar de $\sigma = 10\,000\ km$.

a. ¿Qué porcentaje de los motores tiene una vida útil entre $90\,000\ km$ y $110\,000\ km$? Argumenten su respuesta.

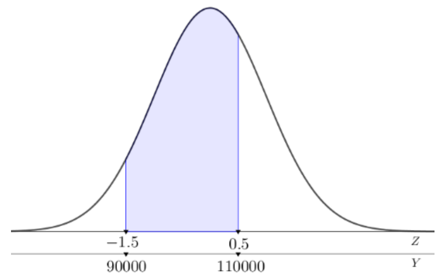
SOLUCIÓN:

Sea:

$$Y: Vida\ útil\ de\ los\ motores\ de\ un\ automóvil, en\ km.$$

Se tiene que $Y \sim N(\mu = 105\,000; \sigma^2 = 10\,000^2)$

Luego, se pide $P(90\,000 < Y < 110\,000)$, así:



$P(90\,000 < Y < 110\,000)$

$= P\left(\frac{90\,000-105\,000}{10\,000} < Z < \frac{110\,000-105\,000}{10\,000}\right)$

$= P(-1{,}5 < Z \leq 0{,}5)$

$= P(Z < 0{,}5) - P(Z \leq -1{,}5)$

$= \phi(0{,}5) - \phi(-1{,}5)$

$= 0{,}6915 - 0{,}0668$

$= 0{,}6247$

z	0	0.01	0.02
-1.6	0.0548	0.0537	0.0526
-1.5	0.0668	0.0655	0.0643
-1.4	0.0808	0.0793	0.0778

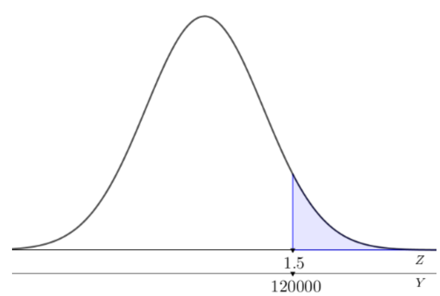
z	0	0.01	0.02
0.4	0.6554	0.6591	0.6628
0.5	0.6915	0.6950	0.6985
0.6	0.7257	0.7291	0.7324

Por lo tanto, el porcentaje de los motores que tiene una vida útil tiene una vida útil entre $90\,000\ km$ y $110\,000\ km$ es de 62,47%.

b. Qué porcentaje de los motores supera la vida útil de $120\,000\ km$? Argumenten su respuesta.

<u>SOLUCIÓN</u>

Usando la misma variable aleatoria definida en a., se pide $P(Y > 120\,000)$, así:



$P(Y > 120\,000) = 1 - P(Y \leq 120\,000)$

$= 1 - P\left(Z \leq \frac{120\,000-105\,000}{10\,000}\right)$

$= 1 - P(Z \leq 1{,}5)$

$= 1 - \phi(1{,}5)$

$= 1 - 0{,}9332$

$= 0{,}0668$

z	0	0.01	0.02
1.4	0.9192	0.9207	0.9222
1.5	0.9332	0.9345	0.9357
1.6	0.9452	0.9463	0.9474

Por lo que, el porcentaje de los motores que supera la vida útil de $120\,000\ km$ es de 6,68%.

c. ¿Qué porcentaje de los motores se desvía del valor esperado en más de $12\,000\ km$? Argumenten su respuesta.

SOLUCIÓN:

Usando la misma variable aleatoria definida en a., se pide

$$P(|Y - 105\,000| > 12\,000)$$

Así:

$P(|Y - 105\,000| > 12\,000)$

$= P\big((Y - 105\,000 < -12\,000) \cup (Y - 105\,000 > 12\,000)\big)$

$= P(Y - 105\,000 < -12\,000) + P(Y - 105\,000 > 12\,000)$

$= P(Y < 93\,000) + P(Y > 117\,000)$

$= P(Y < 93\,000) + [1 - P(Y \leq 117\,000)]$

$= 1 + P(Y < 93\,000) - P(Y \leq 117\,000)$

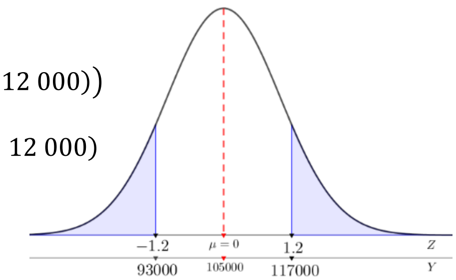
$= 1 + P\left(Z < \frac{93\,000 - 105\,000}{10\,000}\right) - P\left(Z \leq \frac{117\,000 - 105\,000}{10\,000}\right)$

$= 1 + P(Z < -1{,}2) - P(Z \leq 1{,}2)$

$= 1 + \phi(-1{,}2) - \phi(1{,}2)$

$= 1 + 0{,}1151 - 0{,}8849$

$= 0{,}2302$



z	0	0.01	0.02
-1.3	0.0968	0.0951	0.0934
-1.2	0.1151	0.1131	0.1112
-1.1	0.1357	0.1335	0.1314

z	0	0.01	0.02
1.1	0.8643	0.8665	0.8686
1.2	0.8849	0.8869	0.8888
1.3	0.9032	0.9049	0.9066

En consecuencia, el porcentaje de los motores que se desvían del valor esperado en más de $12\,000\ km$ es de $23{,}02\%$.

Nota al docente: Se sugiere recordar la propiedad de desigualdad con valor absoluto, es decir, $|a| > b \Rightarrow a < -b \cup a > b$. Además, se debe utilizar la propiedad del complemento, $P(X > a) = 1 - P(X \leq a)$.

OBJETIVO DE PREGUNTAS COMPLEMENTARIAS

- Las preguntas complementarias permiten la aplicación de la desestandarización, donde los estudiantes deben extraer información de los enunciados para plantear lo que se pide y resolverlo.

DESARROLLO

d. La empresa que fabrica los motores hace constantemente pruebas con ellos, saca los que se encuentran con una vida útil de un 10% o menos. ¿Hasta qué valor de vida útil el motor no pasa las pruebas?

SOLUCIÓN

Usando la variable definida en a., se pide determinar un puntaje y, es decir, un valor de la vida útil de los motores, tal que un 10% de ellos tenga una vida útil menor a esta. Así:

$P(Y \leq y) = 0{,}1$

$\Longrightarrow P\left(Z \leq \frac{y-105\,000}{10\,000}\right) = 0{,}1$

$\Longrightarrow \phi\left(\frac{y-105\,000}{10\,000}\right) = 0{,}1$

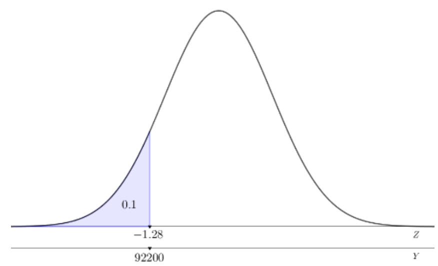
$\Longrightarrow \phi^{-1}\left(\phi\left(\frac{y-105\,000}{10\,000}\right)\right) = \phi^{-1}(0{,}1)$

$\Longrightarrow \frac{y-105\,000}{10\,000} = \phi^{-1}(0{,}1)$

$\Longrightarrow \frac{y-105\,000}{10\,000} = -1{,}28$

$\Longrightarrow y = 105\,000 + 10\,000 \cdot (-1{,}28)$

$\Longrightarrow y = 92\,200$



z	0.07	0.08	0.09
-1.3	0.0853	0.0838	0.0823
-1.2	0.1020	0.1003	0.0985
-1.1	0.1210	0.1190	0.1170

Por lo tanto, el 10% de los motores tiene una vida útil menor a $92\,200\ km$, por ende, no pasará las pruebas, por lo que deben ser sacados.

e. Se considera un riesgo que el motor tenga una vida útil en el $2{,}5\%$ más alto de la distribución, ¿desde qué valor se encuentra en esa categoría?

SOLUCIÓN

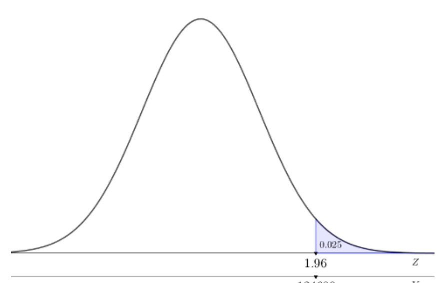
Usando la variable definida en a., se pide determinar un puntaje y, es decir, un valor de la vida útil de los motores, tal que un $2{,}5\%$ de ellos tenga una vida útil mayor a esta. Así:

$P(Y > y) = 0{,}025$

$\Longrightarrow 1 - P(Y \leq y) = 0{,}025$

$\Longrightarrow P(Y \leq y) = 1 - 0{,}025$

$\Longrightarrow P(Y \leq y) = 0{,}975$



z	0.05	0.06	0.07
1.8	0.9678	0.9686	0.9693
1.9	0.9744	0.9750	0.9756
2.0	0.9798	0.9803	0.9808

$$\Rightarrow P\left(Z \leq \frac{y-105\,000}{10\,000}\right) = 0{,}975$$

$$\Rightarrow \phi\left(\frac{y-105\,000}{10\,000}\right) = 0{,}975$$

$$\Rightarrow \phi^{-1}\left(\phi\left(\frac{y-105\,000}{10\,000}\right)\right) = \phi^{-1}(0{,}975)$$

$$\Rightarrow \frac{y-105\,000}{10\,000} = \phi^{-1}(0{,}975)$$

$$\Rightarrow \frac{y-105\,000}{10\,000} = 1{,}96$$

$$\Rightarrow y = 105\,000 + 10\,000 \cdot 1{,}96$$

$$\Rightarrow y = 124\,600$$

Por lo tanto, desde $y = 124\,600\ km$ se encuentra en riesgo el motor.

DESCRIPCIÓN DE LA PREGUNTA COMPLEMENTARIA

- Esta pregunta permite a los estudiantes aplicar lo aprendido a lo largo de esta *actividad 1: Comprender el modelo Normal de probabilidades*. Con ella, calculan probabilidades estandarizando, utilizan los intervalos icónicos, y desestandarizan. Además, viene a potenciar las dos actividades anteriores que tienen una finalidad parecida, correspondiente a aplicar los contenidos de la actividad.

DESARROLLO

C. El consumo promedio de pan en Chile ronda los $90\,kg$ por persona al año, para su elaboración se utiliza principalmente harina. Esta materia prima posee distintos niveles de calidad basado de la cantidad de proteína presentes en ella. Específicamente en Santiago se realizó un estudio con 1500 panaderías sobre la calidad del pan producido en función de la cantidad de harina que usan. La categorización se hizo respecto a la ración de harina requerida para 1 kilogramo de pan, teniéndose los niveles bajo, medio y alto, de modo que el primero concentra el $26{,}11\%$ de las panaderías, el nivel medio considera el $66{,}54\%$ y el resto pertenece al último nivel. Además, se sabe que la cantidad promedio de harina para 1 kilogramo de pan sigue una distribución Normal de parámetros μ desconocido, con una variabilidad de $\sigma = 40\ g$.

a. Si se sabe que 1431 panaderías utilizan entre 525 y 685 gramos de harina, ¿cuál es el promedio de la cantidad de harina utilizado por las panaderías de Santiago para elaborar 1 kilogramo de pan?

SOLUCIÓN:

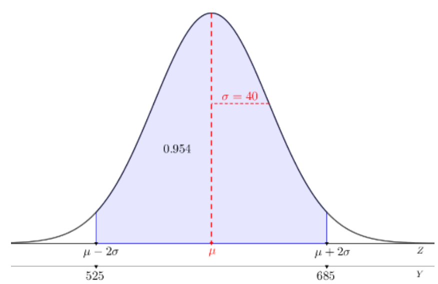
Sea:

$$Y: Cantidad\ de\ harina, en\ gramos, utilizada\ en\ 1\ kg\ de\ pan$$

Se sabe que 1431 panaderías utilizan entre 525 y 685 gramos de harina, por lo que, se determinará su respectivo valor en porcentaje

$$\begin{array}{ccc} 1500 & \longrightarrow & 100\% \\ 1431 & \longrightarrow & y \end{array}$$

Así, $y = \frac{1431 \cdot 100}{1500} = 95{,}4\%$. Luego, 95,4% de las panaderías usan entre 525 y 685 gramos de harina. Ahora, hay que recordar que la probabilidad que la variable aleatoria se aleje de su media en a lo más 2 desviaciones estándar es 95,4%, es decir:



$$P(\mu - 2\sigma < Y < \mu + 2\sigma) = 0{,}954$$

Lo que puesto en contexto es

$$P(525 < Y < 685) = 0{,}954$$

Se puede observar que $\mu - 2\sigma = 525$. Si $\sigma = 40\ g$, entonces

$$\mu - 2 \cdot 40 = 525 \Longrightarrow \mu - 80 = 525 \Longrightarrow \mu = 525 + 80 \Longrightarrow \mu = 605$$

Por lo tanto, se tiene que la cantidad media de harina utilizada para la elaboración de 1 kilogramo de pan es de $605\ g$. Así, $Y \sim N(\mu = 605\ ; \sigma^2 = 40^2)$.

b. Si se escoge una panadería al azar, ¿cuál es la probabilidad de que utilice más de 590 gramos de harina para la producción de $1\ kg$ de pan?

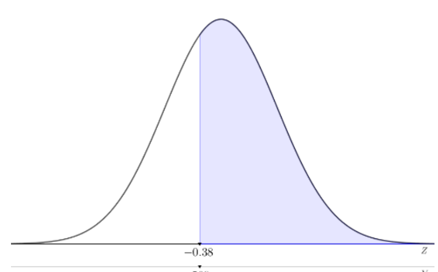
SOLUCIÓN:

Utilizando la variable definida en a. se pide $P(Y > 590)$. Así:

$P(Y > 590)$

$= 1 - P(Y \leq 590)$

$= 1 - P\left(\frac{Y-605}{40} \leq \frac{590-605}{40}\right)$



$= 1 - P\left(Z \leq -\frac{3}{8}\right)$

z	0.07	0.08	0.09
-0.4	0.3192	0.3156	0.3121
-0.3	0.3557	0.3520	0.3483
-0.2	0.3936	0.3897	0.3859

$= 1 - P(Z \leq -0{,}38)$

$= 1 - \phi(-0{,}38)$

$= 1 - 0{,}352$

$= 0{,}648$

Por lo que, la probabilidad de que una panadería utilice menos de $590\ g$ de harina es de un $64{,}8\%$

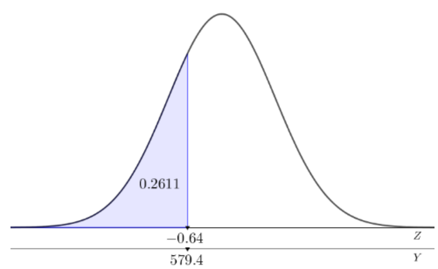
c. Si una panadería pertenece al nivel de calidad medio, ¿Entre qué valores se encuentra la cantidad de harina utilizada por el local? Bosqueja las campanas correspondientes a los tres niveles de calidad según la cantidad de harina que usan.

SOLUCIÓN:

Se utilizará la misma variable definida en a. Para determinar los valores de corte, se requiere de las probabilidades correspondientes a cada nivel de calidad.

El límite inferior del nivel de calidad medio se calculará mediante $P(Y < a) = 0{,}2611$, así:

$P(Y < a) = 0{,}2611$



z	0.03	0.04	0.05
-0.7	0.2327	0.2296	0.2266
-0.6	0.2643	0.2611	0.2578
-0.5	0.2981	0.2946	0.2912

$\Rightarrow P\left(\frac{Y-605}{40} < \frac{a-605}{40}\right) = 0{,}2611$

$\Rightarrow P\left(Z < \frac{a-605}{40}\right) = 0{,}2611$

$\Rightarrow \phi\left(\frac{a-605}{40}\right) = 0{,}2611$

$\Rightarrow \phi^{-1}\left(\phi\left(\frac{a-605}{40}\right)\right) = \phi^{-1}(0{,}2611)$

$\Rightarrow \frac{a-605}{40} = \phi^{-1}(0{,}2611)$

$\Rightarrow \frac{a-605}{40} = -0{,}64$

$\Rightarrow a - 605 = (-0{,}64) \cdot 40$

$\Rightarrow a = 605 + (-0{,}64) \cdot 40$

$\Rightarrow a = 579{,}4$

Luego, el límite inferior de la calidad media es $579{,}4\ g$ de harina. Por otro lado, el límite superior será $P(Y > b) = 0{,}0735$, ya que el nivel de calidad alto está dado por $1 - 0{,}2611 - 0{,}6654 = 0{,}0735$. Así:

$P(Y > b) = 0{,}0735$

$\Rightarrow 1 - P(Y \leq b) = 0{,}0735$

$\Rightarrow P(Y \leq b) = 1 - 0{,}0735$

$\Rightarrow P\left(\frac{Y-605}{40} \leq \frac{b-605}{40}\right) = 0{,}9265$

$\Rightarrow P\left(Z \leq \frac{b-605}{40}\right) = 0{,}9265$

$\Rightarrow \phi\left(\frac{b-605}{40}\right) = 0{,}9265$

$\Rightarrow \phi^{-1}\left(\phi\left(\frac{b-605}{40}\right)\right) = \phi^{-1}(0{,}9265)$

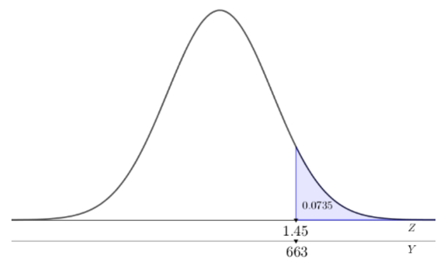
$\Rightarrow \frac{b-605}{40} = \phi^{-1}(0{,}9265)$

$\Rightarrow \frac{b-605}{40} = 1{,}45$

$\Rightarrow b - 605 = 1{,}45 \cdot 40$

$\Rightarrow b = 605 + 1{,}45 \cdot 40$

$\Rightarrow b = 663$

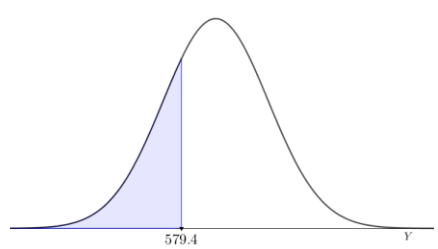


z	0.04	0.05	0.06
1.3	0.9099	0.9115	0.9131
1.4	0.9251	0.9265	0.9279
1.5	0.9382	0.9394	0.9406

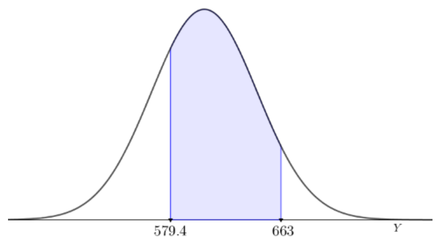
z	0.04	0.05	0.06
1.3	0.9099	0.9115	0.9131
1.4	0.9251	0.9265	0.9279
1.5	0.9382	0.9394	0.9406

Así, el límite superior de la calidad media es $663\ g$ de harina.

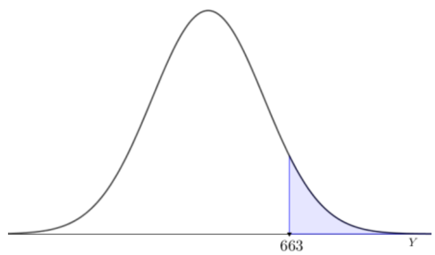
Respecto a los bosquejos de las campanas se tiene que el pan de una panadería será considerado bajo si la cantidad de harina que usan es menor a $579{,}4\ g$ para $1\ kg$ de pan.



Si utiliza entre $579{,}4\ g$ y $663\ g$ (ambos inclusive), el producto de la panadería será considerado de nivel medio.



Por último, si la cantidad de harina supera los $663\,g$, el pan será considerado de un nivel alto.



Nota al docente: Para calcular el límite superior del intervalo de calidad medio, también se podría haber sumado $0{,}2611$ y $0{,}6654$, y haber obtenido la probabilidad $P(Y \leq 0{,}2611 + 0{,}6654)$.

ORIENTACIONES PARA EL DOCENTE DE LA PROPUESTA MINISTERIAL

1. Es posible que los estudiantes ya conozcan la distribución Normal de probabilidades. En estas actividades se propone profundizar en los conceptos y cómo llegar a ellos. Por ejemplo, se parte de experimentos que utilizan variables aleatorias continuas, y se sigue con gráficos de barras que representan experimentos Binomiales, en los cuales, al aumentar el número de repeticiones "n", la forma del gráfico de barra se parece cada vez más a una campana de Gauss.

2. Las actividades propuestas permiten que los jóvenes entiendan, paso a paso, el modelo Normal de probabilidades. Además, discuten si es necesario incorporar la distribución Normal estándar y cómo determinar las probabilidades que corresponden a los intervalos $[\mu - 1 \cdot \sigma, \mu + 1 \cdot \sigma]$, $[\mu - 2 \cdot \sigma, \mu + 2 \cdot \sigma]$ y $[\mu - 3 \cdot \sigma, \mu + 3 \cdot \sigma]$ e interpretar su significado.

3. Se sugiere proponer ejercicios en contexto en los que tengan que aproximar una distribución Binomial a una distribución Normal, para que vean la utilidad de esta transformación cuando el número "n" de observaciones es grande.

4. Para finalizar, se recomienda que resuelvan problemas en contexto en que deban aplicar la distribución Normal con parámetros μ y σ dados. La idea es que

resuelvan cálculos de probabilidad, utilizando la distribución Normal estandarizada.

5. Se sugiere los siguientes indicadores para evaluar formativamente los aprendizajes:

   - Identifican las principales características de una distribución Normal de probabilidades.
   - Resuelven problemas que involucran los modelos Binomial y Normal.

ORIENTACIONES DOCENTES DE LA PROPUESTA COMPLEMENTARIA

1. Se sugiere enfatizar en ítem C del paso 3, sobre todo en la gráfica de la pregunta f., ya que estos conocimientos servirán para institucionalizar posteriormente el teorema de estandarización, así los estudiantes podrán vincular lo que se realizó en esa actividad con lo que sucede cuando se realiza el cambio de forma y escala en la variable aleatoria.
2. Antes de comenzar el paso 4, se recomienda al docente haber institucionalizado la notación $\phi(z)$, de esta forma, los estudiantes podrán familiarizarse con esta notación para enfrentarse a las futuras actividades de esta unidad. También se sugiere solicitar a los estudiantes dibujar a priori la gráfica de la campana de cada ejercicio, de esta forma visualizar la correspondencia del valor en la escala Z y la variable original.
3. Una vez que finalice el paso 5 se sugiere institucionalizar la propiedad de simetría presentándola de forma gráfica y resolviendo algunos ejemplos donde entregue solo la probabilidad de un valor en la tabla Z y se solicite calcular la probabilidad de su opuesto.
4. Antes de iniciar la actividad A del paso 7, se sugiere recordar la definición de cuantil.
5. Se recomienda solicitar a los estudiantes utilizar la notación ϕ y ϕ^{-1} para las diversas actividades que lo requieran.

RECURSOS Y SITIOS WEB

Sitios web sugeridos para estudiantes y profesores

- Generadores de números aleatorios: https://www.curriculumnacional.cl/link/http://www.generarnumerosaleatorios.com/

- https://www.curriculumnacional.cl/link/https://www.augeweb.com/azar/
- https://www.curriculumnacional.cl/link/https://numero-aleatorio.com/generadores/
- https://www.curriculumnacional.cl/link/https://pinetools.com/es/generador-numerosaleatorios
- Cómo usar la fórmula de la distribución Binomial en Excel https://www.curriculumnacional.cl/link/https://www.youtube.com/watch?v=xVwetXD9cis

ACTIVIDAD 2: APLICAR EL MODELO NORMAL EN EL TRANSPORTE DE PERSONAS

PROPÓSITO

Los estudiantes usan la distribución Normal para determinar probabilidades de interés, convirtiendo los valores de la variable aleatoria en su puntuación correspondiente. Se espera que sean empáticos con la situación planteada y que no respondan basándose en prejuicios sobre el peso de las personas y el transporte. La tarea se potencia con la interpretación de los resultados en el contexto y el uso de la planilla de cálculo como apoyo, aprovechando así las herramientas disponibles. También avanzan en el análisis de lo que ocurre con las medias de muestras tomadas al azar.

OBJETIVOS DE APRENDIZAJE

OA 3. Modelar fenómenos o situaciones cotidianas del ámbito científico y del ámbito social, que requieran el cálculo de probabilidades y la aplicación de las distribuciones Binomial y Normal.

- OA b. Resolver problemas que impliquen variar algunos parámetros en el modelo utilizado y observar cómo eso influye en los resultados obtenidos.
- OA e. Construir modelos, realizando conexiones entre variables para predecir posibles escenarios de solución a un problema, y tomar decisiones fundamentadas.
- OA i. Buscar, seleccionar, manejar y producir información matemática/cuantitativa confiable a través de la web.

ACTITUDES

- Trabajar con empatía y respeto en el contexto de la diversidad, eliminando toda expresión de prejuicio y discriminación.
- Aprovechar las herramientas disponibles para aprender y resolver problemas.

DURACIÓN

12 horas pedagógicas.

DESARROLLO

TRANSPORTE EN TELEFÉRICO

Conexión Interdisciplinaria: **Ciencias para la Ciudadanía** OA c, 3° y 4° medio.

El Parque Metropolitano de Santiago es el cuarto parque urbano más grande del mundo, con más de 1785 hectáreas de extensión. Esto lo convierte en el pulmón verde más grande de Latinoamérica, que resalta gracias a los atractivos que posee. Es uno de los principales espacios de recreación, naturaleza, cultura y deporte de la capital. Muchos visitantes usan el teleférico, pues permite ver Santiago desde las alturas y apreciar la belleza de la capital de nuestro país. Cada cabina tiene una capacidad de $480\,kg$ o 6 personas. Por otro lado, varios estudios estadísticos (como la Encuesta Nacional de Salud) indican que la masa promedio de los hombres es de $77{,}3\,kg$, con una desviación estándar de $12{,}9\,kg$ y la masa promedio de las mujeres, $67\,kg$. ¡Imagina qué ocurriría si las normas de seguridad del Parque Metropolitano no tuvieran en cuenta esta información!

1. En grupos, respondan a las siguientes preguntas: ¿Qué información aporta saber que la masa de los hombres se distribuye normalmente? ¿Qué les permite saber y/o inferir? Argumenten su respuesta.

   SOLUCIÓN:

   Que la masa de los hombres distribuya normalmente, ayuda a inferir que la media de la masa de los hombres es también la moda y la mediana, y que la mayor parte de los datos se concentra alrededor de la esperanza, por el contrario, la cantidad de individuos que está en las colas es la menor parte de los datos. Además, nos permite saber la media y desviación estándar de la población en estudio.

   Nota al docente: Se puede solicitar a los estudiantes que bosquejen la gráfica de la Distribución Normal, para que identifiquen los parámetros.

2. ¿Qué sucedería en la cabina si se suben 7 hombres cuya masa total aproximadamente es la media? Argumenten su respuesta.

   SOLUCIÓN:

   Lo más intuitivo desde una perspectiva aritmética es multiplicar la masa de los 7 hombres por la masa promedio aproximada, y darse cuenta de que es mayor a $480\,kg$, ya que:

$$7 \cdot 77{,}3 = 541{,}1\,kg$$

Sin embargo, mirándolo desde un punto de vista lógico, se puede considerar que la número de hombres que subieron a la cabina sobrepasa la cantidad máxima de pasajeros.

Por lo tanto, cualquiera sea el razonamiento usado por los estudiantes, se llega a la misma conclusión, que es probable que el teleférico tenga sobrecarga de masa.

3. ¿Qué sucedería en la cabina si se sube un hombre que tiene una masa mayor que la media; por ejemplo: $80\ kg$? Argumenten su respuesta.

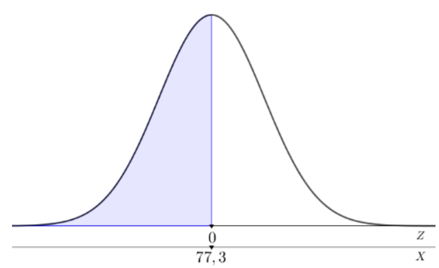
SOLUCIÓN:

Si el hombre es solo uno, y se sube sin nadie más, no afecta en nada al teleférico. Ahora si hay más personas en la cabina, hay que ver que la suma de las masas no exceda la capacidad total.

4. ¿Cuál es la probabilidad de que, al seleccionar aleatoriamente a un hombre, su masa sea menor a $77{,}3\ kg$? Argumenten su respuesta utilizando solo la campana de Gauss.

SOLUCIÓN:

Sea:



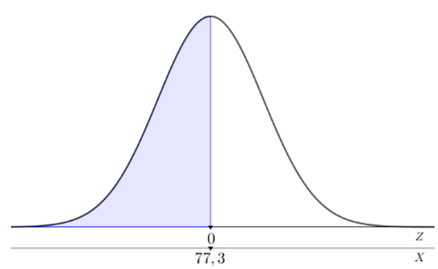
$$X: Masa\ del\ hombre, en\ kg.$$

La probabilidad es inmediata, ya que, bajo la media, se acumula un 50%.

$$P(X < \mu) = P(X < 77{,}3) = 0{,}5$$

a. Calcule la probabilidad pedida haciendo uso de la tabla Z.

SOLUCIÓN:



z	0	0.01	0.02
0.0	0.5000	0.5040	0.5080
0.1	0.5398	0.5438	0.5478
0.2	0.5793	0.5832	0.5871

Los valores son

$X = 77{,}3\ ;\ \mu = 77{,}3\ ;\ \sigma = 12{,}9$

Estandarizando

$Z = \frac{X-\mu}{\sigma} = \frac{77{,}3-77{,}3}{12{,}9} = \frac{0}{12{,}9} = 0$

Luego

$P(Z < 0) = \phi(0) = 0{,}5$

b. ¿Cómo aporta y se usa la tabla de distribución Normal estándar en este caso? Argumenten su respuesta.

SOLUCIÓN:

Para este caso en particular, a través de la aritmetización el aporte de la estandarización de la Distribución Normal permite comprobar teóricamente que la probabilidad encontrada en la gráfica acampanada es concordante con el valor de la tabla Z.

En general, de la tabla se puede inferir que el 0 es la mitad de la campana, ya que su probabilidad se puede encontrar en la parte de los valores positivos como negativos, y acumula $0{,}5$.

c. Usando el resultado anterior, determinen la probabilidad de que se elija al azar a un hombre con una masa mayor a la media. Expliquen los cálculos realizados.

SOLUCIÓN:

La probabilidad es inmediata, porque sobre la media se acumula un 50%, entonces la necesidad de usar tabla no es imprescindible.

$$P(X > \mu) = P(X > 77{,}3) = 0{,}5$$

Sin embargo, el ejercicio espera que utilicen la tabla ya que menciona "usando el resultado anterior", así:

$P(X > \mu) = P(X > 77{,}3)$

$= 1 - P(X \leq 77{,}3)$

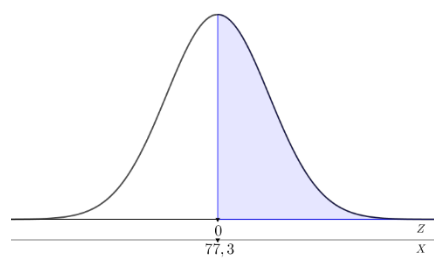
$= 1 - P\left(\frac{X-77{,}3}{12{,}9} \leq \frac{77{,}3-77{,}3}{12{,}9}\right)$

$= 1 - P(Z \leq 0)$

$= 1 - \phi(0)$

$= 1 - 0{,}5$

$= 0{,}5$



z	0	0.01	0.02
0.0	0.5000	0.5040	0.5080
0.1	0.5398	0.5438	0.5478
0.2	0.5793	0.5832	0.5871

Por lo tanto, la probabilidad de que un hombre tenga una masa mayor a la media es de un 50%.

5. ¿Cuál es la probabilidad de que, al seleccionar aleatoriamente a un hombre, su masa sea mayor a $80\ kg$? Expliquen los cálculos realizados.

SOLUCIÓN:

Sea:

$$X\text{: } Masa\ de\ los\ hombres\ que\ suben\ al\ teleférico, en\ kg.$$

De esta forma, $X \sim N(\mu = 77{,}3;\ \sigma)$, y se pide que $P(X > 80)$, por lo que:

$P(X > 80) = 1 - P(X \leq 80)$

$= 1 - P\left(\frac{X-77{,}3}{\sigma} \leq \frac{80-77{,}3}{12{,}9}\right)$

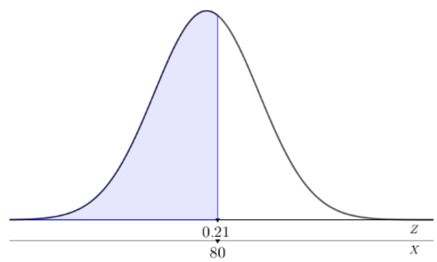
$= 1 - P\left(Z \leq \frac{80-77{,}3}{12{,}9}\right)$

$= 1 - P(Z \leq 0{,}21)$

$= 1 - \phi(0{,}21)$

$= 1 - 0{,}5832$

$= 0{,}4168$



z	0	0.01	0.02
0.1	0.5398	0.5438	0.5478
0.2	0.5793	0.5832	0.5871
0.3	0.6179	0.6217	0.6255

Por lo tanto, la probabilidad de que la masa de un hombre sea mayor a $80\ kg$, es de un $41{,}68\%$.

6. ¿Cómo podrían usar el resultado anterior para evaluar la seguridad del teleférico, según el criterio de capacidad y cantidad de personas? Argumenten su respuesta.

SOLUCIÓN:

La finalidad de esta pregunta es concluir respecto a la seguridad, por lo que, se busca un vínculo entre la Distribución Normal y la vida real.

Se usa para evaluar seguridad del teleférico, en otras palabras, que no exceda la capacidad máxima de masa permitida. Según el resultado anterior, es más probable que un hombre tenga una masa menor a $80\ kg$, por lo que el teleférico no debería tener problemas con las normas de seguridad.

DETERMINAR PROBABILIDADES QUE IMPLIQUEN EL USO DEL TEOREMA DEL LÍMITE CENTRAL

1. ¿Qué sucedería si se suben 6 hombres a una cabina y todos ellos tienen una masa mayor que la media; por ejemplo: $80\ kg$? Expliquen la situación.

   SOLUCIÓN:

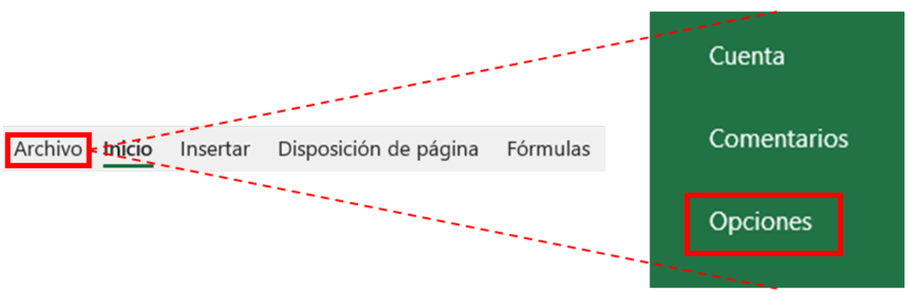
   El teleférico tendría problemas de capacidad máxima de carga, pues entre los 6 hombres se estaría superando la masa máxima que aguanta el teleférico.

2. En Excel, generen muchas muestras aleatorias como la situación de estudio: masas de 6 hombres, distribuidos normalmente, con media $77{,}3\ kg$ y desviación estándar $12{,}9$.

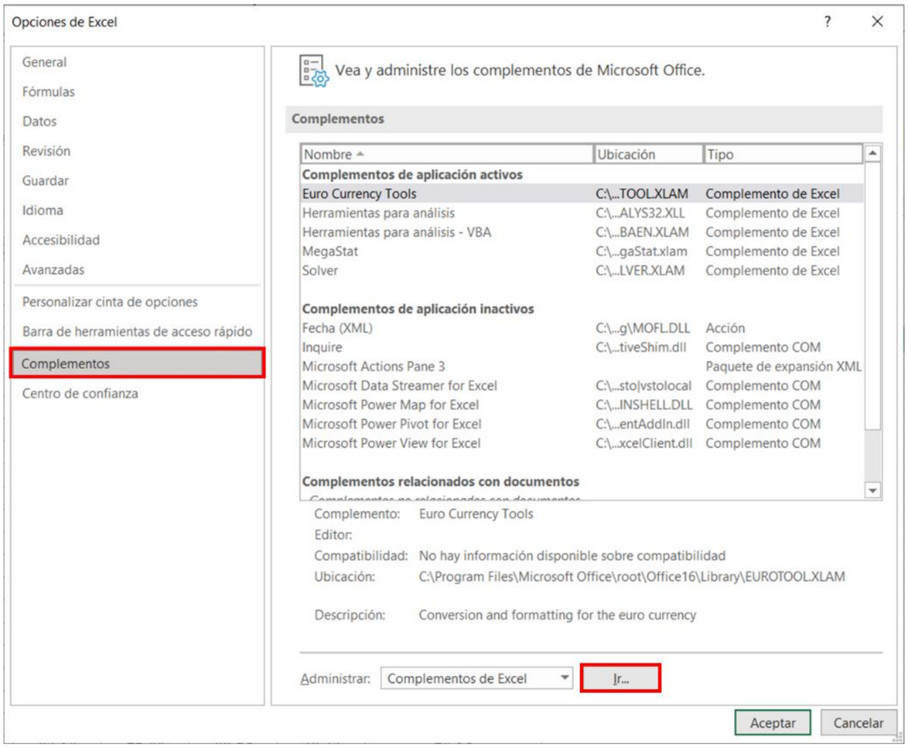
   a. Usen la herramienta *Análisis de datos* en la pestaña *Datos*. En la ventana emergente, seleccionen *Generación de números aleatorios*.

   SOLUCIÓN:

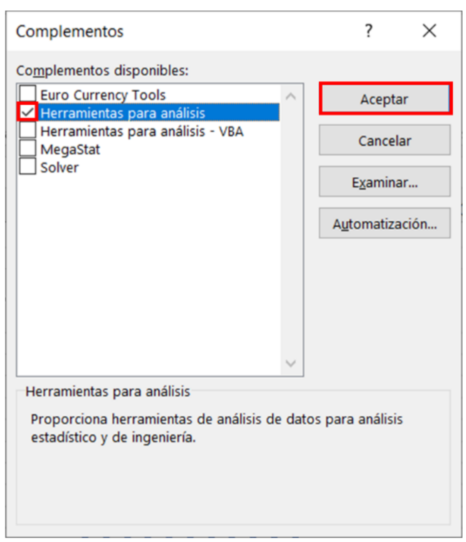
   Para activar la función *Análisis de datos* en Excel, se hace clic en la parte superior, en la pestaña *Archivo ⟶ Opciones* (al final de la parte izquierda de la pantalla).



   Luego, se desplegará una ventana llamada *Opciones de Excel* se hace clic en *Complementos ⟶ Ir…*



Se desplegará una última ventana, donde se debe seleccionar la casilla *Herramientas para análisis ⟶ Aceptar*.



Para usar la función *Análisis de datos*, se hace clic en la pestaña *Datos ⟶ Análisis de datos*.

b. En la nueva ventana, ajusten los valores como en la figura 1. El número de variables corresponde a la cantidad de elementos de la muestra: seis masas.

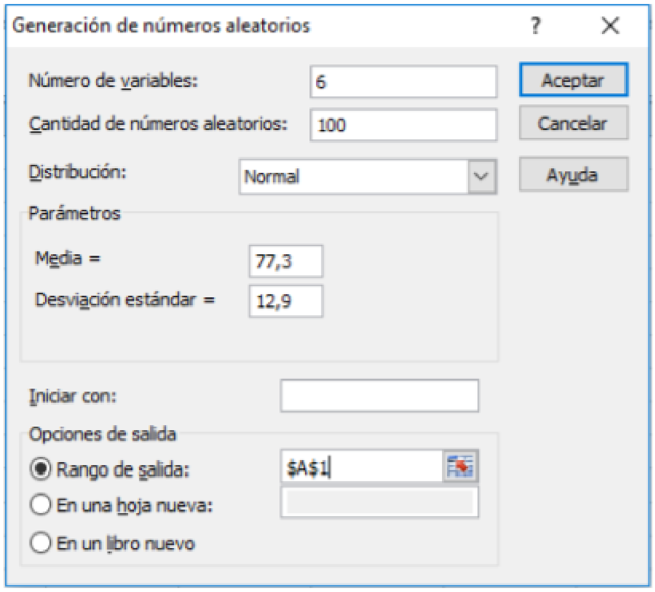
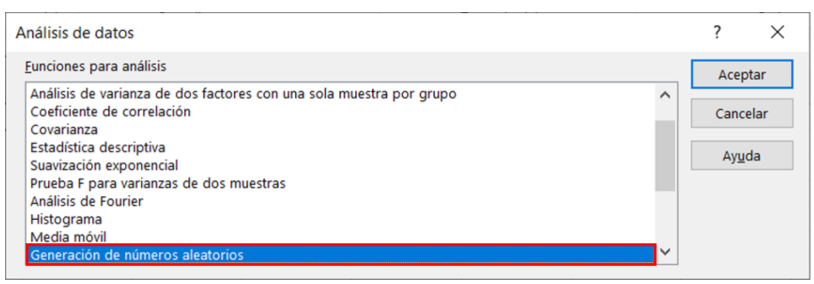


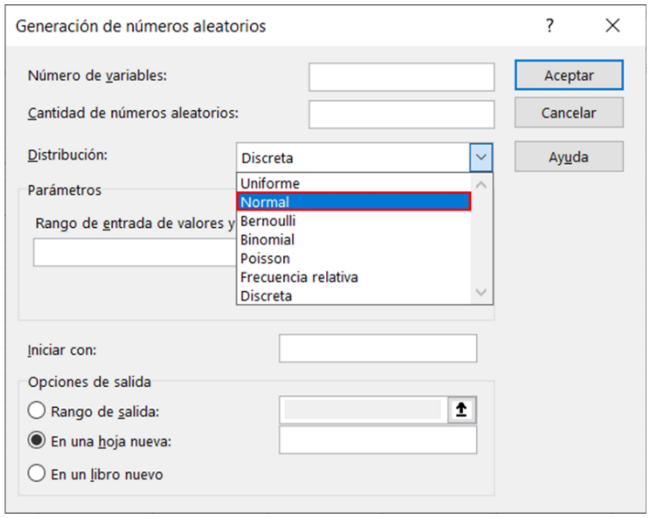
Figura 1: Distribución Normal de datos aleatorios, con media y desviación conocida.

SOLUCIÓN:

Luego de hacer clic en *Análisis de datos* se despliega la siguiente ventana, donde se debe seleccionar la opción *Generación de números aleatorios ⟶ Aceptar*



Después, se desplegará otra ventana, como se muestra a continuación. En la opción *Distribución* debe desplegarse la lista, y seleccionar *Normal*, y completar la ventana con los datos necesarios.



Al seleccionar la opción *Normal*, se despliegan las casillas de los parámetros a completar, como se muestra en la figura 1.

3. ¿Qué necesitan saber para determinar la distribución de las medias de las masas de las muestras de 6 hombres?

SOLUCIÓN:

Se necesita el promedio de cada muestra. Además, por definición, para saber la distribución de las medias muestrales se requiere la esperanza (en este caso $\mu = 77{,}3\ kg$), la desviación estándar (en este caso $\sigma = 12{,}9\ kg$) y la cantidad de individuos en cada muestra (en este caso $n = 6$).

a. Determinen la media de cada muestra, usando la función Promedio en Excel.

SOLUCIÓN:

El resultado de la simulación se encuentra en el anexo 2. Para determinar la media de cada muestra, al función por usar es

$$= PROMEDIO(celdas\ donde\ están\ los\ datos)$$

De esta forma, como son 100 muestras, se tendrán 100 promedios muestrales ($\bar{X}$), y debiese ser más frecuente que la mayoría de los $\bar{X}$ estén en torno a la media poblacional. Así, los valores de los promedios muestrales son:

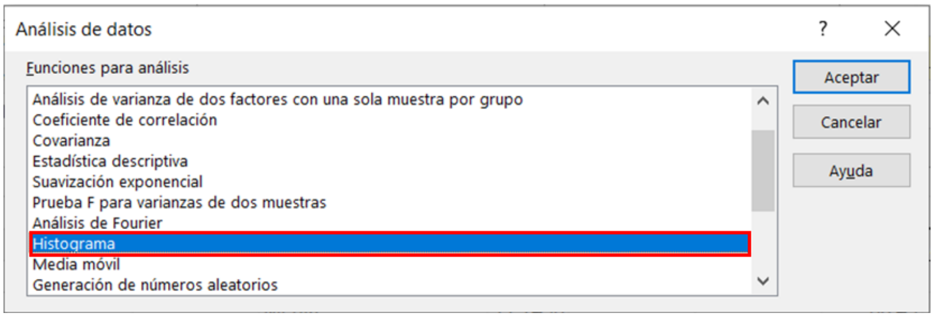
Tabla con 100 promedios muestrales

Nº de muestra	Promedio	Nº de muestra	Promedio	Nº de muestra	Promedio	Nº de muestra	Promedio
1	72,26	**26**	77,99	**51**	82,52	**76**	80,44
2	79,99	**27**	76,67	**52**	67,94	**77**	85,33
3	77,36	**28**	78,11	**53**	82,37	**78**	76,97
4	70,24	**29**	78,59	**54**	79,31	**79**	68,81
5	79,95	**30**	80,28	**55**	74,63	**80**	76,80
6	83,01	**31**	75,90	**56**	73,45	**81**	73,90
7	75,08	**32**	87,26	**57**	77,99	**82**	78,10
8	75,66	**33**	75,55	**58**	76,98	**83**	72,37
9	77,21	**34**	77,69	**59**	69,52	**84**	74,26
10	81,30	**35**	80,02	**60**	70,12	**85**	74,75
11	73,71	**36**	70,44	**61**	85,36	**86**	83,97
12	76,65	**37**	78,28	**62**	88,62	**87**	76,85
13	76,06	**38**	82,04	**63**	77,13	**88**	82,55
14	72,33	**39**	76,52	**64**	70,12	**89**	74,82
15	76,52	**40**	75,38	**65**	77,34	**90**	72,35
16	75,74	**41**	80,20	**66**	81,72	**91**	78,18
17	89,22	**42**	69,28	**67**	77,32	**92**	80,42
18	73,96	**43**	78,70	**68**	79,90	**93**	69,16
19	76,13	**44**	66,18	**69**	79,94	**94**	77,56
20	79,36	**45**	75,30	**70**	75,69	**95**	82,37
21	83,56	**46**	84,02	**71**	86,71	**96**	79,30
22	70,15	**47**	83,82	**72**	77,32	**97**	80,16
23	78,16	**48**	84,89	**73**	77,20	**98**	87,64
24	78,33	**49**	82,93	**74**	78,51	**99**	76,53
25	81,60	**50**	79,28	**75**	69,50	**100**	85,31

b. Con la herramienta *Análisis de datos*, seleccionen *Histograma* y grafiquen los valores de las medias obtenidas en 3.a.

SOLUCIÓN:

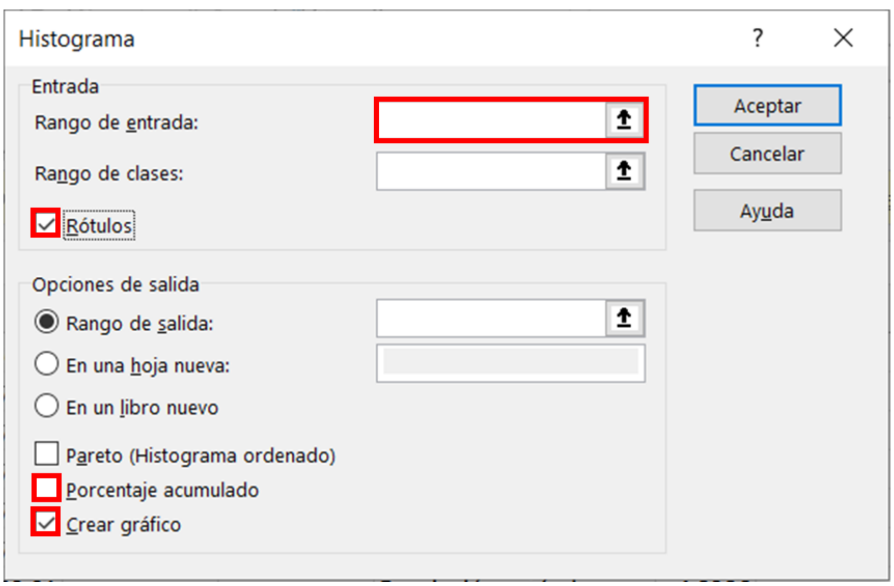
Para obtener el histograma ingresar nuevamente a la opción *Análisis de datos* se despliega la siguiente ventana, donde se debe seleccionar la opción *Histograma ⟶ Aceptar*.



Luego, aparece la siguiente ventana, donde deben seleccionarse las siguientes opciones:

- *Rótulos*, si es que en el *Rango de entrada* se seleccionaron los datos con el título de lo que corresponde.
- *Crear gráfico*, para que, al finalizar, aparezca el gráfico.
- Opcionalmente, se puede seleccionar la casilla *Porcentaje acumulado*, que entrega la frecuencia acumulada, expresada en porcentaje.

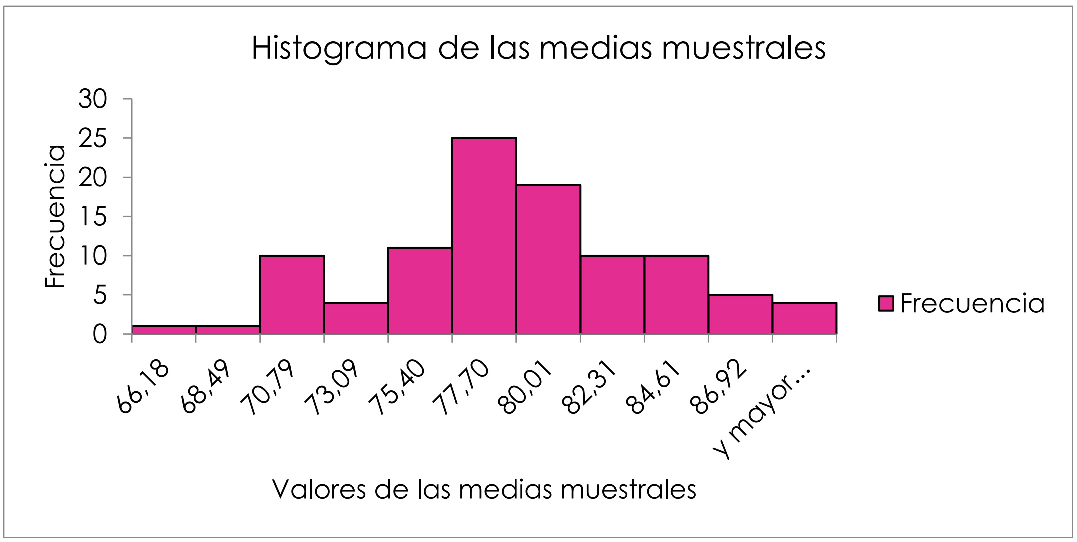
Una vez se esté listo, se presiona *Aceptar*.



El gráfico que corresponde a los valores expresado en 3.a., se muestra a continuación. Además, el programa entrega una tabla resumen de los valores, que también se adjunta.

Tabla de frecuencia de los 100 promedios muestrales

Clase	*Frecuencia*	*% acumulado*
66,18	1	1,00%
68,49	1	2,00%
70,79	10	12,00%
73,09	4	16,00%
75,40	11	27,00%
77,70	25	52,00%
80,01	19	71,00%
82,31	10	81,00%
84,61	10	91,00%
86,92	5	96,00%
y mayor...	4	100,00%



c. Describan la distribución de las medias de las muestras.

SOLUCIÓN:

La distribución de las medias muestrales tiene una forma medianamente acampanada, que deja pocos valores en los extremos, y la mayoría de los promedios se encuentra en el centro de la distribución, cercana al valor de la media poblacional, que es $77{,}3\ kg$.

4. ¿En qué les aporta saber que las medias de las muestras se distribuyen normalmente para entender la situación de estudio? Argumenten su respuesta.

SOLUCIÓN:

Su aporte radica en que la forma de campana es directa en comparación con aquellos casos donde las muestras provengan de poblaciones no normales, pues la convergencia a la forma de campana requiere de más repeticiones.

5. Determinen la media de las medias de las muestras, usando la herramienta Promedio.

SOLUCIÓN:

Utilizando la misma fórmula que en 3.a., se tiene que la media de los promedios muestrales es $77{,}75\ kg$.

a. Comparen las medias de todas las masas y la media obtenida de las medias de las muestras.

SOLUCIÓN:

La comparación se realizará entre la media poblacional, que es $\mu = 77{,}3\ kg$, y la media de los promedios muestrales es $77{,}75\ kg$. Es posible apreciar que los valores difieren en $0{,}45\ kg$, por lo que se puede concluir que los valores son prácticamente iguales. Esto corresponde a que, en teoría, la media de la población, y la media de los promedios muestrales para ambos casos es la misma, que es lo que estipula la teoría, $\mu_{\bar{X}} = E(\bar{X}) = \mu_X$.

Nota al docente: Cuando el enunciado hace referencia a "las medias de todas las masas" se considera la media poblacional μ.

b. Determinen la desviación estándar de las medias de las muestras, usando la función Desvest en Excel.

SOLUCIÓN:

Para determinar la desviación estándar de las medias muestrales, utilizaremos la función

$$= DESVEST.P(celdas\ donde\ están\ los\ datos)$$

Así, la desviación estándar de los promedios muestrales tiene un valor aproximado de $4{,}81\ kg$.

c. Comparen la desviación estándar de todas las masas y la desviación estándar obtenida de las medias de las muestras.

SOLUCIÓN:

La comparación se realizará entre la desviación estándar poblacional, que es $\sigma_X = 12{,}9\ kg$, y la desviación estándar de los promedios muestrales es $4{,}81\ kg$. Es posible apreciar que los valores difieren en un gran valor $8{,}09 kg$, sin embargo, esto sucede porque no se puede hacer el mismo análisis que se realizó a la media. La teoría estipula que la desviación estándar de los promedios muestrales corresponderá a la desviación estándar poblacional σ_X dividido en $\sqrt{n}$, donde n es la cantidad de individuos de cada muestra. Si calculamos $\sigma_{\bar{X}} = \frac{\sigma_X}{n} = \frac{12{,}9}{\sqrt{6}} = 5{,}27\ kg$. Es posible apreciar que los valores difieren en $0{,}45\ kg$, por lo que se puede concluir que los valores son prácticamente iguales.

Nota al docente: Cuando el enunciado hace referencia a "la desviación estándar de todas las masas" se considera la desviación estándar poblacional σ_X. Cabe destacar que este ejercicio es un caso particular ya que la diferencia entre la media poblacional μ_X y la media de los promedios muestrales $\mu_{\bar{X}}$ y la desviación estándar poblacional σ_X y la desviación de los promedios muestrales $\sigma_{\bar{X}}$ no siempre coincide.

6. Prueben generando otras muestras que cumplan con lo enunciado en el punto 2 (media y desviación estándar dadas).

   a. Verifiquen si obtienen los mismos resultados que en los puntos 3 y 5.

SOLUCIÓN:

- Simulación 1:

  La tabla con los valores de cada muestra se adjunta al final, en el anexo 3. Los promedios muestrales se muestran en la siguiente tabla:

Tabla con 100 promedios muestrales

Nº de muestra	Promedio	Nº de muestra	Promedio	Nº de muestra	Promedio	Nº de muestra	Promedio
1	78,33	**26**	74,00	**51**	75,43	**76**	78,57
2	74,89	**27**	81,70	**52**	77,98	**77**	77,17
3	79,84	**28**	74,95	**53**	80,76	**78**	82,02
4	77,99	**29**	71,67	**54**	87,55	**79**	74,02
5	76,54	**30**	73,93	**55**	68,56	**80**	79,63

6	75,30	**31**	73,17	**56**	76,83	**81**	68,45
7	71,09	**32**	74,71	**57**	66,14	**82**	78,14
8	75,93	**33**	77,18	**58**	71,66	**83**	81,19
9	78,89	**34**	81,25	**59**	79,80	**84**	72,72
10	73,82	**35**	74,94	**60**	78,19	**85**	66,56
11	77,79	**36**	75,16	**61**	82,82	**86**	81,38
12	79,54	**37**	73,39	**62**	78,14	**87**	84,66
13	75,95	**38**	80,79	**63**	77,89	**88**	76,79
14	70,99	**39**	74,09	**64**	85,81	**89**	74,76
15	85,17	**40**	78,48	**65**	76,65	**90**	75,11
16	83,29	**41**	86,41	**66**	73,52	**91**	71,74
17	77,60	**42**	83,77	**67**	75,22	**92**	73,94
18	75,09	**43**	75,19	**68**	69,84	**93**	79,06
19	71,45	**44**	67,48	**69**	70,92	**94**	84,37
20	73,43	**45**	70,46	**70**	80,75	**95**	73,35
21	79,93	**46**	73,22	**71**	66,09	**96**	78,39
22	85,46	**47**	72,81	**72**	75,91	**97**	82,62
23	81,03	**48**	72,05	**73**	73,61	**98**	73,55
24	76,82	**49**	81,59	**74**	82,01	**99**	81,60
25	71,37	**50**	75,50	**75**	80,27	**100**	82,14

Al calcular la media de estos promedios se obtiene $\mu = 76{,}70\ kg$, y una desviación estándar de $\sigma = 4{,}69\ kg$.

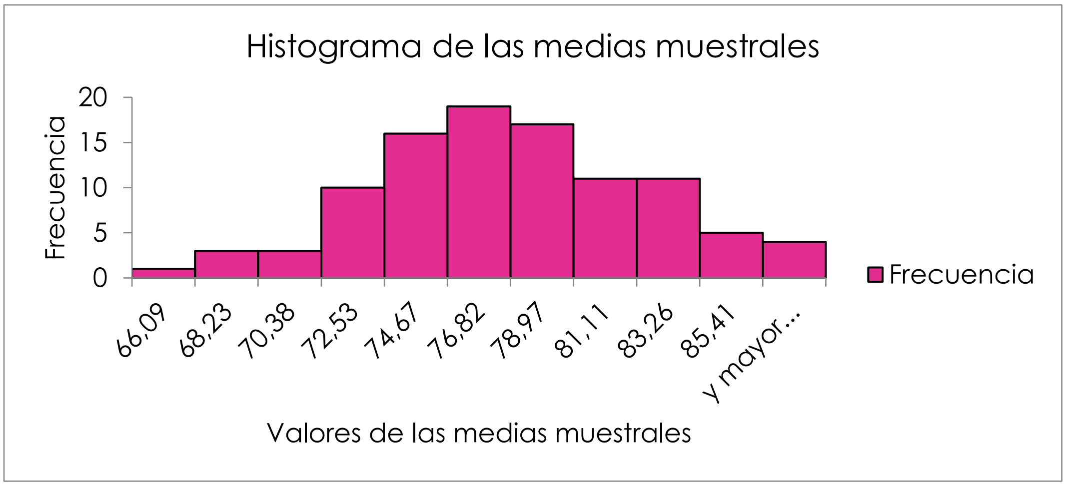
Como es posible ver, los valores no son iguales, pero si difieren en una cantidad cercana a cero respecto a los obtenidos en el punto 3. y 5.

Respecto al gráfico y tabla se tiene:

Tabla de frecuencia de los 100 promedios muestrales

Clase	*Frecuencia*	*% acumulado*
66,1	1	1,01%
68,5	4	5,05%
70,9	3	8,08%
73,2	13	21,21%
75,6	25	46,46%
78,0	15	61,62%
80,4	14	75,76%
82,8	14	89,90%

85,2	5	94,95%
y mayor...	5	100,00%



Se puede ver que el histograma de las medias muestrales tiene una forma acampanada.

- Simulación 2:

La tabla con los valores de cada muestra se adjunta al final, en el anexo 4. Los promedios muestrales se adjuntan en la siguiente tabla:

Tabla con 100 promedios muestrales

Nº de muestra	Promedio	Nº de muestra	Promedio	Nº de muestra	Promedio	Nº de muestra	Promedio
1	79,67	**26**	72,84	**51**	76,75	**76**	83,59
2	79,84	**27**	80,91	**52**	77,93	**77**	92,13
3	72,88	**28**	72,49	**53**	80,68	**78**	76,12
4	80,50	**29**	67,65	**54**	68,87	**79**	79,12
5	80,26	**30**	79,15	**55**	72,29	**80**	80,48
6	81,73	**31**	79,91	**56**	52,30	**81**	72,01
7	89,73	**32**	67,80	**57**	70,24	**82**	77,39
8	77,70	**33**	84,35	**58**	89,41	**83**	71,28
9	85,04	**34**	70,89	**59**	83,52	**84**	75,99
10	77,63	**35**	78,21	**60**	73,81	**85**	83,29
11	79,34	**36**	74,17	**61**	81,54	**86**	69,57
12	76,95	**37**	79,06	**62**	84,61	**87**	73,35

13	75,34	**38**	70,20	**63**	76,34	**88**	82,34
14	77,46	**39**	80,76	**64**	88,02	**89**	77,58
15	71,08	**40**	63,45	**65**	77,66	**90**	91,76
16	79,60	**41**	83,17	**66**	73,45	**91**	77,71
17	73,91	**42**	71,42	**67**	82,24	**92**	74,34
18	73,24	**43**	68,46	**68**	78,35	**93**	79,75
19	67,15	**44**	74,45	**69**	77,28	**94**	87,85
20	84,02	**45**	86,46	**70**	69,72	**95**	86,90
21	71,43	**46**	83,68	**71**	85,16	**96**	76,23
22	84,64	**47**	75,16	**72**	77,15	**97**	74,70
23	72,95	**48**	73,14	**73**	79,38	**98**	83,35
24	65,41	**49**	80,40	**74**	69,70	**99**	88,36
25	82,03	**50**	74,70	**75**	81,12	**100**	65,57

Al calcular la media de estos promedios se obtiene $\mu = 77{,}41\ kg$, y una desviación estándar de $\sigma = 6{,}57\ kg$.

Como es posible ver, los valores no son iguales, la media difiere en una cantidad mínima respecto a la calculada en el punto 3 y 5. Sin embargo, para la desviación estándar la diferencia es de casi dos unidades, por lo que se puede mencionar que en esta muestra los datos están más dispersos respecto a la desviación estándar teórica $\sigma_{\bar{X}} = \frac{\sigma_X}{\sqrt{n}}$.

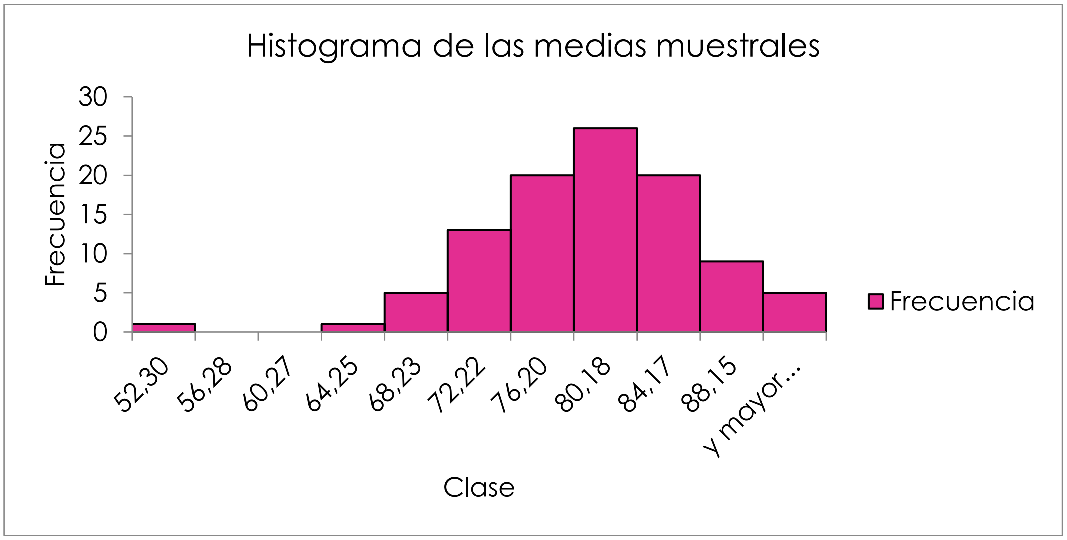
Nota al docente: Se sugiere que, para las simulaciones realizadas se solicite a los estudiantes comparar los valores obtenidos, para que se den cuenta que los valores observados en la práctica pueden diferir en una cantidad significativa de los teóricos.

Respecto al gráfico y tabla se tiene:

Tabla de frecuencia de los 100 promedios muestrales

Clase	*Frecuencia*	*% acumulado*
66,1	1	1,01%
68,5	4	5,05%
70,9	3	8,08%
73,2	13	21,21%
75,6	25	46,46%
78,0	15	61,62%

80,4	14	75,76%
82,8	14	89,90%
85,2	5	94,95%
y mayor...	5	100,00%



En esta ocasión, se puede ver que el histograma de las medias muestrales posee una forma acampanada.

b. Generalicen a partir de sus resultados y los de sus compañeros.

SOLUCIÓN:

La generalización debe ser que:

$$\bar{X} \xrightarrow{d} N\left(\mu_X; \frac{\sigma_X^2}{n}\right)$$

O en su defecto que

$$\frac{\bar{X} - \mu_X}{\sigma_X} \cdot \sqrt{n} \xrightarrow{d} N(0,1)$$

Cuando este proviene de cualquier población, no necesariamente Normal. En palabras, se puede mencionar que la media de los promedios muestrales debe ser igual a la media poblacional, por otro lado, la desviación estándar de los promedios muestrales es el cociente entre la desviación estándar poblacional y $\sqrt{n}$, con n como la cantidad de individuos de cada muestra generada.

7. ¿Cuál es la probabilidad de que, al seleccionar 6 hombres al azar, estos tengan una masa de más de $80\ kg$ en promedio? Argumenten su respuesta.

SOLUCIÓN:

Sea:

X_i: *Peso del hombre i, en kg* con $i = 1,2,3,4,5,6$

Luego, se tiene que

$\mu_X = E(X_i) = 77{,}3\ kg$

$\sigma_X = \sqrt{Var(X_i)} = 12{,}9\ kg$

Ahora, se pide $P(\bar{X} > 80)$, así

$P(\bar{X} > 80) = 1 - P(\bar{X} \leq 80)$

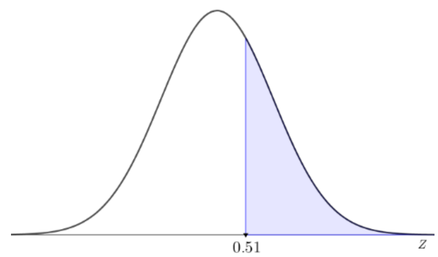
$= 1 - P\left(\frac{\bar{X}-77{,}3}{12{,}9} \cdot \sqrt{6} \leq \frac{80-77{,}3}{12{,}9} \cdot \sqrt{6}\right)$

$= 1 - P(Z \leq 0{,}51)$

$= 1 - \phi(0{,}51)$

$= 1 - 0{,}6950$

$= 0{,}305$



z	0	0.01	0.02
0.4	0.6554	0.6591	0.6628
0.5	0.6915	0.6950	0.6985
0.6	0.7257	0.7291	0.7324

Por lo tanto, la probabilidad de que al seleccionar 6 hombres al azar, estos tengan una masa promedio de más de $80\ kg$ es de un $30{,}5\%$

a. ¿Cómo pueden usar las conclusiones del teorema del límite central para determinar probabilidades? Argumenten su respuesta.

SOLUCIÓN:

Se espera que los estudiantes identifiquen que el Teorema Central del Límite señala que, la distribución de la media aritmética de una variable aleatoria tiende a una Distribución Normal cuando la cantidad de variables es muy grande, en otras palabras, garantiza una Distribución Normal cuando n es suficientemente grande, a partir de $n > 30$. En consecuencia, facilita el proceso para determinar la probabilidad que sea pedida, ya que la estandarización contribuye a ello.

b. ¿Qué parámetros se debe considerar para obtener la puntuación z de la media de las masas de una muestra? Argumenten su respuesta.

SOLUCIÓN:

Los parámetros son $\mu_X = E(X_i) = 77{,}3\ kg$, $\sigma_X = \sqrt{Var(X_i)} = 12{,}9\ kg$ y $n = 6$, ya que la forma de estandarizar es $Z = \frac{X-\mu_X}{\sigma_X} \cdot \sqrt{n}$.

c. Según la probabilidad encontrada, ¿qué les parece la seguridad del teleférico? Argumenten su respuesta.

SOLUCIÓN:

La seguridad del teleférico es apropiada, ya que es más probable que los seis hombres pesen en promedio menos de $80\,kg$, lo que corresponde a casi un 70%. Además, se debe considerar que las masas de las personas que se pueden subir a un teleférico se compensan ellos.

DESCRIPCIÓN DE LA SUBACTIVIDAD:

- Esta subactividad permite que los estudiantes comprendan que el Teorema Central del Límite también se cumple cuando la población de origen no distribuye Normal.
- Aplicar el Teorema Central del Límite a partir de un estudio de caso que distribuye Binomial utilizando la herramienta digital Excel.

DESARROLLO

8. Se desea realizar una prueba con la intención de medir las habilidades matemáticas adquiridas por los estudiantes hasta segundo medio a 200 establecimientos educacionales de la Región Metropolitana. Esta evaluación consta de 40 preguntas que tienen 3 alternativas y solo una de ellas es correcta. Se han escogido al azar 30 personas de los colegios para rendir dicha evaluación. Suponga independencia entre los estudiantes que rinden la prueba.

a. Define la variable aleatoria X que se está midiendo e identifica su distribución y parámetros respectivos.

SOLUCIÓN:

La variable aleatoria es

$$X\text{: }\mathit{Número\ de\ preguntas\ respondidas\ correctamente\ de\ un\ total\ de\ }40.$$

Por la forma que tiene X se puede afirmar que posee una Distribución Binomial de parámetros $n = 40$ y $p = \frac{1}{3}$.

b. Con los parámetros encontrados en a. calcula la esperanza y desviación estándar de la población.

SOLUCIÓN:

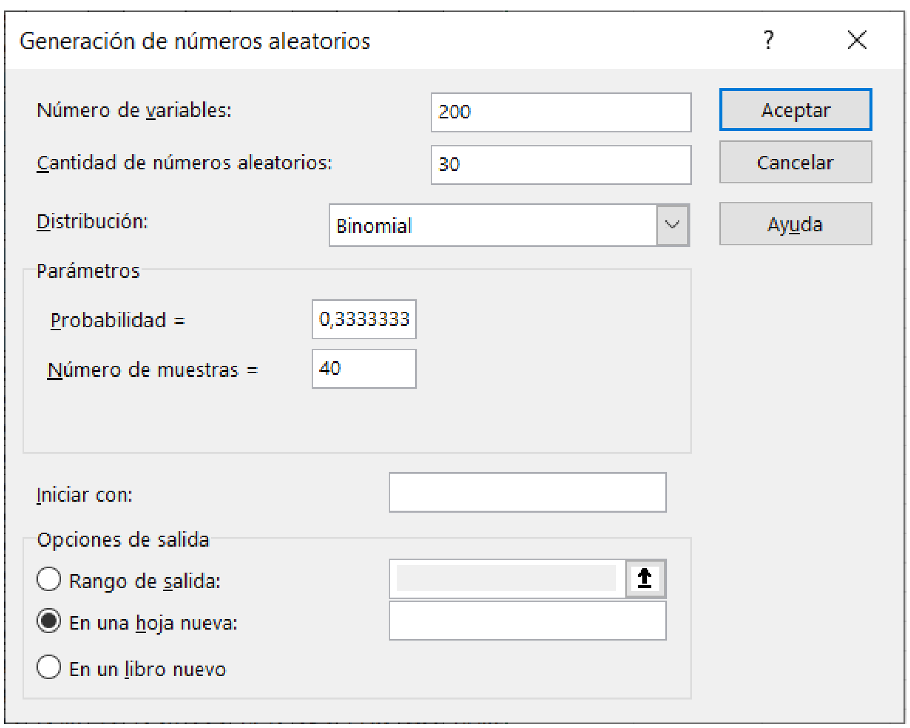
La esperanza es

$$\mu = n \cdot p = 40 \cdot \frac{1}{3} = \frac{40}{3} = 13{,}3333 \; respuestas \; correctas$$

Por otra parte, la desviación estándar es

$$\sigma = \sqrt{n \cdot p \cdot q} = \sqrt{40 \cdot \frac{1}{3} \cdot \frac{2}{3}} = 2{,}9814 \; respuestas \; correctas$$

c. Ahora, en Excel, generen 200 muestras de 30 individuos cada una con la herramienta Análisis de datos ⟶ Generación de números aleatorios y completen los datos como se muestra en la figura adjunta considerando para la sección Parámetros la información contestada en a.



d. Calcula los promedios de cada muestra con la función *PROMEDIO* de Excel.

SOLUCIÓN:

Tabla con 200 promedios muestrales

Nº	Promedio	Nº	Promedio	Nº	Promedio	Nº	Promedio	Nº	Promedio
1	13,67	**41**	13,00	**81**	13,20	**121**	14,27	**161**	13,63
2	13,13	**42**	13,13	**82**	14,10	**122**	13,17	**162**	13,17
3	13,00	**43**	13,17	**83**	14,33	**123**	12,97	**163**	13,40
4	13,60	**44**	13,80	**84**	13,97	**124**	13,97	**164**	13,83
5	13,83	**45**	14,00	**85**	12,57	**125**	12,67	**165**	12,97
6	13,30	**46**	13,37	**86**	13,50	**126**	13,90	**166**	12,87
7	13,03	**47**	13,17	**87**	12,83	**127**	14,50	**167**	13,33
8	12,83	**48**	12,50	**88**	13,33	**128**	12,40	**168**	13,20
9	12,77	**49**	13,53	**89**	12,80	**129**	12,73	**169**	13,30
10	13,60	**50**	13,73	**90**	13,47	**130**	14,30	**170**	14,30
11	13,03	**51**	13,23	**91**	13,90	**131**	13,27	**171**	13,00
12	12,77	**52**	13,17	**92**	13,73	**132**	13,60	**172**	14,37
13	13,03	**53**	13,07	**93**	14,03	**133**	13,10	**173**	12,90
14	13,67	**54**	12,67	**94**	12,43	**134**	13,00	**174**	13,73
15	12,83	**55**	13,40	**95**	13,80	**135**	12,47	**175**	12,50
16	13,77	**56**	13,07	**96**	13,57	**136**	13,50	**176**	13,93
17	13,23	**57**	13,13	**97**	14,03	**137**	12,33	**177**	13,40
18	13,30	**58**	12,77	**98**	14,10	**138**	13,30	**178**	13,00
19	13,60	**59**	13,00	**99**	12,97	**139**	12,43	**179**	14,20
20	13,20	**60**	13,00	**100**	12,73	**140**	12,83	**180**	13,70
21	13,23	**61**	13,47	**101**	12,93	**141**	12,87	**181**	13,67
22	13,57	**62**	13,20	**102**	13,07	**142**	13,17	**182**	13,67
23	12,87	**63**	12,13	**103**	13,47	**143**	13,10	**183**	13,13
24	13,37	**64**	13,63	**104**	12,60	**144**	12,77	**184**	13,03
25	13,40	**65**	12,83	**105**	13,33	**145**	12,23	**185**	12,50
26	12,27	**66**	13,83	**106**	13,00	**146**	12,97	**186**	14,17
27	13,47	**67**	13,60	**107**	13,13	**147**	12,90	**187**	12,97
28	13,23	**68**	12,63	**108**	13,10	**148**	13,27	**188**	14,10
29	13,73	**69**	14,70	**109**	13,73	**149**	13,07	**189**	13,47
30	13,73	**70**	13,27	**110**	13,67	**150**	12,97	**190**	13,33
31	12,27	**71**	14,30	**111**	13,47	**151**	14,43	**191**	13,50
32	13,40	**72**	13,30	**112**	14,23	**152**	12,63	**192**	14,40
33	13,37	**73**	13,33	**113**	13,83	**153**	13,93	**193**	12,47
34	13,03	**74**	13,60	**114**	13,10	**154**	13,47	**194**	13,43
35	13,43	**75**	13,13	**115**	13,77	**155**	13,33	**195**	12,87
36	12,93	**76**	13,13	**116**	13,67	**156**	12,80	**196**	13,03
37	14,47	**77**	12,97	**117**	13,87	**157**	13,20	**197**	12,00
38	13,77	**78**	13,47	**118**	14,17	**158**	13,23	**198**	13,60
39	13,23	**79**	14,07	**119**	12,63	**159**	12,70	**199**	12,50
40	12,57	**80**	13,97	**120**	12,93	**160**	14,57	**200**	13,73

e. Realiza el histograma de los promedios muestrales y describe qué forma tiene la gráfica.

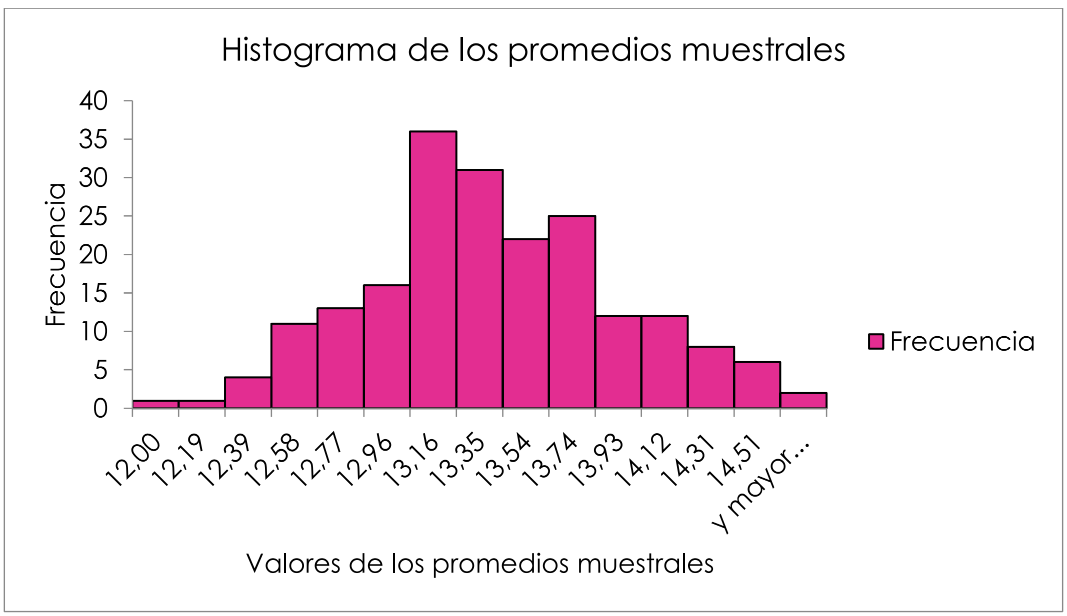
SOLUCIÓN:

La tabla que resume los promedios es:

Tabla de frecuencia de los 200 promedios muestrales

Clase	*Frecuencia*	*% acumulado*
12,00	1	0,50%
12,19	1	1,00%
12,39	4	3,00%
12,58	11	8,50%
12,77	13	15,00%
12,96	16	23,00%
13,16	36	41,00%
13,35	31	56,50%
13,54	22	67,50%
13,74	25	80,00%
13,93	12	86,00%
14,12	12	92,00%
14,31	8	96,00%
14,51	6	99,00%
y mayor...	2	100,00%

Y el histograma es el siguiente:



La forma que posee el histograma concentra la mayoría de los datos entorno a la media, y deja una menor información en los extremos. En otras palabras, tiene una forma acampanada.

f. Obtén la media y la desviación estándar de los promedios muestrales, utilizando la función $PROMEDIO$ y $DESVEST.P$ respectivamente.

SOLUCIÓN:

Los valores son:

$$\mu = 13{,}3135\ respuestas\ correctas$$

$$\sigma = 0{,}5319\ respuestas\ correctas$$

g. Comparen los valores obtenidos en la pregunta f. con los valores teóricos del Teorema Central del Límite. Concluyan si la diferencia entre ambos valores es significativa o no (considere como significativa la diferencia cuando es mayor a $0{,}02$).

SOLUCIÓN:

Los valores teóricos según el TCL son:

$$\mu_{\bar{X}} = \frac{40}{3} = 13{,}3333\ respuestas\ correctas$$

$$\sigma_{\bar{X}} = \frac{\sigma_X}{\sqrt{n}} = \frac{2{,}9814}{\sqrt{30}} = 0{,}5443\ respuestas\ correctas$$

Los valores obtenidos se encuentran en la pregunta f. y corresponden a:

$$\mu = 13{,}3135\ respuestas\ correctas$$

$$\sigma = 0{,}5319\ respuestas\ correctas$$

Por lo tanto, es posible evidenciar que las medias son relativamente iguales, ya que su diferencia absoluta es de $0{,}0198$. Mientras que la diferencia absoluta entre las desviaciones estándar es de $0{,}0124$, es decir, también son relativamente iguales. Por lo que se puede concluir que ambas diferencias no son significativas.

Nota al docente: Se sugiere que utilicen el mismo contexto del problema para plantear ejercicios donde se requiera calcular probabilidades con el TCL.

ORIENTACIONES PARA EL DOCENTE

1. Se espera que los estudiantes determinen la probabilidad de solo un dato para que apliquen la estandarización y encuentren la puntuación Z, usando la tabla de distribución Normal Estándar, que se puede obtener en el Anexo 3.
2. Se sugiere orientar la forma de determinar probabilidades en el caso de la distribución Normal mediante la estandarización, pero si los alumnos no están listos para esta tarea, se recomienda remitirse a las actividades propuestas en el programa de asignatura de Matemática de plan común de 4° medio. En ellas se trabaja de forma exhaustiva en estos cálculos y su interpretación.
3. Se recomienda poner el foco en determinar medias de muestras y luego, en la media de dichas medias. Esto puede ser confuso si no se tiene claridad absoluta de lo que se está intentando hacer. El esquema muestra el orden en que se analiza los datos y qué media se menciona en cada oportunidad.

4. Se argumenta de forma muy simple el teorema del límite central, comprobando que se cumple para casos puntuales. Se eligió en esta oportunidad el caso en que los datos se encuentran distribuidos normalmente; el otro caso queda como un desafío para trabajar en las clases posteriores o se lo menciona, sin profundizar en él. Lo importante es que noten el aporte de contar con este teorema para entender cómo se comportan las medias de muestras, provengan o no de datos distribuidos normalmente.
5. Se propone que usen Excel, porque genera datos aleatorios rápidamente y sus funciones permiten determinar medias y desviaciones estándar.
6. La indicación de observar la gráfica de los valores determinados –sea con las masas o con la media de las muestras de las masas– es que visualicen la distribución de los datos. Ambos casos tienen distribución Normal, pero en las medias de las muestras se nota menos exactitud; por ende, es aún más valioso haber usado valores "reales", pues así es como ocurre en realidad en las investigaciones ajenas al contexto escolar. Se requiere muchos más datos para obtener la campana de Gauss de forma exacta.

7. Para cerrar las actividades, tanto la individual como la colaborativa, se pregunta por la seguridad de las cabinas. Es importante que los argumentos se basen en las probabilidades determinadas, para que vean cómo el hecho de conocer las distribuciones de los datos y el cálculo de probabilidades ayuda para tomar decisiones fundadas y no basadas solamente en la intuición.

8. Si bien en los casos "probabilidad de que un hombre tenga una masa de más de $80\ kg$" y "probabilidad de que 6 hombres tengan una masa total de más de $80\ kg$", las probabilidades no son suficientemente pequeñas como para desestimar un posible exceso de carga en las cabinas, no se debe olvidar que se ha analizado el caso más extremo, en el que solo suben hombres con masas superiores a la media, lo cual no debe ser la regla habitual. Muchas mujeres y niños son usuarios de este atractivo turístico. Por otro lado, las cabinas tienen una estructura muy segura y soportan más de lo expuesto al público.

9. Se sugiere los siguientes indicadores para evaluar formativamente los aprendizajes:
    - Interpretan información estadística que involucra distribuciones de probabilidad Binomial y el Normal.
    - Modelan fenómenos o situaciones cotidianas, científicas y sociales mediante distribuciones Binomiales y Normales.

RECURSOS Y SITIOS WEB

Sitios web sugeridos para estudiantes y profesores

- Generación de datos Normales aleatorios en Excel, alternativa sin Análisis de datos https://www.curriculumnacional.cl/link/https://www.youtube.com/watch?v=d5KaLnAEJIg
- Tabla de probabilidades de distribución Normal Estándar. https://www.curriculumnacional.cl/link/http://www.est.uc3m.es/esp/nueva_docencia/comp_c
ol_leg/ing_tec_inf_gestion/estadistica/Documentacion/Tablas/tablas2caras.pd

ACTIVIDAD 3: APROXIMAR LA DISTRIBUCIÓN BINOMIAL POR LA DISTRIBUCIÓN NORMAL.

PROPÓSITO

Se espera que los estudiantes profundicen aún más en el uso y el aporte de las distribuciones de probabilidad para modelar fenómenos cotidianos. También se pretende que usar la distribución Normal como aproximación de la distribución Binomial, parece un mejor camino según las condiciones del problema. En esta oportunidad, se propone que descubran por qué la técnica que ya han usado es válida y cómo, en términos generales, se debe ajustar un modelo matemático para adecuarse a los requerimientos de lo que pretende modelar.

OBJETIVOS DE APRENDIZAJE

OA 3. Modelar fenómenos o situaciones cotidianas del ámbito científico y del ámbito social, que requieran el cálculo de probabilidades y la aplicación de las distribuciones Binomial y Normal.

- OA b. Resolver problemas que impliquen variar algunos parámetros en el modelo utilizado y observar cómo eso influye en los resultados obtenidos.
- OA e. Construir modelos, realizando conexiones entre variables para predecir posibles escenarios de solución a un problema, y tomar decisiones fundamentadas.
- OA i. Buscar, seleccionar, manejar y producir información matemática/cuantitativa confiable a través de la web.

ACTITUDES

- Pensar con conciencia, reconociendo que los errores ofrecen oportunidades para el aprendizaje.
- Aprovechar las herramientas disponibles para aprender y resolver problemas.

DURACIÓN

12 horas pedagógicas

DESARROLLO

APROXIMANDO UNA DISTRIBUCIÓN BINOMIAL MEDIANTE UNA DISTRIBUCIÓN NORMAL

¿Eres feliz? No siempre es fácil responder esta simple pregunta, pues cada persona tiene su propia idea de lo que es la felicidad; además, nuestras concepciones al respecto van cambiando constantemente, sobre todo a medida que pasan los años. En 2019, *Activa Research* y *WIN* publicó los resultados de una investigación que buscaba averiguar sobre la "Felicidad en Chile y el mundo". Los resultados mostraron que, de los $30\,890$ entrevistados en el mundo, el 52% se considera feliz. En Chile, los resultados indican que, de $1\,032$ encuestados, el 58% se considera feliz. ¿Te parece que estos datos se acercan a lo que ocurre en tu entorno? ¿Qué tan probable será que todos los que conoces estén en ese 58%, o que no estén en ese porcentaje? ¿Qué tan probable será que, si encuestas a algunas personas, ellas sean felices?

1. ¿Qué información te aporta saber que el 58% de los entrevistados es feliz?

   SOLUCIÓN:

   Se puede decir que el 42% de los entrevistados no son felices, y que más de la mitad de la población es feliz. Esta pregunta va dirigida a lo que es la estimación de la probabilidad de éxito p.

   a. ¿Cómo puedes usar esta información para referirte a la población? ¿Qué restricciones debes tener en cuenta?

      SOLUCIÓN:

      El 58% es una estimación del verdadero parámetro p de la población. Como restricciones se tiene que la probabilidad debe estar entre 0 y 1.

   b. ¿Cómo puedes usar esta información para determinar la probabilidad de que, al elegir a una persona al azar y preguntarle si es feliz, su respuesta sea afirmativa?

      SOLUCIÓN:

      Se puede usar como una estimación de que una persona elegida al azar de la población conteste si es feliz o no, considerando que hay mayor probabilidad en que la persona encuestada diga que sí.

2. Se entrevista a 5 personas independientes entre sí y se les pregunta si son felices o no. Se sabe de antemano que el 58% de las personas es feliz y asumimos que el 42% no lo es.

   a. ¿Cuál es la variable aleatoria en este caso? ¿Qué valores podría tomar?

      SOLUCIÓN:

      En este caso, la variable aleatoria es

      $$X\text{: } número\ de\ personas\ felices, en\ un\ total\ de\ 5\ encuestados.$$

      Los valores que se podrían tomar van desde el 0 hasta el 5. Recordar que $Rec(X) = \{0,1,2, \dots, n\}$ si $X \sim Bin(n, p)$.

   b. ¿Se puede modelar esta situación con una distribución Binomial? ¿En qué aportaría conocer la distribución de los datos para entender el problema? Argumenta.

      SOLUCIÓN:

      Sí, se puede modelar con una Distribución Binomial, ya que las condiciones se cumplen, en otras palabras:

      - La independencia entre sucesos, ya que se menciona que las personas son independientes entre sí.
      - La repetición de un fenómeno dicotómico (si es feliz o no).
      - La probabilidad de éxito es constante para cada persona.

      Aportaría en poder asociar una probabilidad al número de personas que sean felices, siempre dentro del recorrido de la variable aleatoria.

3. Determina las probabilidades de los valores de la variable aleatoria si se entrevista a 5 personas.

   SOLUCIÓN:

   Las probabilidades se calculan mediante la función de probabilidad de la Distribución Binomial. A partir del contexto, se entrevista a 5 personas, por lo tanto $n = 5$, además el 58% de las personas es feliz, por lo que $p = 0{,}58$. De esta forma, se puede armar la siguiente función, considerando la variable aleatoria definida en a.:

$$P(X=x)=\begin{cases}\binom{5}{x}\cdot(0{,}58)^x\cdot(0{,}42)^{5-x} & si\ x=0{,}1{,}2{,}3{,}4{,}5\\ 0 & en\ otro\ caso\end{cases}$$

Luego, se pide $P(X=0)$, $P(X=1)$, $P(X=2)$, $P(X=3)$, $P(X=4)$, y $P(X=5)$, así:

$$P(X=0)=\binom{5}{0}\cdot(0{,}58)^0\cdot(0{,}42)^{5-0}=1\cdot(0{,}58)^0\cdot(0{,}42)^5=0{,}0131$$

$$P(X=1)=\binom{5}{1}\cdot(0{,}58)^1\cdot(0{,}42)^{5-1}=5\cdot(0{,}58)^1\cdot(0{,}42)^4=0{,}0902$$

$$P(X=2)=\binom{5}{2}\cdot(0{,}58)^2\cdot(0{,}42)^{5-2}=10\cdot(0{,}58)^2\cdot(0{,}42)^3=0{,}2492$$

$$P(X=3)=\binom{5}{3}\cdot(0{,}58)^3\cdot(0{,}42)^{5-3}=10\cdot(0{,}58)^3\cdot(0{,}42)^2=0{,}3442$$

$$P(X=4)=\binom{5}{4}\cdot(0{,}58)^4\cdot(0{,}42)^{5-4}=5\cdot(0{,}58)^4\cdot(0{,}42)^1=0{,}2376$$

$$P(X=5)=\binom{5}{5}\cdot(0{,}58)^5\cdot(0{,}42)^{5-5}=1\cdot(0{,}58)^5\cdot(0{,}42)^0=0{,}0656$$

a. Grafica los resultados en la Figura 1.

SOLUCIÓN:

Con las probabilidades calculadas en el apartado anterior y considerando que

$$X: número\ de\ personas\ felices, en\ un\ total\ de\ 5\ encuestados$$

De esta forma, se pueden graficar las probabilidades utilizando GeoGebra:

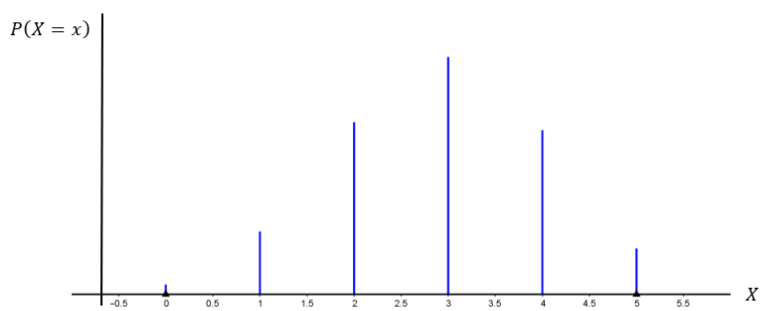


Figura. 1: Distribución de probabilidad de X cuando $n=5$

Nota al docente: Se espera que los estudiantes mantengan la escala en el eje de las probabilidades $P(X=x)$ (eje de las ordenadas) para que puedan visualizar de mejor manera la figura obtenida.

b. Describe la distribución de probabilidad de los datos.

SOLUCIÓN:

La distribución de probabilidad tiene probabilidad máxima en el valor $X = 3$, y aparentemente posee un leve sesgo hacia la derecha.

c. Describe tu apreciación sobre lo laborioso que es determinar probabilidades usando la forma algebraica de la distribución Binomial. Proyéctalo al trabajo que implica determinar probabilidades cuando la variable aleatoria sea muy grande.

SOLUCIÓN:

Cuando los valores del recorrido son pequeños, es fácil calcular las probabilidades de forma algebraica, sin embargo, cuando el n crece, y los valores son de la forma menor (o mayor) que o sus equivalencias, implica mayor tiempo de trabajo y el proceso se vuelve tedioso.

4. Usando GeoGebra, determina las distribuciones de probabilidad de la variable aleatoria cuando $n = 5$, $n = 10$, $n = 50$, $n = 250$, $n = 500$, $n = 1000$.

a. Puedes configurar las opciones en GeoGebra como muestra la figura 2.

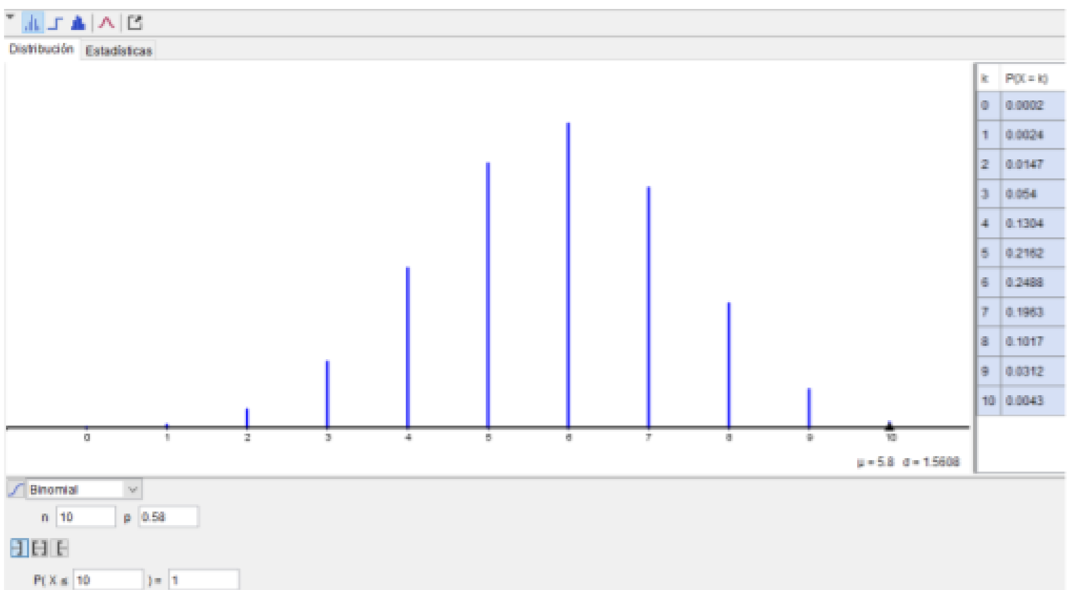


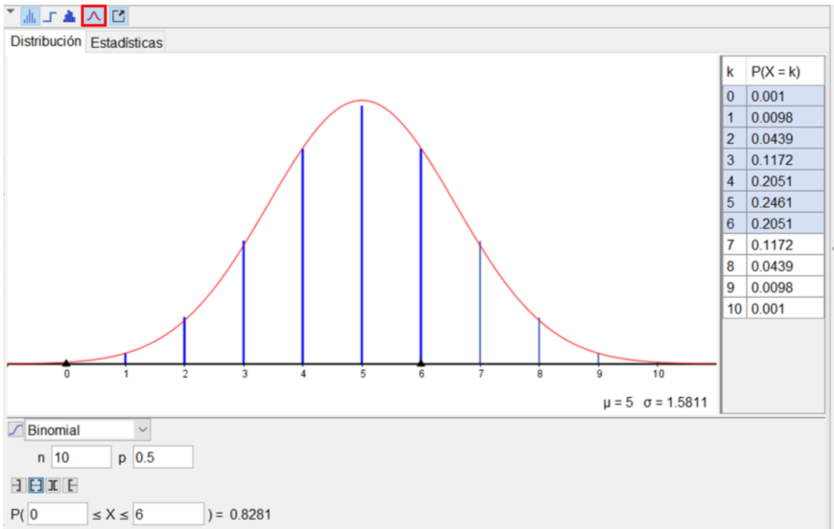
Fig. 2: Configuración en GeoGebra para determinar la distribución de probabilidad Binomial.

Nota al docente: En la página 28 se muestra cómo ingresar a la vista Cálculos de probabilidad en el software GeoGebra e ingresar los valores para el gráfico de barras, dirigirse a la Actividad 1 ⟶ Paso 2 ⟶ apartado b.

b. Selecciona en GeoGebra la opción "Superposición de curva Normal".

SOLUCIÓN:

Para superponer la "curva Normal", se presiona el botón que parece en la parte superior, que tiene la campana de Gauss, como se muestra en la siguiente figura. Este acción ubicará una campana de color rojo en la ventana donde está el gráfico de barras de la Distribución Binomial.

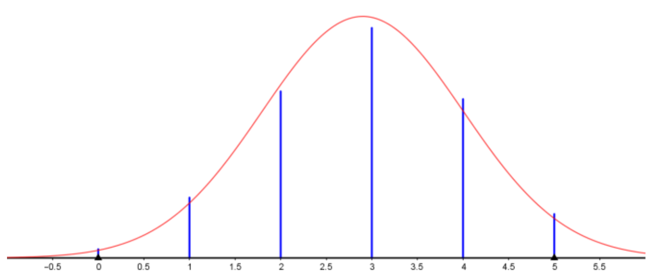


c. Bosqueja las gráficas en la figura 3 con diferentes n, según corresponda.

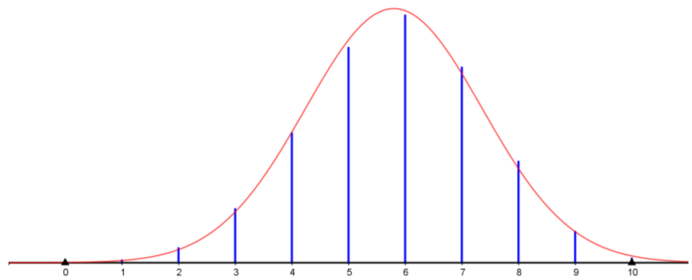
SOLUCIÓN:

Este ejercicio pide bosquejar las distribuciones de probabilidad de:

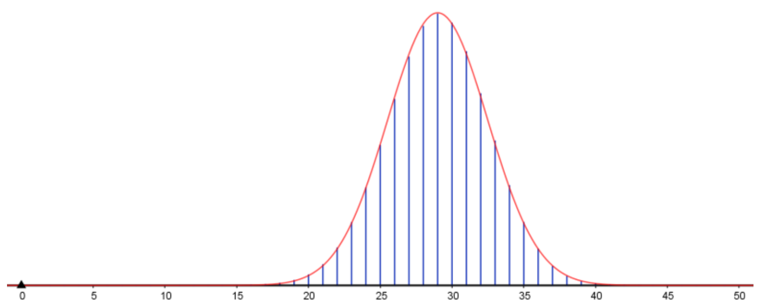
- $n = 5$, obteniendo el siguiente gráfico



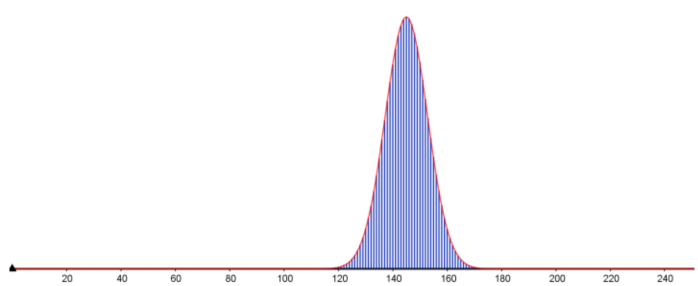
- $n = 10$, obteniendo el siguiente gráfico



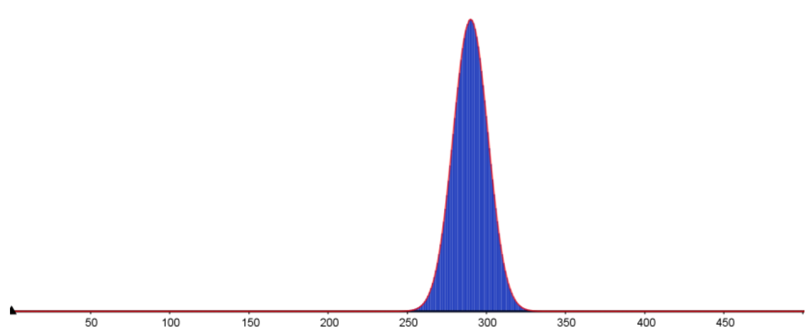
- $n = 50$, obteniendo el siguiente gráfico



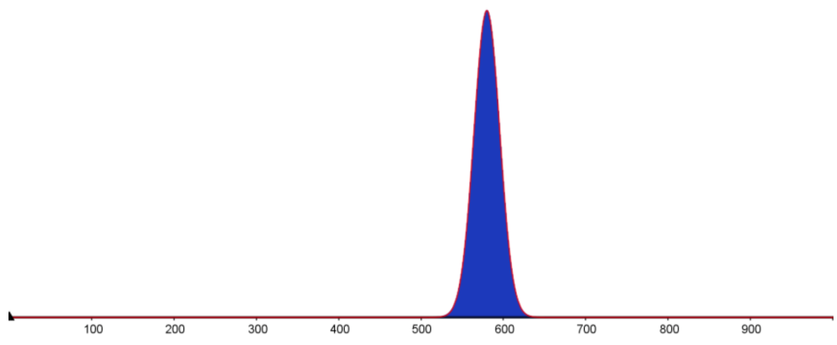
- $n = 250$, obteniendo el siguiente gráfico



- $n = 500$, obteniendo el siguiente gráfico



- $n = 1000$ obteniendo el siguiente gráfico



5. ¿Cómo varían las formas gráficas de la distribución de probabilidad de la variable aleatoria a medida que n crece?

   a. ¿A qué forma se acercan las gráficas a medida que n crece?

   SOLUCIÓN:

   A medida que n crece, las gráficas adquieren forma cada vez más simétrica, y acampanada.

   b. Indica a qué otra distribución de probabilidades se asemeja la distribución Binomial cuando n es muy grande.

   SOLUCIÓN:

   Se asemeja cada vez más a una Distribución Normal.

   c. Conjetura cómo podrías usar este hallazgo para determinar probabilidades y el aporte que puede dar esta aproximación.

   SOLUCIÓN:

   Las posibles conjeturas que el estudiante puede considerar son:

   - Que es imposible aproximar una a la otra, ya que la Distribución Binomial es discreta, y la Distribución Normal es continua.
   - La facilidad de los cálculos, ya que el trabajo con la tabla Z de la Distribución Normal, es inmediata en comparación al cálculo de probabilidades con la función de la Binomial.

   Nota al docente: Se sugiere realizar la institucionalización correcta del contenido para aclarar las conjeturas erradas como la que señala en el primer punto

CÁLCULO DE PROBABILIDADES MEDIANTE LA APROXIMACIÓN NORMAL DE UNA DISTRIBUCIÓN BINOMIAL

1. ¿Cuáles son los parámetros que se necesita para determinar probabilidades según cada distribución?

   a. Indica los parámetros usados en la distribución Binomial.

      SOLUCIÓN:

      Para la Binomial, se necesitan los parámetros n y p, donde n corresponde al número de repeticiones del experimento dicotómico, y p corresponde a la probabilidad de éxito.

   b. Indica los parámetros usados en la distribución Normal.

      SOLUCIÓN:

      Para la Normal, se necesitan los parámetros μ y σ^2, que corresponden a la media y varianza o desviación estándar respectivamente.

OBJETIVO PREGUNTAS COMPLEMENTARIAS

- Las preguntas permiten determinar la relación entre los parámetros de la Distribución Binomial y Normal a partir de la visualización de gráficas en el software GeoGebra.

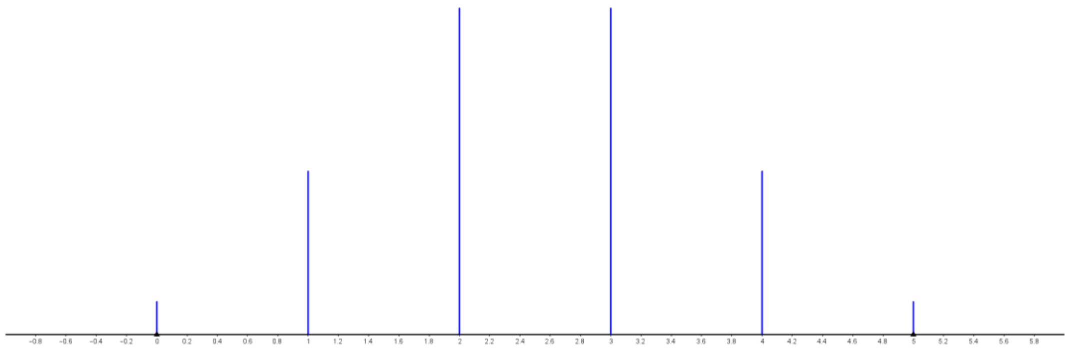
DESARROLLO

c. En GeoGebra, abran la vista cálculos de probabilidad, simulen una gráfica Binomial con $p = 0{,}5$ y los siguientes valores de n.

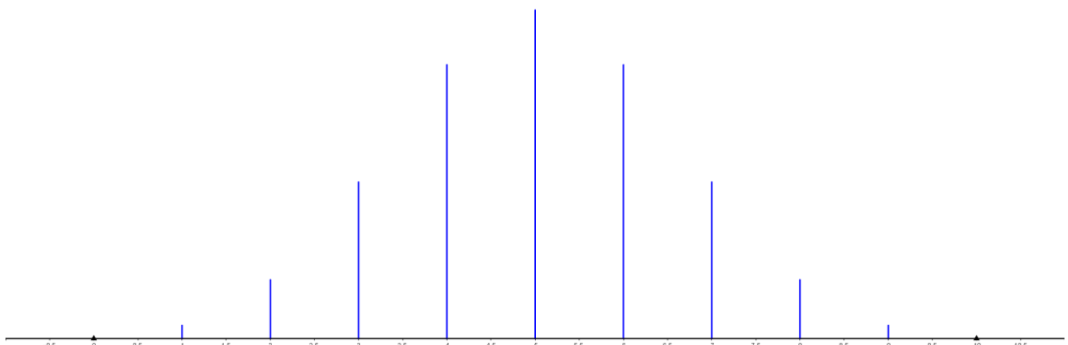
   - Con $n = 5$
   - Con $n = 10$
   - Con $n = 20$
   - Con $n = 30$

   SOLUCIÓN:

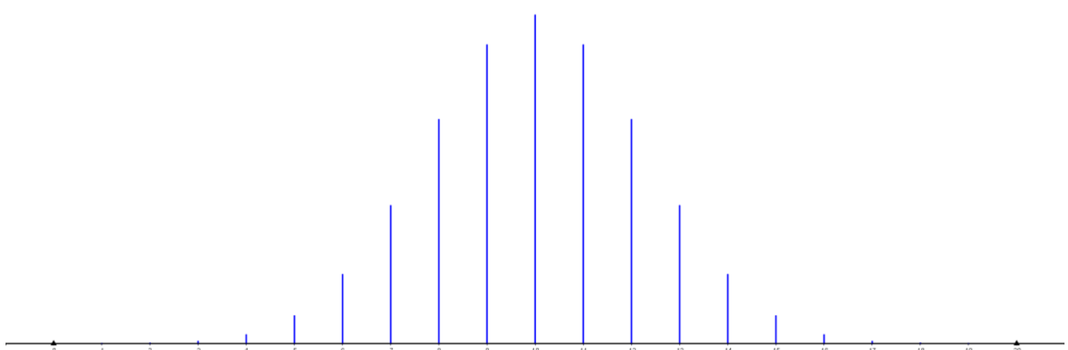
   - Con $n = 5$



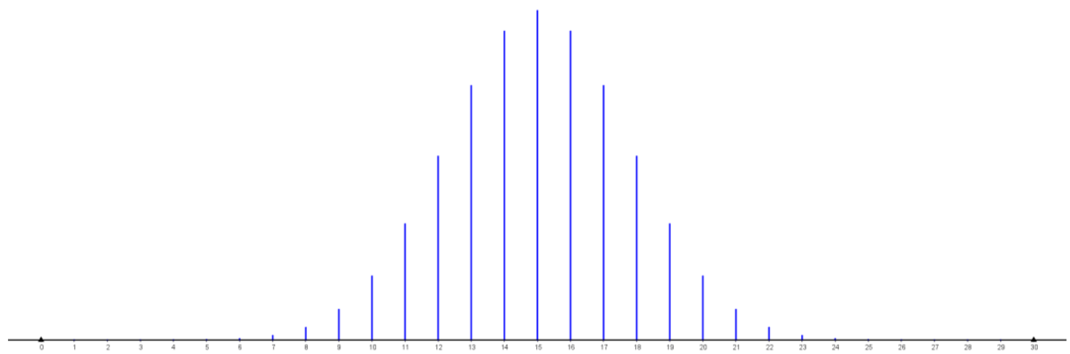
- Con $n = 10$



- Con $n = 20$



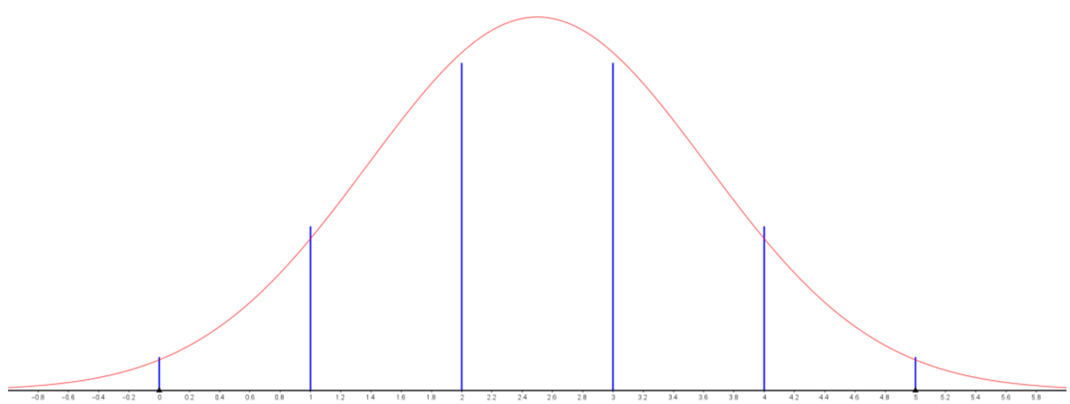
- Con $n = 30$



d. Luego seleccionen la opción superposición de curva Normal y estimen donde se encuentra el centro de la campana superpuesta.

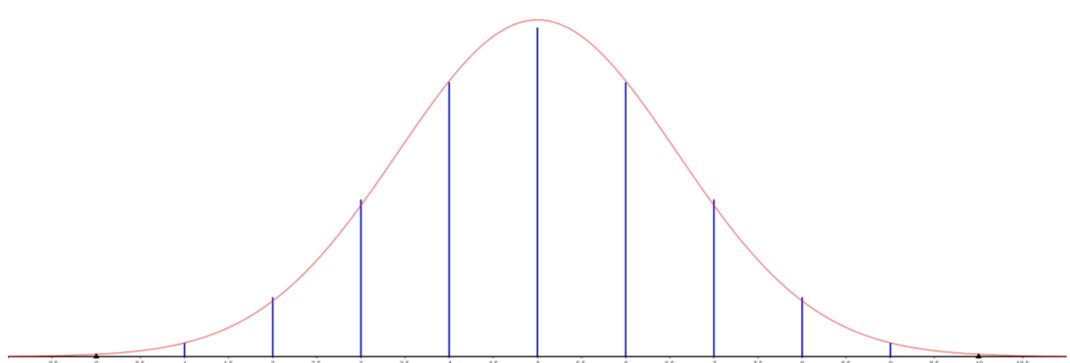
SOLUCIÓN:

- Con $n = 5$



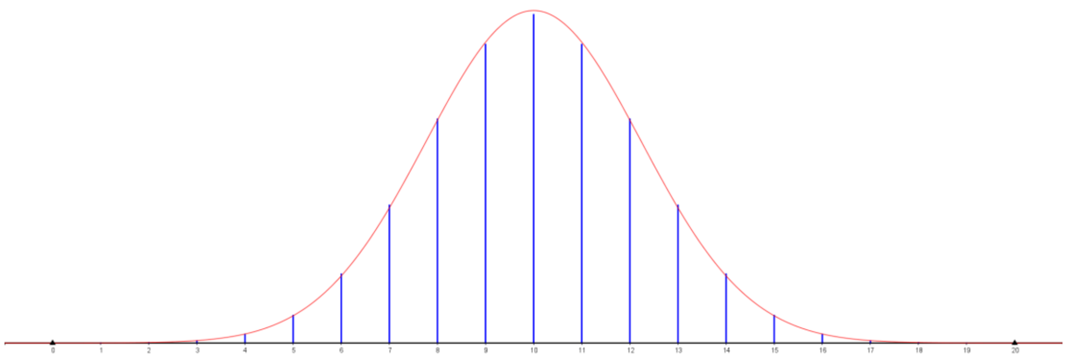
El centro de la campana superpuesta está entre 2 y 3.

- Con $n = 10$



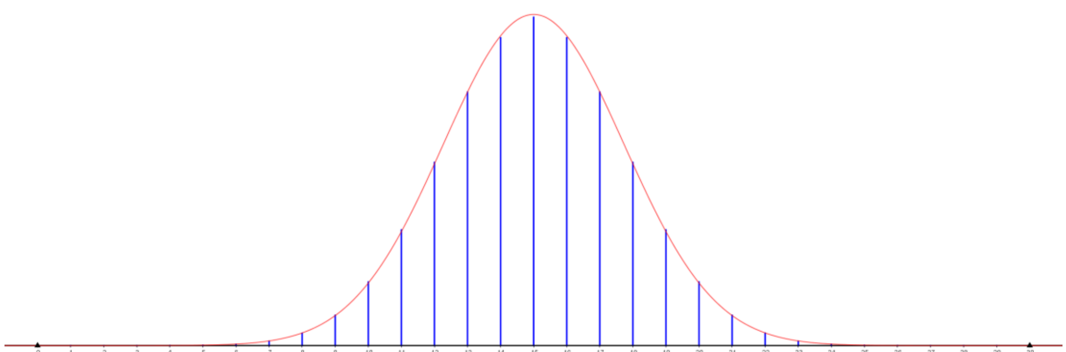
El centro de la campana superpuesta está cercano al 5.

- Con $n = 20$



El centro de la campana está cercano al 10.

- Con $n = 30$



El centro de la campana está cercano al 15.

e. Ahora, calcula la esperanza y desviación estándar de la Distribución Binomial y compárenlos con los parámetros dados de la Distribución Normal.

- Con $n = 5$, los parámetros de la Normal son $\mu = \frac{5}{2}$ y $\sigma = \frac{\sqrt{5}}{2}$.
- Con $n = 10$, los parámetros de la Normal son $\mu = 5$ y $\sigma = \frac{\sqrt{10}}{2}$.
- Con $n = 20$, los parámetros de la Normal son $\mu = 10$ y $\sigma = \sqrt{5}$.
- Con $n = 30$, los parámetros de la Normal son $\mu = 15$ y $\sigma = \frac{\sqrt{30}}{2}$.

SOLUCIÓN:

- Con $n = 5$, los parámetros de la Normal son $\mu = \frac{5}{2}$ y $\sigma = \frac{\sqrt{5}}{2}$.

  Como $p = \frac{1}{2}$ se tiene que:

$$\mu = n \cdot p = 5 \cdot \frac{1}{2} = \frac{5}{2} = 2{,}5$$

$$\sigma = \sqrt{n \cdot p \cdot q} = \sqrt{5 \cdot \frac{1}{2} \cdot \left(1 - \frac{1}{2}\right)} = \frac{\sqrt{5}}{2} = 1{,}118$$

  Por lo tanto, los parámetros de la Distribución Binomial coindicen con los parámetros de la Distribución Normal.

- Con $n = 10$, los parámetros de la Normal son $\mu = 5$ y $\sigma = \frac{\sqrt{10}}{2}$.

  Como $p = \frac{1}{2}$ se tiene que:

$$\mu = n \cdot p = 10 \cdot \frac{1}{2} = 5$$

$$\sigma = \sqrt{n \cdot p \cdot q} = \sqrt{10 \cdot \frac{1}{2} \cdot \left(1 - \frac{1}{2}\right)} = \frac{\sqrt{10}}{2} = 1{,}581$$

  Por lo tanto, los parámetros de la Distribución Binomial coindicen con los parámetros de la Distribución Normal.

- Con $n = 20$, los parámetros de la Normal son $\mu = 10$ y $\sigma = \sqrt{5}$.

  Como $p = \frac{1}{2}$ se tiene que:

$$\mu = n \cdot p = 20 \cdot \frac{1}{2} = 10$$

$$\sigma = \sqrt{n \cdot p \cdot q} = \sqrt{20 \cdot \frac{1}{2} \cdot \left(1 - \frac{1}{2}\right)} = \sqrt{5} = 2{,}236$$

  Por lo tanto, los parámetros de la Distribución Binomial coindicen con los parámetros de la Distribución Normal.

- Con $n = 30$, los parámetros de la Normal son $\mu = 15$ y $\sigma = \frac{\sqrt{30}}{2}$.

  Como $p = \frac{1}{2}$ se tiene que:

$$\mu = n \cdot p = 30 \cdot \frac{1}{2} = 15$$

$$\sigma = \sqrt{n \cdot p \cdot q} = \sqrt{30 \cdot \frac{1}{2} \cdot \left(1 - \frac{1}{2}\right)} = \frac{\sqrt{30}}{2} = 2{,}739$$

Por lo tanto, los parámetros de la Distribución Binomial coindicen con los parámetros de la Distribución Normal.

f. ¿Cómo se puede determinar los parámetros requeridos en la distribución Normal, usando los parámetros de la distribución Binomial?

SOLUCIÓN:

A partir de las preguntas anteriores, se espera que los estudiantes señalen que los parámetros de la Distribución Normal, en función de los de la Binomial son:

$$\mu = n \cdot p$$

$$\sigma^2 = n \cdot p \cdot q$$

2. Usa la aproximación normal de la Binomial para determinar la probabilidad de que, al elegir al azar 50 personas, 25 respondan como máximo que son felices. Registra los cálculos.

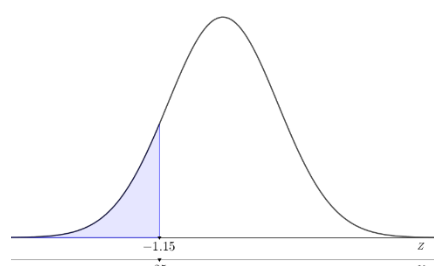
SOLUCIÓN:

Sea:

$$Y\text{: } número\ de\ personas\ que\ son\ felices, de\ un\ total\ de\ 50.$$

Se tiene que $Y \sim Bin(n = 50; p = 0{,}58)$, por lo tanto, los parámetros de la Normal son:

$\mu = 50 \cdot 0{,}58 = 29$ y $\sigma = \sqrt{50 \cdot 0{,}58 \cdot 0{,}42} = 3{,}49$



luego se tiene que:

$P(Y \leq 25) = P\left(\frac{Y-29}{3{,}49} \leq \frac{25-29}{3{,}49}\right)$

$\approx P(Z \leq -1{,}15)$

$= \phi(-1{,}15)$

$= 0{,}1251$

z	0.04	0.05	0.06
-1.2	0.1075	0.1056	0.1038
-1.1	0.1271	0.1251	0.1230
-1.0	0.1492	0.1469	0.1446

En consecuencia, la probabilidad que como máximo 25 de las 50 personas respondan que son felices es de un 12,51%.

Si calculamos la probabilidad de la Binomial con GeoGebra, se obtiene aproximadamente $0{,}158$

a. Si hubieras determinado esta probabilidad mediante la distribución Binomial, ¿qué dificultades habrías tenido?

SOLUCIÓN:

El cálculo del número combinatorio dificulta el proceso que utiliza la Distribución Binomial para determinar la probabilidad, pues no cuenta con una expresión compacta para la función de distribución acumulada (que calcula probabilidades de la forma menor que, como máximo, entre otras equivalencia) por lo que sumar 26 probabilidades (desde el 0 al 25), requiere de mayor trabajo.

b. ¿En qué te aporta determinar la probabilidad de $P(X \leq k)$ mediante la distribución Normal y no mediante la distribución Binomial? Argumenta.

SOLUCIÓN:

La rapidez en obtener la probabilidad pedida, ya que, una vez obtenidos los parámetros, solo basta estandarizar y observar la tabla. Sin embargo, siempre se debe tener en consideración que k es cualquier valor del recorrido de la Distribución Binomial. En general el aporte es el cálculo de las probabilidades de la forma menor o igual o sus equivalentes.

3. Determina la probabilidad de que, al elegir al azar a 100 personas, exactamente 75 de ellas respondan que son felices. Argumenta.

a. Sin hacer los cálculos, describe cómo los harías usando la distribución Binomial.

SOLUCIÓN:

Se espera que un estudiante mencione que utilizará la función de probabilidad de la Distribución Binomial, identificando la variable aleatoria, y sus parámetros, luego define la f.d.p. y finalmente realiza los cálculos correspondientes.

La descripción anterior matemáticamente corresponde al siguiente proceso:

Sea:

X: *número de personas que responden que son felices, de un total de* 100 *encuestados*

Luego, se tiene que $X \sim Bin(n = 100, p = 0{,}58)$, por lo que, la f.d.p. corresponde a:

$$P(X=x)=\begin{cases}\binom{100}{x}\cdot(0{,}58)^x\cdot(0{,}42)^{100-x} & si\ x=0,1,2,\dots,100\\ 0 & en\ otro\ caso\end{cases}$$

Como se pide $P(X=75)$ se tiene que:

$$P(X=75)=\binom{100}{75}\cdot(0.58)^{75}\cdot(0{,}42)^{100-75}=\binom{100}{75}\cdot(0.58)^{75}\cdot(0{,}42)^{25}$$

b. Conjetura una forma de calcular la probabilidad buscada, usando la distribución Normal.

SOLUCIÓN:

Las posibles conjeturas que un estudiante puede considerar son:

- Realizar todas las cuentas, descubrir que $P(X=75)$ es 0, y quedarse con ese resultado.
- Buscar el área de un rectángulo (que corresponde a base por altura), y de esa forma se puede aproximar la Binomial a la Normal, trabajando en un intervalo conveniente, para que la base sea 1 y la altura $P(X=x)$.

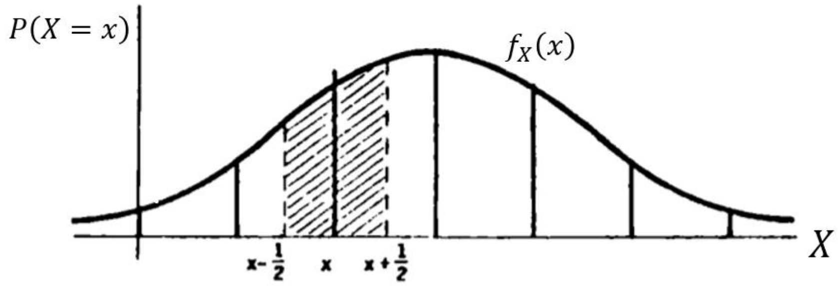
Nota al docente: Para la primera conjetura se sugiere preguntar al estudiante: ¿y si son 99?, ¿o 5?, ¿y con 10?, luego de escuchar sus respuestas, utilice las siguientes devoluciones ¿y en todas nos da 0?, ¿de qué forma se puede hacer que no sea cero? Esta última pregunta induce a la conjetura del segundo punto.

c. ¿Cómo puedes hacer que la variable aleatoria discreta que estás considerando sea "ajustada" para ser continua? ¿Cómo podrías crear intervalos adecuados para ello?

SOLUCIÓN:

En primer lugar, para ajustar la variable aleatoria discreta a la continua se debe considerar un intervalo conveniente cuya área es la probabilidad buscada. Por otra parte, la forma de crear un intervalo adecuado para el valor x debe ser un rectángulo de base 1, por lo que, se considerará una distancia de $0{,}5$ a cada lado del valor buscado.

Nota al docente: Se sugiere apoyarse en la siguiente gráfica para la compresión del contenido.



d. Elabora una regla para hacer una corrección de continuidad y compárala con las ideas de tus compañeros. Luego, en consenso, lleguen a la mejor corrección posible.

SOLUCIÓN:

La regla más apropiada, considerando todo lo anteriormente expuesto es la siguiente:

$$P(X = a) \approx P(a - 0{,}5 < X < a + 0{,}5)$$

e. Vuelve a determinar la probabilidad pedida en el punto 2, ahora aplicando la corrección por continuidad, y compara ambos resultados.

SOLUCIÓN:

Sea:

$$X\text{: número de personas que son felices, de un total de } 50.$$

Se tiene que $X \sim N(\mu = 29; \sigma^2 = 3{,}49^2)$, por lo tanto, lo pedido es $P(X \leq 25)$, si se realiza la corrección por continuidad, se tiene que:

$P(X \leq 25) \approx P(X \leq 25 + 0{,}5)$

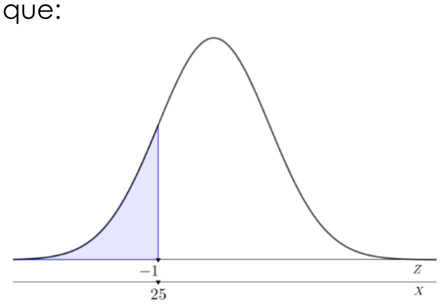
$= P(X \leq 25{,}5)$

$\approx P\left(Z \leq \frac{25{,}5 - 29}{3{,}49}\right)$

$= P(Z \leq -1{,}00)$

$= \phi(-1{,}00)$

$= 0{,}1587$



z	0	0.01	0.02
-1.1	0.1357	0.1335	0.1314
-1.0	0.1587	0.1562	0.1539
-0.9	0.1841	0.1814	0.1788

En consecuencia, la probabilidad que 25 de las 50 personas respondan como máximo que son felices es de un 15,87%

El resultado obtenido en el punto 2 es $0{,}1251$, que corresponde al valor sin la corrección, por el contrario, el valor con la corrección es $0{,}1587$. La diferencia es de $0{,}0336$, lo que no es un porcentaje importante de pérdida de información, no obstante, el valor que más se acerca al calculado con GeoGebra $(0{,}158)$, es el que utiliza la corrección por continuidad.

4. ¿En qué te basarías para decidir desde qué valor de n vale la pena utilizar la aproximación Normal de la distribución Binomial? Argumenta.

SOLUCIÓN:

En que el valor de n complique el cálculo del número combinatorio, ya sea de forma mental, manualmente o con calculadora.

a. Completa la tabla 1 con la probabilidad de distintos valores de la variable aleatoria y para un n dado.

SOLUCIÓN:

Para el desarrollo de esta tabla en la parte de la Distribución Normal se considera la corrección por continuidad que hace trabajar en el intervalo $[X-0{,}5; X+0{,}5]$

Tabla 1: Comparación de probabilidades según cada distribución de probabilidad, variando n.

Distribución de Probabilidad		
$\boldsymbol{n}$	$Bin(n,p)$	$N(\mu,\sigma^2)$
5	$P(X=3)=\binom{5}{3}\cdot(0{,}58)^3\cdot(0{,}42)^{5-3}$ $=10\cdot(0{,}58)^3\cdot(0{,}42)^2$ $=0{,}3442$	$P(X=3)\approx P(3-0{,}5<X<3+0{,}5)$ $=P(2{,}5<X<3{,}5)$ $\approx P\left(\frac{2{,}5-5\cdot 0{,}58}{\sqrt{5\cdot 0{,}58\cdot 0{,}42}}<Z<\frac{3{,}5-5\cdot 0{,}58}{\sqrt{5\cdot 0{,}58\cdot 0{,}42}}\right)$ $=P(-0{,}36<Z<0{,}54)$ $=P(Z<0{,}54)-P(Z<-0{,}36)$ $=\phi(0{,}54)-\phi(-0{,}36)$ $=0{,}7054-0{,}3594$ (**) $=0{,}346$
10	$P(X=5)=\binom{10}{5}\cdot(0{,}58)^5\cdot(0{,}42)^{10-5}$ $=252\cdot(0{,}58)^5\cdot(0{,}42)^5$	$P(X=5)\approx P(5-0{,}5<X<5+0{,}5)$ $=P(4{,}5<X<5{,}5)$

	$= 0{,}2162$	$\approx P\left(\frac{4{,}5-10\cdot 0{,}58}{\sqrt{10\cdot 0{,}58\cdot 0{,}42}} < Z < \frac{5{,}5-10\cdot 0{,}58}{\sqrt{10\cdot 0{,}58\cdot 0{,}42}}\right)$ $= P(-0{,}83 < Z < -0{,}19)$ $= P(Z < -0{,}19) - P(Z < -0{,}83)$ $= \phi(-0{,}19) - \phi(-0{,}83)$ $= 0{,}4247 - 0{,}2033$ (**) $= 0{,}2214$
50	$P(X = 25) = \binom{50}{25} \cdot (0{,}58)^{25} \cdot (0{,}42)^{50-25}$ $= \binom{50}{25} \cdot (0{,}58)^{25} \cdot (0{,}42)^{25}$ $= 0{,}0587$	$P(X = 25) \approx P(25 - 0{,}5 < X < 25 + 0{,}5)$ $= P(24{,}5 < X < 25{,}5)$ $\approx P\left(\frac{24{,}5-50\cdot 0{,}58}{\sqrt{50\cdot 0{,}58\cdot 0{,}42}} < Z < \frac{25{,}5-50\cdot 0{,}58}{\sqrt{50\cdot 0{,}58\cdot 0{,}42}}\right)$ $= P(-1{,}29 < Z < -1{,}00)$ $= P(Z < -1{,}00) - P(Z < -1{,}29)$ $= \phi(-1{,}00) - \phi(-1{,}29)$ $= 0{,}1587 - 0{,}0985$ (**) $= 0{,}0602$
100	$P(X = 50) = \binom{100}{50} \cdot (0{,}58)^{50} \cdot (0{,}42)^{100-50}$ $= \binom{100}{50} \cdot (0{,}58)^{50} \cdot (0{,}42)^{50}$ $= 0{,}0218$	$P(X = 50) \approx P(50 - 0{,}5 < X < 50 + 0{,}5)$ $= P(49{,}5 < X < 50{,}5)$ $\approx P\left(\frac{49{,}5-100\cdot 0{,}58}{\sqrt{100\cdot 0{,}58\cdot 0{,}42}} < Z < \frac{50{,}5-100\cdot 0{,}58}{\sqrt{100\cdot 0{,}58\cdot 0{,}42}}\right)$ $= P(-1{,}72 < Z < -1{,}52)$ $= P(Z < -1{,}52) - P(Z < -1{,}72)$ $= \phi(-1{,}52) - \phi(-1{,}72)$ $= 0{,}0643 - 0{,}0427$ (**) $= 0{,}0216$
250	$P(X = 125) = \binom{250}{125} \cdot (0{,}58)^{125} \cdot (0{,}42)^{250-125}$ $= \binom{250}{125} \cdot (0{,}58)^{125} \cdot (0{,}42)^{125}$ $= 0{,}002\ (*)$	$P(X = 125) \approx P(125 - 0{,}5 < X < 125 + 0{,}5)$ $= P(124{,}5 < X < 125{,}5)$ $\approx P\left(\frac{124{,}5-250\cdot 0{,}58}{\sqrt{250\cdot 0{,}58\cdot 0{,}42}} < Z < \frac{125{,}5-250\cdot 0{,}58}{\sqrt{250\cdot 0{,}58\cdot 0{,}42}}\right)$ $= P(-2{,}63 < Z < -2{,}50)$ $= P(Z < -2{,}50) - P(Z < -2{,}63)$ $= \phi(-2{,}50) - \phi(-2{,}63)$ $= 0{,}0062 - 0{,}0043$ (**) $= 0{,}0019$
500	$P(X = 250) = \binom{500}{250} \cdot (0{,}58)^{250} \cdot (0{,}42)^{500-250}$ $= \binom{500}{250} \cdot (0{,}58)^{250} \cdot (0{,}42)^{250}$ $= 0{,}001\ (*)$	$P(X = 250) \approx P(250 - 0{,}5 < X < 250 + 0{,}5)$ $= P(249{,}5 < X < 250{,}5)$ $\approx P\left(\frac{249{,}5-500\cdot 0{,}58}{\sqrt{500\cdot 0{,}58\cdot 0{,}42}} < Z < \frac{250{,}5-500\cdot 0{,}58}{\sqrt{500\cdot 0{,}58\cdot 0{,}42}}\right)$ $= P(-3{,}67 < Z < -3{,}58)$

		$= P(Z < -3{,}58) - P(Z < -3{,}67)$ $= \phi(-3{,}58) - \phi(-3{,}67)$ $= 0{,}0002 - 0{,}0001$ (**) $= 0{,}0001$
1000	$P(X = 500) = \binom{1000}{500} \cdot (0{,}58)^{500} \cdot (0{,}42)^{1000-500}$ $= \binom{1000}{500} \cdot (0{,}58)^{500} \cdot (0{,}42)^{500}$ $= 0$ (*)	$P(X = 500) \approx P(500 - 0{,}5 < X < 500 + 0{,}5)$ $= P(499{,}5 < X < 500{,}5)$ $\approx P\left(\frac{499{,}5-1000\cdot 0{,}58}{\sqrt{1000\cdot 0{,}58\cdot 0{,}42}} < Z < \frac{500{,}5-1000\cdot 0{,}58}{\sqrt{1000\cdot 0{,}58\cdot 0{,}42}}\right)$ $= P(-5{,}16 < Z < -5{,}09)$ $= P(Z < -5{,}09) - P(Z < -5{,}16)$ $= \phi(-5{,}09) - \phi(-5{,}16)$ $= 0 - 0$ $= 0$

(*): Valores obtenidos con GeoGebra.

(**): Valores extraídos de la tabla Z (ver Anexo 1).

b. ¿Desde qué valor de n la diferencia entre la probabilidad determinada por $Bin(n, p)$ y $N(\mu, \sigma^2)$ es más pequeña que una centésima? Argumenta.

SOLUCIÓN:

La diferencia comienza a verse desde $n = 5$, ya que

$$|0{,}3442 - 0{,}346| < 0{,}01 \Longrightarrow |-0{,}0018| < 0{,}01 \Longrightarrow 0{,}0018 < 0{,}01$$

c. Para ese valor de n determina $n \cdot p$ y $n \cdot q$:

SOLUCIÓN:

$$n \cdot p \ = \ 5 \cdot 0{,}58 = 2{,}9$$

$$n \cdot (1 - p) = \ 5 \cdot 0{,}42 = 2{,}1$$

d. Prueba con algunos valores distintos de n entre 10 y 50 y analiza desde qué n las probabilidades determinadas por ambas distribuciones son muy cercanas.

SOLUCIÓN:

En este caso, cualquier valor que escoja el estudiante, y que se encuentre entre 10 y 50 servirá para analizar desde que n las probabilidades son muy cercana, ya que por efectos del valor de p, la aproximación se visualiza al tiro. No obstante, se ofrecen algunos valores que podrían probarse para poder instruir la regla, y como valor de X se mantiene la mitad de n, es decir:

Para $n = 20$ se calculará $P(X = 10)$.

Para $n = 30$ se calculará $P(X = 15)$.

Para $n = 40$ se calculará $P(X = 20)$.

Distribución de Probabilidad		
$\boldsymbol{n}$	$Bin(n,p)$	$N(\mu,\sigma^2)$
20	$P(X=10)=\binom{20}{10}\cdot(0{,}58)^{10}\cdot(0{,}42)^{20-10}$ $=184\,756\cdot(0{,}58)^{10}\cdot(0{,}42)^{10}$ $=0{,}1359$	$P(X=10)\approx P(10-0{,}5<X<10+0{,}5)$ $=P(9{,}5<X<10{,}5)$ $\approx P\left(\frac{9{,}5-20\cdot0{,}58}{\sqrt{20\cdot0{,}58\cdot0{,}42}}<Z<\frac{10{,}5-20\cdot0{,}58}{\sqrt{20\cdot0{,}58\cdot0{,}42}}\right)$ $=P(-0{,}95<Z<-0{,}50)$ $=P(Z<-0{,}50)-P(Z<-0{,}95)$ $=\phi(-0{,}50)-\phi(-0{,}95)$ $=0{,}3085-0{,}1711$ (*) $=0{,}1374$
30	$P(X=15)=\binom{30}{15}\cdot(0{,}58)^{15}\cdot(0{,}42)^{30-15}$ $=\binom{30}{15}\cdot(0{,}58)^{15}\cdot(0{,}42)^{15}$ $=0{,}0979$	$P(X=15)\approx P(15-0{,}5<X<15+0{,}5)$ $=P(14{,}5<X<15{,}5)$ $\approx P\left(\frac{14{,}5-30\cdot0{,}58}{\sqrt{30\cdot0{,}58\cdot0{,}42}}<Z<\frac{05{,}5-30\cdot0{,}58}{\sqrt{30\cdot0{,}58\cdot0{,}42}}\right)$ $=P(-1{,}07<Z<-0{,}70)$ $=P(Z<-0{,}70)-P(Z<-1{,}07)$ $=\phi(-0{,}70)-\phi(-1{,}07)$ $=0{,}2420-0{,}1423$ (*) $=0{,}0997$
40	$P(X=20)=\binom{40}{20}\cdot(0{,}58)^{20}\cdot(0{,}42)^{40-20}$ $=\binom{40}{20}\cdot(0{,}58)^{20}\cdot(0{,}42)^{20}$	$P(X=20)\approx P(20-0{,}5<X<20+0{,}5)$ $=P(19{,}5<X<20{,}5)$ $\approx P\left(\frac{19{,}5-40\cdot0{,}58}{\sqrt{40\cdot0{,}58\cdot0{,}42}}<Z<\frac{20{,}5-40\cdot0{,}58}{\sqrt{40\cdot0{,}58\cdot0{,}42}}\right)$

	$= 0{,}0746$	$= P(-1{,}19 < Z < -0{,}86)$ $= P(Z < -0{,}86) - P(Z < -1{,}19)$ $= \phi(-0{,}86) - \phi(-1{,}19)$ $= 0{,}1949 - 0{,}1170$ (*) $= 0{,}0779$

(*): Valores extraídos de la tabla Z (ver anexo 1).

Como se puede ver, a medida que n crece, las probabilidades con la Distribución Binomial y la Distribución Normal se acercan cada vez más.

e. Conjetura una regla práctica para determinar desde qué valor de n conviene usar la aproximación Normal de la Binomial.

SOLUCIÓN:

En rigor, no existe una regla clara para aproximar la Binomial a la Normal, sin embargo, se utilizará la mencionada en el texto del estudiante de formación general, que indica lo siguiente:

Sea Y una variable aleatoria discreta que se distribuye de forma binomial $B(n, p)$ de parámetros n, p y cumple las siguientes condiciones:

$$n \cdot p \geq 5 \quad \text{y} \quad n \cdot q \geq 5$$

Entonces la variable aleatoria Y se puede aproximar a una distribución normal de media np y desviación típica $\sqrt{n \cdot p \cdot q}$, es decir:

$$B(n, p) \rightarrow N(np, \sqrt{n \cdot p \cdot q})$$

De modo que, para poder aproximar, se hace necesario considerar que:

$$n \cdot p \geq 5$$

$$n \cdot q \geq 5$$

5. Vuelve a responder la pregunta: ¿En qué te aporta determinar la probabilidad de $P(X \leq k)$ mediante la distribución Normal y no mediante la distribución Binomial?

SOLUCIÓN:

El proceso llevado a cabo es más eficiente porque al ser una probabilidad acumulada, los valores a partir de una Distribución Normal se calculan en base a la tabla Z, sin embargo, con la Binomial, se debe sumar la probabilidad en cada punto, ya que no cuenta con una expresión para su función de distribución acumulada, dicho de otra manera, para calcular probabilidades de la forma $P(X \leq k)$.

ORIENTACIONES PARA EL DOCENTE

1. Se sugiere vincular datos estadísticos con el cálculo de probabilidades, usando un contexto sencillo pero cercano a los estudiantes. Aunque ya deberían haber hecho esta tarea en niveles anteriores, sirve para reforzar nuevamente cómo, desde la estadística y eligiendo de forma adecuada la muestra (usando métodos probabilísticos), se puede obtener conclusiones acerca de la población, por ahora informalmente.
2. Se comienza preguntando si la situación se puede modelar con la distribución Binomial y qué aportaría hacerlo, a fin de que revisen los conceptos previos que esta actividad requiere y también para que se acostumbren a hacer metacognición.
3. Para la parte *Cálculo de Probabilidades mediante la Aproximación Normal de una Distribución Binomial*, en las preguntas 1.c., 1.d., y 1.e., se recomienda darle otros valores a p, como $0{,}1$ o $0{,}8$, idealmente no próximos a $0{,}5$, para que los estudiantes visualicen que, al variar este valor, la aproximación no es inmediata. Adicionalmente, comprenden que se puede aproximar con cualquier probabilidad de éxito.
4. Usando algún programa, se puede comparar distintas gráficas de la distribución Binomial a medida que n crece considerablemente (de 5 a 1000) y observar cómo los datos van tendiendo a distribuirse Normalmente; la herramienta de curva Normal de GeoGebra ayuda a hacerlo. Deberán concluir que los datos se distribuyen normalmente cuando n es lo suficientemente grande. Si la probabilidad de éxito se aleja de $0{,}5$, se necesita que n sea más grande para que la aproximación sea razonable. Conviene que, usando también GeoGebra, prueben con probabilidades distintas de $0{,}5$ para notarlo.
5. En la primera actividad, se espera que usen la aproximación Normal de la Binomial; por eso, se les pide determinar una probabilidad que, si se hiciera con la distribución Binomial, comportaría una tarea tan laboriosa que sería muy complejo cumplirla sin un recurso digital. Acá importa discutir con los alumnos sobre los aportes de las tecnologías actuales, pues hoy se puede determinar

probabilidades con la distribución Binomial para un n cualquiera, por grande que sea. La idea es mostrar cómo se usaría la aproximación Normal si no se cuenta con dichos recursos digitales y, al mismo tiempo, que los estudiantes sepan cómo se ha ido construyendo la historia de las probabilidades, por qué antes era tan importante poder aproximar distribuciones Binomiales a Normales para determinar probabilidades, y por qué esto ha hecho que la distribución Normal sea tan importante y estudiada con tanto detalle.

6. Las actividades siguientes buscan establecer las reglas necesarias para determinar probabilidades de datos discretos a datos continuos; por ende, se analiza la corrección por continuidad. Se recomienda hacer esta tarea más de una vez, hasta verificar que todos la comprenden. El profesor puede complementar esta parte con gráficos, ver las barras (Binomial) y el histograma (Normal), y discutir sobre los ajustes que se debe hacer.
7. La regla práctica – que se vincula con lo mencionado en el punto 3 sobre el valor de n y la probabilidad cercana o lejana a $0{,}5$ – se debe notar desde una tabla con distintas probabilidades determinadas, usando la distribución Binomial y la Normal. Si usan GeoGebra, no necesitan hacer la corrección de continuidad, pues permite determinar los valores buscados en forma inmediata. Una alternativa es usar GeoGebra solo para los cálculos con la Binomial y calcular la Normal a mano; así aprovechan de estandarizar y aplicar la corrección de continuidad.
8. Finalmente, más allá de todos los cálculos, se espera que los alumnos demuestren el aporte de aproximar la distribución Binomial mediante una distribución Normal, por lo que el docente debe propiciar una discusión al respecto.
9. Se sugiere los siguientes indicadores para evaluar formativamente los aprendizajes:
    - Modelan situaciones que involucran aproximar una distribución Binomial mediante el modelo Normal.
    - Argumentan cuándo se puede modelar una situación o fenómeno con una distribución Binomial o normal.

RECURSOS Y SITIOS WEB

Sitios web sugeridos para estudiantes y profesores

- GeoGebra online https://www.curriculumnacional.cl/link/https://www.geogebra.org/graphing?lang=es
- Tabla de probabilidades de distribución Normal Estándar (solo las 2 primeras hojas) https://www.curriculumnacional.cl/link/http://www.est.uc3m.es/esp/nueva_docencia/comp_col_leg/ing_tec_inf_gestion/estadistica/Documentacion/Tablas/tablas2caras.pdf

ACTIVIDAD DE EVALUACIÓN

Objetivos de Aprendizaje	Indicadores de evaluación
OA 3. Modelar fenómenos o situaciones cotidianas del ámbito científico y del ámbito social, que requieran el cálculo de probabilidades y la aplicación de las distribuciones Binomial y Normal. OA b. Resolver problemas que impliquen variar algunos parámetros en el modelo utilizado y observar cómo eso influye en los resultados obtenidos. OA e. Construir modelos, realizando conexiones entre variables para predecir posibles escenarios de solución a un problema, y tomar decisiones fundamentadas. OA i. Buscar, seleccionar, manejar y producir información matemática/cuantitativa confiable a través de la web.	Identifican las principales características de los modelos Bernoulli y Binomial de probabilidades. Identifican las principales características de una distribución Normal de probabilidades. Interpretan información estadística que involucra distribuciones de probabilidad Binomial y el Normal. Resuelven problemas que involucran los modelos Binomial y Normal. Argumentan cuándo se puede modelar una situación o fenómeno con una distribución Binomial o Normal. Modelan fenómenos o situaciones cotidianas, científicas y sociales mediante distribuciones Binomiales y Normales. Modelan situaciones que involucran aproximar una distribución Binomial mediante el modelo Normal.

DURACIÓN

6 horas pedagógicas.

DESARROLLO

Se puede usar las siguientes actividades como ejemplos de evaluaciones para la unidad 3, cada una por sí misma o en conjunto. Conviene que trabajen colaborativamente algunas para que discutan y propongan estrategias que permitan llegar a la o las soluciones posibles.

I.IDENTIFICAR EL SENTIDO DE CONTAR CON DATOS DISTRIBUIDOS MEDIANTE UNA DISTRIBUCIÓN BINOMIAL O UNA NORMAL.

4. Se sabe que una variable aleatoria se distribuye normalmente:

   a. ¿Qué se puede decir de la variable aleatoria en términos generales?

   SOLUCIÓN:

   Se espera que los estudiantes respondan algunas de las siguientes afirmaciones:

   - Que la media, mediana y moda coinciden en su valor (que es μ).
   - Que, en general se debe estandarizar la variable para poder calcular las probabilidades.

   b. ¿En qué aporta saber que los datos son Normales en su distribución para entender cómo se distribuyen?

   SOLUCIÓN:

   Se espera que los estudiantes respondan algunas de las siguientes afirmaciones:

   - Que la campana es simétrica y centrada en el valor μ
   - Que la mayoría de los valores se encuentran en torno al valor μ, y el porcentaje que queda en las colas, es muy pequeño.
   - Que hay probabilidades icónicas encerradas en intervalos que se encuentran a $1,2$ y 3 desviaciones estándar hacia la derecha y la izquierda.

   c. ¿Qué parámetros necesitas para determinar probabilidades?

   SOLUCIÓN:

   Se necesitan de dos parámetros para calcular las probabilidades mediante la estandarización, estos corresponden a la media μ y la varianza σ^2 o desviación estándar σ.

   d. ¿Cuáles son las principales diferencias entre variables aleatorias que corresponden a modelos Binomiales de distribución, respecto de las que corresponden a modelos Normales?

   SOLUCIÓN:

La primera diferencia radica en que, la Distribución Normal se utiliza para modelar fenómenos continuos y la Binomial para fenómenos discretos dicotómicos. En segundo lugar, difieren en su proceso para calcular probabilidades, pues la Normal utiliza la tabla Z, en cambio, la Binomial el número combinatorio y las probabilidades de éxito y fracaso. Por otro lado, ambas usan distintos parámetros para definir su f.d.p.

5. La duración de los embarazos se distribuye normalmente con media de $268\ días$ y desviación estándar $15\ días$.

   a. ¿Cómo se interpreta que los días de duración de los embarazos se distribuyan normalmente? ¿Cuál es el beneficio de saberlo?

   SOLUCIÓN:

   Su interpretación señala que la mayor parte de los embarazos se encuentran alrededor de los 268 días de gestación.

   El beneficio de saber que los embarazos distribuyen Normal proporciona información inmediata porque, conocer el valor de la media, en este caso 268 días, indica que el tiempo esperado de gestación es de 38 semanas con una variabilidad aproximada de 2 semanas, asimismo, lo más frecuente es que los embarazos duren la misma cantidad del tiempo esperado.

   b. ¿Cuál es la probabilidad de que un embarazo dure exactamente 9 meses (de 30 días cada uno)?

   SOLUCIÓN:

   Inmediatamente se puede responder que es 0, ya que se está pidiendo la probabilidad en un valor puntual del recorrido, es decir, que sea exactamente $270\ días$, en otras palabras, $P(X = 270)$ con X: $Duración\ de\ los\ embarazos, en\ días$.

   c. En Chile, una mujer que tenga sistema de salud Fonasa puede optar a financiar su parto mediante el bono PAD (Pago asociado al diagnóstico). Sin embargo, este no es válido si el parto se produce antes de las 37 semanas. ¿Cuál es la probabilidad de que una mujer elegida al azar no pueda usar el bono PAD (considerando que la media y desviación estándar son las mismas para este grupo de mujeres)?

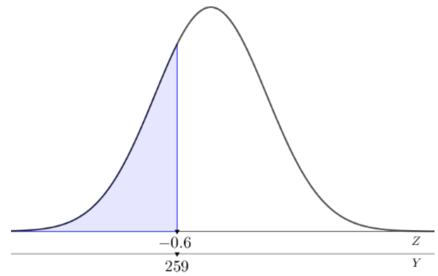
   SOLUCIÓN:

Sea:

$$Y: Duración\ de\ los\ embarazos,$$

$en\ días.$

Se tiene que $Y \sim N(\mu = 268; \sigma^2 = 15^2)$, luego, como se menciona que el bono PAD no será válido antes de las 37 semanas, este valor debe pasarse a días, así, $37 \cdot 7 = 259\ días$. De esta forma, se pide $P(Y < 259)$. Así:



z	0	0.01	0.02
-0.7	0.2420	0.2389	0.2358
-0.6	0.2743	0.2709	0.2676
-0.5	0.3085	0.3050	0.3015

$P(Y < 259) = P\left(Z < \frac{259-268}{15}\right)$

$= P(Z < -0{,}6)$

$= 0{,}2743$

Por lo tanto, la probabilidad de que una mujer no pueda usar el bono PAD es de un 27,43%

II.MODELAR SITUACIONES CON UNA DISTRIBUCIÓN BINOMIAL Y UNA DISTRIBUCIÓN NORMAL

3. En 2019, una empresa debe reducir personal por problemas económicos. Toma la decisión según los sueldos, por lo que despedirá a todos los que ganen más de $\$\,6\,000\,000$, y considera que los sueldos se distribuyen normalmente con desviación estándar $\$1\,800\,000$ en la empresa. Esa reducción de personal corresponderá al 10% de los trabajadores.

   a. ¿Qué porcentaje de trabajadores gana entre la media de sueldos y los $\$6\,000\,000$?

   SOLUCIÓN:

   Sea:

   $$X: \textit{Sueldo de los trabajadores de la empresa, en pesos.}$$

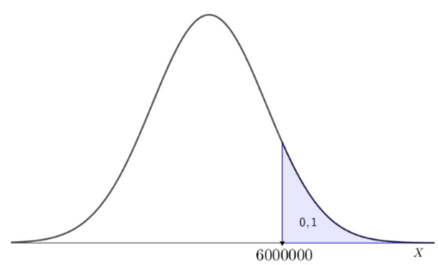
   Luego, se tiene que $X \sim N(\mu; \sigma^2 = 1\,800\,000^2)$. En este caso, no se necesita el valor de la media ya que, el enunciado menciona que la reducción del personal corresponderá al 10% de los trabajadores que reciban un sueldo superior a seis millones de pesos, así:

   $P(X \geq 6\,000\,000) = 0{,}1$

   $\Rightarrow 1 - P(X < 6\,000\,000) = 0{,}1$

   $\Rightarrow P(X < 6\,000\,000) = 0{,}9$

   $\Rightarrow P(X < 6\,000\;\,000) = 0{,}9$



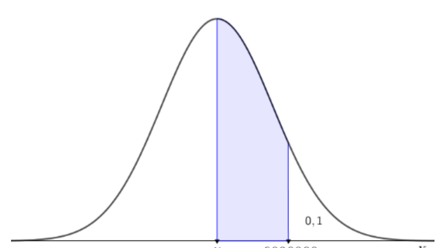
   Ahora, se pide la probabilidad de personas que reciben entre el sueldo medio y los $\$6\,000\,000$, así:

   $P(\mu \leq X \leq 6\,000\,000)$

   $= P(X \leq 6\,000\,000) - P(X \leq \mu)$

   $= 0{,}9 - 0{,}5$

   $= 0{,}4$

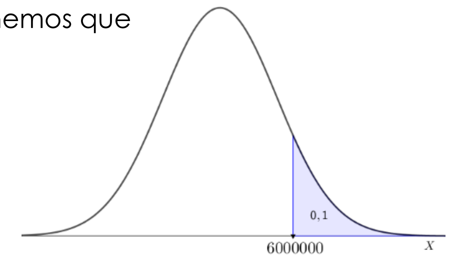


   Por lo tanto, el 40% de los trabajadores recibe un sueldo entre la media y los $\$6\,000\,000$.

b. ¿Cuál es la puntuación Z en $6 000 000? (Busca el porcentaje de la curva Normal que se considera hasta ese sueldo).

SOLUCIÓN:

Para el sueldo de $\$6\,000\,000$ en puntaje Z, tenemos que



z	0.07	0.08	0.09
1.1	0.8790	0.8810	0.8830
1.2	0.8980	0.8997	0.9015
1.3	0.9147	0.9162	0.9177

$P(X \geq 6\,000\,000) = 0{,}1$

$\Longrightarrow 1 - P(X < 6\,000\,000) = 0{,}1$

$\Longrightarrow P(X < 6\,000\,000) = 0{,}9$

$\Longrightarrow P\left(Z < \frac{6\,000\,000 - \mu}{1\,800\,000}\right) = 0{,}9$

$\Longrightarrow \phi\left(\frac{6\,000\,000 - \mu}{1\,800\,000}\right) = 1{,}28$

Por lo que el sueldo de $\$6\,000\,000$ correspondiente a la variable X, en puntaje Z es $1{,}28$

c. ¿Cuál es el promedio actual de sueldos en esta empresa?

SOLUCIÓN:

Para calcular el sueldo promedio de la empresa, se utiliza la información entregada en el enunciado, así:

$P(X \geq 6\,000\,000) = 0{,}1$

$\Longrightarrow 1 - P(X < 6\,000\,000) = 0{,}1$

$\Longrightarrow P(X < 6\,000\,000) = 0{,}9$

$\Longrightarrow P\left(Z < \frac{6\,000\,000 - \mu}{1\,800\,000}\right) = 0{,}9$

$\Longrightarrow \phi\left(\frac{6\,000\,000 - \mu}{1\,800\,000}\right) = 0{,}9$

$\Longrightarrow \frac{6\,000\,000 - \mu}{1\,800\,000} = 1{,}28$

$\Longrightarrow 6\,000\,000 - \mu = 1\,800\,000 \cdot 1{,}28$

$\Longrightarrow \mu = 6\,000\,000 - 2\,304\,000$

$\Longrightarrow \mu = 3\,696\,000$

Por lo tanto, el promedio actual de los sueldos de la empresa es de $\$3\,696\,000$ y de esta forma, se tiene que $X \sim N(\mu = 3\,696\,000; \sigma^2 = 1\,800\,000^2)$

d. ¿Cómo crees que las grandes empresas toman decisiones fundamentadas matemáticamente para mejorar sus ganancias? ¿Qué decisión tomarías tú en su lugar, basándote en datos estadísticos?

SOLUCIÓN:

A continuación, se entrega una posible respuesta dada por los estudiantes:

- Las empresas distribuyen sus sueldos de tal forma que las personas que reciben los sueldos más altos o bajos son pocas (similar a lo que ocurre en la Distribución Normal con las colas), por el contrario, la mayoría de los sueldos de las personas se concentran entorno a la remuneración promedio (similar a lo que ocurre con la Distribución Normal en el centro).

Para la segunda pregunta, no se otorgan respuestas especificas porque se pretende que la decisión de cada estudiante varíe según sus propias perspectivas.

III. MODELAR SITUACIONES, USANDO LA DISTRIBUCIÓN NORMAL Y EL TEOREMA DEL LÍMITE CENTRAL, Y APROXIMAR UNA DISTRIBUCIÓN BINOMIAL MEDIANTE UNA NORMAL

1. En promedio, la temperatura corporal de los niños es de $37\,°C$, con una desviación estándar de $0{,}34\,°C$. Si se selecciona al azar una muestra de 80 niños, ¿cuál es la probabilidad de obtener una media menor a $36\,°C$?

   SOLUCIÓN

   Sea:

   X_i: *Temperatura corporal del niño i en* °C, con $i = 1,2,3,\dots,80$

   Luego, se tiene que

$$E(X_i) = 37°\mathrm{C}$$

$$\sigma = \sqrt{Var(X_i)} = 0{,}34°\mathrm{C}$$

   Ahora, se pide $P(\overline{X} < 36)$, así

   $P(\overline{X} < 36) = P\left(\frac{\overline{X}-37}{0{,}34}\cdot\sqrt{80} < \frac{36-37}{0{,}34}\cdot\sqrt{80}\right)$

   $= P\left(\frac{\overline{X}-37}{0{,}34}\cdot\sqrt{80} < \frac{36-37}{0{,}34}\cdot\sqrt{80}\right)$

   $\approx P(Z < -26{,}31)$

   $= \phi(-26{,}31)$

   $= 0$

   Por lo tanto, la probabilidad de obtener una media menor a $36\,°C$ es de 0%.

   a. Para determinar esa probabilidad, ¿hay que saber la distribución de las temperaturas de la población?

      SOLUCIÓN:

      No es necesario saber la distribución que posea la población de la que se extraen las muestras, pues se utiliza el Teorema Central del Límite que asegura la convergencia a una Distribución Normal, cuando se cumpla la condición $n > 30$.

   b. Argumenten por qué pueden aplicar el teorema del límite central en este caso.

SOLUCIÓN:

Se puede aplicar ya que se pregunta por la probabilidad de obtener una media menor a $36\ °C$, cuando $n > 30$. Además, son sucesos independientes.

c. ¿Qué parámetros usarán para la distribución Normal de las medias de las muestras?

SOLUCIÓN:

Se necesita $\mu = E(X_i) = 37°\mathrm{C}$ y $\sigma = \sqrt{\frac{Var(X_i)}{n}} = \sqrt{\frac{(0{,}34)^2}{80}} = \sqrt{\frac{289}{200\ 000}}\ °\mathrm{C}$.

d. ¿Por qué es necesario estandarizar para conocer $P(\bar{X} < 36)$?

SOLUCIÓN:

Para poder utilizar la tabla Z, y estimar dicha probabilidad, ya que la variable aleatoria $Z = \frac{\bar{X}-\mu}{\sigma} \cdot \sqrt{n}$ se aproxima a una $N(0{,}1)$.

e. ¿Cómo se interpreta la probabilidad obtenida?

SOLUCIÓN:

La probabilidad de obtener una temperatura media corporal menor a $36\ °\mathrm{C}$ en los 80 niños seleccionados es de un 0%, en otras palabras, es casi imposible que los niños tengan una temperatura media corporal menor a $36\ °\mathrm{C}$.

f. ¿En qué ayudó conocer la distribución de las medias de las muestras?

SOLUCIÓN:

Ayuda a poder calcular la probabilidad pedida, sin necesidad de saber la distribución original de la población.

2. Una empresa de agua mineral sabe que las botellas pequeñas se llenan con $350\ cc$ de agua, con una desviación estándar de $3\ cc$.

a. ¿Cuál es la probabilidad de que una muestra de 50 botellas pequeñas tenga una cantidad media de, al menos, $354\ cc$?

SOLUCIÓN:

Sea:

$X_i: Cantidad\ de\ agua\ mineral\ en\ una\ botella\ pequeña, en\ cc$, con $i = 1,2,3, \ldots ,50$

Luego, se tiene que

$E(X_i) = 350\ cc$

$\sigma = \sqrt{Var(X_i)} = 3\ cc$

Ahora, se pide $P(\bar{X} \geq 354)$, así

$P(\bar{X} \geq 354) = 1 - P(\bar{X} < 354)$

$= 1 - P\left(\frac{\bar{X}-350}{3} \cdot \sqrt{50} < \frac{354-350}{3} \cdot \sqrt{50}\right)$

$\approx 1 - P(Z < 9{,}43)$

$= 1 - \phi(9{,}43)$

$= 1 - 1$

$= 0$

En consecuencia, la probabilidad de que una muestra de 50 botellas pequeñas tenga una cantidad media de, al menos, $354\ cc$ es de 0%.

b. A partir del resultado anterior, ¿será razonable pensar que las botellas pequeñas en realidad tienen una media de $350\ cc$?

SOLUCIÓN:

Es razonable pensar que ese es su valor real, porque a partir de los cálculos realizados, se aprecia que la probabilidad que la media de las 50 botellas sea menor a $354\ cc$ es segura, en otras palabras 100%.

c. Si la media no es $350\ cc$, ¿se engaña a los consumidores?

SOLUCIÓN:

Sí, se engaña a los consumidores.

d. ¿En qué podría afectar a la empresa considerar una media diferente a la "media real"?

SOLUCIÓN:

Podría llevar a pérdidas o ganancias millonarias, si se mira desde que una producción masiva que posea este error.

e. ¿Cómo se puede usar este modelo de probabilidad en problemas de contextos similares? ¿Cuál es su aporte?

SOLUCIÓN:

Se puede usar para modelar fenómenos parecidos al que se presenta en el ejemplo, y también del consumo de agua, electricidad. El aporte es saber si la media que informan las empresas es verídica o no, tanto para ellos como para los consumidores.

3. Una máquina industrial de deshidratado de manzanas presenta una falla que no se detectó a tiempo. Esto causó que $\frac{1}{8}$ de la producción de 4000 kilogramos presentara problemas de secado.

   a. Antes de calcular las probabilidades, ¿cómo se podría ajustar la distribución de los datos para modelarlos mediante una distribución Normal?

   SOLUCIÓN:

   Considerando que la distribución que mejor se ajusta a este ejercicio es una Binomial, ya que se entrega la probabilidad de éxito, y el número de repeticiones. Sea:

$$X: \textit{Número de kilogramos que tuvieron problema de secado, de } 4000\ kg.$$

   De forma que $X \sim Bin\left(n = 4\,000; p = \frac{1}{8}\right)$, y su función de probabilidad está dada por:

$$P(X = x) = \begin{cases} \binom{4\,000}{x} \cdot \left(\frac{1}{8}\right)^{x} \cdot \left(\frac{7}{8}\right)^{4000-x} & si\ x = 0,1,2,\dots,4\,000 \\ 0 & en\ otro\ caso \end{cases}$$

   Y desde allí extraer los parámetros de la Distribución Normal que corresponden a:

$$\mu = n \cdot p$$

$$\sigma = \sqrt{n \cdot p \cdot q}$$

   b. ¿Cuáles son las restricciones para hacerlo? ¿Se cumplen esas restricciones?

   SOLUCIÓN:

   La restricción para hacerlo viene dada por la corrección de continuidad. Para el caso anterior las restricciones se cumplen porque

$$n = 4\,000 > 10$$

$$n \cdot p = 4\,000 \cdot \frac{1}{8} = 500 \geq 5$$

$$n \cdot q = 4\,000 \cdot \frac{7}{8} = 3\,500 \geq 5$$

c. ¿Cuál sería el aporte de hacerlo?

SOLUCIÓN:

Simplificar el cálculo de probabilidades, ya que, si se modela con una Distribución Binomial, el número combinatorio y las potencias serán números excesivamente grandes y pequeñas respectivamente.

d. ¿Cuál es la probabilidad de que, al tomar una muestra de 25 kilogramos, se encuentre un máximo de 3 kilogramos con problemas de secado?

SOLUCIÓN:

Sea:

$$Y: \textit{Número de kilogramos que tuvieron problema de secado, de } 25\ kg.$$

De forma que $Y \sim Bin\left(n = 25; p = \frac{1}{8}\right)$. En este caso, no se puede aproximar, ya que la condición $n \cdot p \geq 5$ no se cumple. Por lo que la función de probabilidad de Y está dada por:

$$P(Y = y) = \begin{cases} \binom{25}{y} \cdot \left(\frac{1}{8}\right)^{y} \cdot \left(\frac{7}{8}\right)^{25-y} & si\ y = 0,1,2, \ldots, 25 \\ 0 & en\ otro\ caso \end{cases}$$

Luego, la probabilidad pedida es $P(Y \leq 3)$, así:

$$P(Y \leq 3) = P(Y = 0) + P(Y = 1) + P(Y = 2) + P(Y = 3)$$

Calculando las probabilidad por separado:

$$P(Y = 0) = \binom{25}{0} \cdot \left(\frac{1}{8}\right)^{0} \cdot \left(\frac{7}{8}\right)^{25-0} = 1 \cdot 1 \cdot \left(\frac{7}{8}\right)^{25} = \left(\frac{7}{8}\right)^{25} = 0{,}0355$$

$$P(Y = 1) = \binom{25}{1} \cdot \left(\frac{1}{8}\right)^{1} \cdot \left(\frac{7}{8}\right)^{25-1} = 25 \cdot \left(\frac{1}{8}\right)^{1} \cdot \left(\frac{7}{8}\right)^{24} = 0{,}1268$$

$$P(Y = 2) = \binom{25}{2} \cdot \left(\frac{1}{8}\right)^{2} \cdot \left(\frac{7}{8}\right)^{25-2} = 300 \cdot \left(\frac{1}{8}\right)^{2} \cdot \left(\frac{7}{8}\right)^{23} = 0{,}2173$$

$$P(Y = 3) = \binom{25}{3} \cdot \left(\frac{1}{8}\right)^{3} \cdot \left(\frac{7}{8}\right)^{25-3} = 2300 \cdot \left(\frac{1}{8}\right)^{3} \cdot \left(\frac{7}{8}\right)^{22} = 0{,}238$$

Finalmente,

$$P(Y \leq 3) = 0{,}0355 + 0{,}1268 + 0{,}2173 + 0{,}238 = 0{,}6176$$

Por lo tanto, la probabilidad que en una muestra de $25\ kg$, se encuentre como máximo $3\ kg$ con problemas de secado es de $61{,}76\%$.

e. ¿Cuál es la probabilidad de que, al tomar una muestra de 100 kilogramos, se encuentre máximo 50 kilogramos con problemas de secado?

SOLUCIÓN:

Sea:

$$W\text{: }Número\ de\ kilogramos\ que\ tuvieron\ problema\ de\ secado, de\ 100\ kg.$$

De forma que $W \sim Bin\left(n = 100; p = \frac{1}{8}\right)$. En este caso, se puede aproximar a una Distribución Normal, ya que las condiciones se cumplen

$$n = 100 > 10$$

$$n \cdot p = 100 \cdot \frac{1}{8} = 12{,}5 \geq 5$$

$$n \cdot q = 100 \cdot \frac{7}{8} = 87{,}5 \geq 5$$

Por lo que, para la resolución del ejercicio se utilizará la aproximación de la Binomial a la Normal. Los parámetros de la Normal corresponden a:

$$\mu = 100 \cdot \frac{1}{8} = 12{,}5$$

$$\sigma = \sqrt{100 \cdot \frac{1}{8} \cdot \frac{7}{8}} = 3{,}3072$$

Luego, la probabilidad pedida es $P(X \leq 50)$, así, haciendo la corrección por continuidad se obtiene:

$P(W \leq 50) \approx P(W \leq 50 + 0{,}5)$

$= P(W \leq 50{,}5)$

$= P\left(\frac{W - 12{,}5}{3{,}3072} \leq \frac{50{,}5 - 12{,}5}{3{,}3072}\right)$

$\approx P(Z \leq 11{,}49)$

$= \phi(11{,}49)$

$= 1$

Por lo tanto, la probabilidad que en una muestra de $100\ kg$, se encuentre como máximo $25\ kg$ con problemas de secado es de 100%.

4. El cáncer a la piel aumenta considerablemente cada año y causa la muerte en pacientes que no lo detectan a tiempo. Si se descubre a tiempo, el porcentaje de supervivencia es del 90% ¿Cuál es la probabilidad de que 200 o más personas sobrevivan al cáncer, de una muestra de 300 personas diagnosticadas a tiempo?

SOLUCIÓN:

Sea:

$$X\text{: número de personas sobrevivientes del cáncer a la piel.}$$

Se tiene que $X \sim Bin(n = 300, p = 0{,}9)$, y se pide $P(X \geq 200)$

Como los valores (para el número combinatorio), son muy altos, se utilizará la aproximación a la Distribución Normal, verificando que se cumplan las condiciones.

$$n = 300 > 10$$

$$n \cdot p = 300 \cdot 0{,}9 = 270 \geq 5$$

$$n \cdot q = 300 \cdot 0{,}1 = 30 \geq 5$$

En efecto, las condiciones se cumplen, así:

$P(X \geq 200) = 1 - P(X < 200)$

$\approx 1 - P(X < 200 + 0{,}5)$

$= 1 - P(X < 200{,}5)$

$= 1 - P\left(\frac{X-(300\cdot 0{,}9)}{\sqrt{300\cdot 0{,}9\cdot 0{,}1}} < \frac{200{,}5-(300\cdot 0{,}9)}{\sqrt{300\cdot 0{,}9\cdot 0{,}1}}\right)$

$\approx 1 - P\left(Z < \frac{200{,}5-270}{\sqrt{27}}\right)$

$= 1 - P(Z < -13{,}38)$

$= 1 - \phi(-13{,}38)$

$= 1 - 0$

$= 1$

Por lo tanto, la probabilidad que 200 o más personas sobrevivan al cáncer, de una muestra de 300 personas es de un 100%

a. Esta situación, ¿se puede aproximar mediante una distribución Normal? Argumenta e indica los beneficios de hacerlo.

SOLUCIÓN:

Sí, porque se cumplen las condiciones. En efecto, para este ejercicio los parámetros son $n = 300$ y $p = 0{,}9$, entonces

$$n = 300 > 10$$

$$n \cdot p = 300 \cdot 0{,}9 = 270 \geq 5$$

$$n \cdot (1 - p) = 300 \cdot (1 - 0{,}9) = 30 \geq 5$$

Respecto a los beneficios se observa la simplicidad del cálculo, porque no es conveniente utilizar la Distribución Binomial para la probabilidad buscada $P(X \leq 200)$ con el valor de $n = 300$. Cabe destacar que la calculadora no permite determinar los números combinatorios desde $x = 40$.

b. ¿Qué otras preguntas de interés podrías responder en esta situación? ¿Tienes suficientes datos para ello? Argumenta.

SOLUCIÓN:

Se pueden formular preguntas parecidas a la que se menciona arriba, y no se necesitan más datos para poder responder, ya que el enunciado entrega todos los parámetros necesarios para realizar el cálculo de probabilidades. Aquí se propone un ejemplo, usando la variable definida en 4.a.:

¿Cuál es la probabilidad de que los sobrevivientes de cáncer a partir de la muestra de 300 personas, esté entre 150 y 260 personas?

Con lo definido arriba, se pide:

$P(150 < X < 260) \approx P(150 - 0{,}5 < X < 260 + 0{,}5)$

$= P(149{,}5 < X < 260{,}5)$

$= P\left(\frac{149{,}5-(300 \cdot 0{,}9)}{\sqrt{300 \cdot 0{,}9 \cdot 0{,}1}} < \frac{X-(300 \cdot 0{,}9)}{\sqrt{300 \cdot 0{,}9 \cdot 0{,}1}} < \frac{260{,}5-(300 \cdot 0{,}9)}{\sqrt{300 \cdot 0{,}9 \cdot 0{,}1}}\right)$

$\approx P\left(\frac{149{,}5-270}{\sqrt{27}} < Z < \frac{260{,}5-270}{\sqrt{27}}\right)$

z	0.02	0.03	0.04
-1.9	0.0274	0.0268	0.0262
-1.8	0.0344	0.0336	0.0329
-1.7	0.0427	0.0418	0.0409

$= P(-23{,}19 < Z < -1{,}83)$

$= P(Z < -1{,}83) - P(Z < -23{,}19)$

$= \phi(-1{,}83) - \phi(-23{,}19)$

$= 0{,}0336 - 0$

$= 0{,}0336$

Por lo tanto, la probabilidad de que los sobrevivientes de cáncer estén entre 150 y 260 persona de una muestra de 300 personas es de un $3{,}36\%$

BIBLIOGRAFÍA

Batanero, C. y Godino, J. (2001). Análisis de Datos y su Didáctica. Recuperado de https://www.ugr.es/~batanero/pages/ARTICULOS/Apuntes.pdf

Meyer, P. L. (1992). Probabilidad y Aplicaciones Estadísticas. Addison-Wesley Iberoamericana.

Osorio, G., Norambuena, P., Romante, M., Gaete, D., Díaz, J., Celedón, J., et al. (2019). *Texto del estudiante Matemática 3° y 4° medio*. Santiago de Chile: SM.

Rincón, L. (2013). *Introducción a la probabilidad*. Ciudad de México, México.

Walpole, R., Myers, R., Myers, S. y Ye, K. (2007). Probabilidad y estadística para ingeniería y ciencias. México: Pearson Educación.

ANEXOS

ANEXO 1: TABLA Z

Valores de la función distribución Normal Estándar

$$\phi(z) = P(Z \leq z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du$$

z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-3.9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
-3.8	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
-3.7	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
-3.6	0.0002	0.0002	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
-3.5	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
-3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
-3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
-3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
-3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
-3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
-2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
-2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
-2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
-2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
-2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
-2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
-2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
-2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
-1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
-1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
-1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
-1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
-1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
-1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
-1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
-1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
-1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
-0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
-0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
-0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
-0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
-0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
-0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
-0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
-0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
-0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641

z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

ANEXO 2: SIMULACIÓN. ACTIVIDAD 2, DETERMINAR PROBABILIDADES QUE IMPLIQUEN EL USO DEL TEOREMA DEL LÍMITE CENTRAL. PREGUNTA 2.

Nº de Muestra	Ind.1	Ind.2	Ind.3	Ind.4	Ind.5	Ind.6	Promedio de la muestra
1	71,62	78,70	63,29	104,25	56,28	59,43	72,26
2	82,44	69,41	97,55	71,99	77,56	80,96	79,99
3	93,15	79,08	85,85	56,81	82,58	66,71	77,36
4	63,74	63,00	79,77	55,29	85,60	74,05	70,24
5	78,21	68,67	99,50	62,52	77,89	92,90	79,95
6	96,00	78,25	79,94	69,12	98,31	76,44	83,01
7	71,68	75,05	83,66	49,47	96,40	74,19	75,08
8	61,68	85,24	85,84	70,68	83,87	66,69	75,66
9	77,11	56,50	96,71	80,96	63,39	88,57	77,21
10	84,47	86,83	72,01	86,23	77,64	80,60	81,30
11	84,25	75,46	55,58	73,37	82,78	70,84	73,71
12	71,09	83,83	67,54	78,70	81,16	77,56	76,65
13	73,42	101,00	79,57	68,90	76,97	56,52	76,06
14	81,98	72,42	72,45	87,12	68,67	51,32	72,33
15	71,68	74,77	98,57	93,07	67,90	53,13	76,52
16	84,86	66,04	69,87	81,23	96,66	55,79	75,74
17	97,38	90,68	102,73	83,44	91,21	69,91	89,22
18	61,01	47,92	86,26	93,37	78,24	76,97	73,96
19	76,58	106,20	70,12	82,04	55,97	65,85	76,13
20	95,05	65,02	104,29	81,84	75,66	54,31	79,36
21	81,22	59,47	89,02	76,03	88,57	107,03	83,56
22	61,98	66,14	83,10	72,00	89,25	48,43	70,15
23	86,91	70,34	76,16	89,08	74,75	71,71	78,16
24	88,05	68,90	76,38	58,19	109,65	68,79	78,33
25	70,55	73,32	93,54	73,85	71,60	106,72	81,60
26	90,92	78,08	89,11	66,58	65,80	77,44	77,99
27	68,58	94,38	87,45	77,63	54,21	77,74	76,67
28	94,27	87,51	94,09	40,67	83,21	68,90	78,11
29	82,33	84,53	71,92	95,50	81,91	55,36	78,59
30	86,86	73,87	85,31	85,01	71,28	79,33	80,28
31	50,56	87,78	76,28	83,71	57,44	99,66	75,90
32	72,92	98,26	102,80	77,59	87,09	84,89	87,26
33	61,36	73,50	75,02	81,78	80,63	80,99	75,55
34	73,02	83,42	81,45	85,75	69,44	73,08	77,69
35	79,13	75,34	70,14	70,69	92,28	92,54	80,02
36	76,39	58,64	60,92	63,57	68,63	94,48	70,44
37	76,59	66,48	57,04	98,42	86,65	84,52	78,28

38	80,26	80,01	92,00	68,52	86,26	85,16	82,04
39	81,28	61,68	89,74	76,73	75,07	74,62	76,52
40	91,90	53,60	66,55	79,34	92,23	68,68	75,38
41	65,09	87,47	87,72	91,23	54,04	95,63	80,20
42	88,94	79,98	61,32	57,47	61,67	66,30	69,28
43	88,27	67,12	73,47	92,51	67,38	83,45	78,70
44	74,14	86,44	62,29	69,49	45,91	58,83	66,18
45	57,21	45,73	79,59	91,22	100,64	77,40	75,30
46	77,76	78,64	99,97	82,38	91,94	73,44	84,02
47	83,46	87,59	75,29	63,72	93,59	99,29	83,82
48	78,64	94,68	83,20	77,22	100,51	75,10	84,89
49	98,54	78,84	55,39	99,81	83,03	81,95	82,93
50	77,13	94,94	74,36	57,88	99,50	71,86	79,28
51	81,81	62,80	90,25	77,46	77,09	105,70	82,52
52	76,28	81,75	71,08	65,47	55,13	57,92	67,94
53	83,94	79,78	77,62	90,28	73,49	89,13	82,37
54	88,17	59,74	86,56	104,22	77,66	59,54	79,31
55	68,26	78,91	78,95	76,61	70,03	75,01	74,63
56	87,09	76,69	74,62	60,53	64,21	77,54	73,45
57	96,08	79,37	76,93	63,48	80,78	71,28	77,99
58	67,14	90,16	69,22	93,80	72,02	69,57	76,98
59	71,53	67,60	49,39	84,97	69,43	74,23	69,52
60	87,65	74,87	67,00	61,55	59,80	69,86	70,12
61	74,07	67,09	96,06	90,20	94,88	89,87	85,36
62	74,19	97,88	71,43	75,62	112,76	99,83	88,62
63	81,91	67,06	66,26	112,63	69,30	65,61	77,13
64	77,75	50,86	71,11	78,53	73,65	68,81	70,12
65	106,37	78,06	50,99	88,34	72,96	67,31	77,34
66	83,11	91,75	73,06	89,00	76,25	77,18	81,72
67	63,13	79,23	66,36	68,33	82,42	104,47	77,32
68	67,10	83,16	94,12	82,04	92,55	60,41	79,90
69	69,54	100,45	82,72	74,85	79,39	72,70	79,94
70	72,54	81,37	71,87	71,19	77,49	79,68	75,69
71	93,96	78,90	78,41	88,82	84,82	95,33	86,71
72	80,91	56,86	77,41	69,19	88,15	91,38	77,32
73	87,79	87,72	77,91	74,25	70,29	65,25	77,20
74	89,35	85,08	86,44	57,50	79,89	72,81	78,51
75	83,85	59,22	47,01	88,29	82,13	56,50	69,50
76	84,06	81,14	74,37	79,16	78,92	85,00	80,44
77	92,33	76,59	85,01	82,77	90,13	85,13	85,33
78	62,76	70,18	77,91	87,17	75,44	88,34	76,97
79	61,44	64,26	67,99	77,74	61,56	79,89	68,81
80	66,95	68,43	85,32	91,50	60,57	88,06	76,80
81	63,93	69,66	67,06	72,35	93,70	76,68	73,90

82	68,95	68,14	65,54	85,46	71,88	108,65	78,10
83	73,25	77,54	86,59	55,96	78,04	62,86	72,37
84	83,20	55,82	57,15	84,61	103,31	61,45	74,26
85	81,06	93,66	82,69	46,69	61,77	82,66	74,75
86	51,84	82,11	107,22	88,88	88,90	84,84	83,97
87	87,08	71,93	90,89	52,09	91,93	67,21	76,85
88	62,70	68,41	86,03	97,50	96,96	83,69	82,55
89	73,95	71,25	57,80	90,72	68,85	86,39	74,82
90	79,75	66,42	76,23	83,31	66,05	62,33	72,35
91	83,29	66,45	90,72	74,45	67,00	87,15	78,18
92	75,02	66,65	91,22	96,44	65,74	87,48	80,42
93	60,65	84,97	97,14	64,40	43,85	63,96	69,16
94	78,41	87,96	82,06	60,07	80,97	75,87	77,56
95	67,35	104,31	66,66	77,27	86,82	91,82	82,37
96	67,17	80,22	83,06	108,67	70,32	66,38	79,30
97	86,88	79,03	63,15	88,72	73,12	90,08	80,16
98	86,36	68,56	95,12	72,04	86,37	117,38	87,64
99	80,19	54,21	61,05	104,30	82,90	76,53	76,53
100	76,49	81,86	92,29	90,87	84,09	86,24	85,31

ANEXO 3: SIMULACIÓN 1. ACTIVIDAD 2, DETERMINAR PROBABILIDADES QUE IMPLIQUEN EL USO DEL TEOREMA DEL LÍMITE CENTRAL. PREGUNTA 6.A.

Nº de Muestra	Ind.1	Ind.2	Ind.3	Ind.4	Ind.5	Ind.6	Promedio de la muestra
1	76,61	102,78	83,41	73,67	53,35	80,17	78,33
2	82,00	69,94	78,08	58,67	88,69	72,00	74,89
3	71,33	88,40	64,98	90,15	74,36	89,84	79,84
4	53,24	95,33	82,56	67,54	83,74	85,53	77,99
5	96,07	91,05	81,30	54,38	54,69	81,78	76,54
6	76,72	67,23	64,36	77,04	85,52	80,95	75,30
7	80,11	59,90	64,64	61,22	78,68	81,98	71,09
8	45,69	86,51	95,31	86,23	61,55	80,30	75,93
9	79,60	83,49	86,05	64,10	64,60	95,51	78,89
10	77,07	61,76	60,64	78,13	89,86	75,47	73,82
11	79,77	60,27	96,01	73,55	70,09	87,02	77,79
12	63,25	92,02	88,43	68,39	86,66	78,48	79,54
13	65,32	55,27	82,40	100,95	86,58	65,18	75,95
14	47,80	73,45	91,54	73,26	66,91	72,95	70,99
15	62,96	108,64	90,30	89,22	71,89	88,02	85,17
16	89,08	87,96	77,13	85,67	87,49	72,43	83,29
17	93,74	105,01	82,22	62,56	59,43	62,67	77,60
18	76,85	84,76	68,38	58,72	81,19	80,60	75,09
19	64,72	73,65	62,89	83,23	63,99	80,22	71,45
20	81,49	64,26	54,95	69,63	76,68	93,60	73,43
21	87,75	76,70	73,88	92,60	83,36	65,28	79,93
22	78,04	102,27	83,18	71,05	97,46	80,75	85,46
23	101,03	72,04	86,62	79,81	67,43	79,26	81,03
24	83,17	80,38	80,03	69,03	63,48	84,84	76,82
25	97,63	53,07	62,61	73,76	77,79	63,37	71,37
26	78,95	64,13	69,09	74,69	70,17	86,95	74,00
27	65,26	78,00	47,82	92,70	100,09	106,36	81,70
28	80,37	89,35	64,64	50,56	75,71	89,06	74,95
29	68,17	72,57	77,43	64,32	81,95	65,60	71,67
30	61,84	78,58	74,34	75,30	85,69	67,82	73,93
31	71,87	76,80	80,69	65,97	66,92	76,78	73,17
32	63,17	89,01	65,30	75,26	78,19	77,32	74,71
33	86,14	69,03	60,21	79,92	78,37	89,40	77,18
34	88,05	96,48	59,57	72,42	83,72	87,29	81,25
35	61,57	58,83	78,37	97,52	73,29	80,08	74,94
36	74,35	61,51	75,57	85,61	80,46	73,47	75,16

37	86,04	79,84	76,46	82,28	54,21	61,51	73,39
38	86,24	51,50	84,32	74,67	89,46	98,56	80,79
39	60,35	99,64	76,19	65,66	66,48	76,21	74,09
40	85,11	68,08	96,08	73,92	69,86	77,82	78,48
41	71,67	69,48	102,21	85,91	97,42	91,77	86,41
42	93,40	93,38	82,46	81,81	85,92	65,63	83,77
43	69,29	69,08	87,52	75,02	62,91	87,30	75,19
44	60,39	70,26	51,18	63,62	78,01	81,40	67,48
45	83,63	82,23	48,87	79,96	65,33	62,75	70,46
46	82,45	81,70	74,29	71,73	65,66	63,50	73,22
47	58,33	72,84	69,19	82,67	72,39	81,44	72,81
48	70,91	80,91	78,41	63,54	61,79	76,75	72,05
49	112,35	83,86	60,72	68,49	82,08	82,03	81,59
50	75,32	59,76	70,62	92,55	90,89	63,86	75,50
51	61,06	87,78	72,41	75,53	74,14	81,65	75,43
52	83,95	79,73	68,42	68,66	92,81	74,32	77,98
53	95,83	78,69	74,65	74,30	82,61	78,46	80,76
54	82,07	84,99	95,73	78,60	98,52	85,42	87,55
55	75,67	68,39	60,47	65,61	75,88	65,37	68,56
56	75,99	79,07	61,47	75,06	76,62	92,75	76,83
57	86,78	57,74	58,84	80,80	52,98	59,67	66,14
58	71,90	81,87	61,75	74,58	68,10	71,77	71,66
59	81,33	86,16	83,46	76,57	66,08	85,17	79,80
60	82,72	99,84	59,21	69,43	87,20	70,72	78,19
61	67,25	79,94	98,22	83,97	86,05	81,48	82,82
62	78,83	82,80	80,97	79,20	87,72	59,32	78,14
63	63,60	99,11	80,93	81,76	71,15	70,77	77,89
64	90,04	74,89	79,19	82,02	98,24	90,49	85,81
65	85,54	84,23	64,46	74,15	71,82	79,72	76,65
66	65,40	81,58	71,00	64,86	72,05	86,23	73,52
67	85,64	76,98	76,50	69,87	83,95	58,38	75,22
68	67,11	62,35	76,21	76,53	60,71	76,15	69,84
69	80,11	70,67	66,30	42,29	86,30	79,84	70,92
70	79,19	84,76	64,18	97,33	69,37	89,70	80,75
71	65,01	78,56	73,79	61,69	62,24	55,22	66,09
72	79,67	65,98	98,39	82,70	80,15	48,61	75,91
73	75,47	80,33	81,61	63,20	52,88	88,17	73,61
74	80,69	102,28	81,62	78,29	66,28	82,88	82,01
75	82,90	76,26	55,05	86,41	83,64	97,38	80,27
76	85,05	70,19	83,93	59,19	96,03	77,06	78,57
77	83,39	72,23	76,52	85,67	74,05	71,14	77,17
78	81,06	95,00	96,22	63,57	67,60	88,65	82,02
79	51,51	80,82	71,69	87,72	88,98	63,39	74,02
80	81,46	70,53	93,34	86,01	71,77	74,69	79,63
81	68,40	68,27	65,04	61,32	65,57	82,08	68,45

82	83,70	70,03	83,58	68,49	87,89	75,14	78,14
83	69,68	64,32	83,83	89,47	92,76	87,09	81,19
84	60,37	61,54	87,84	61,53	93,63	71,42	72,72
85	64,80	71,22	65,61	57,70	66,15	73,89	66,56
86	86,99	96,07	73,53	74,47	63,98	93,22	81,38
87	56,07	66,88	122,29	87,96	76,87	97,90	84,66
88	65,29	79,55	95,24	74,70	72,94	73,03	76,79
89	84,08	79,67	68,06	70,08	90,24	56,41	74,76
90	71,25	86,67	72,54	69,81	76,20	74,21	75,11
91	76,82	94,22	61,28	70,30	71,46	56,34	71,74
92	62,11	78,32	92,57	69,96	79,85	60,83	73,94
93	59,80	83,02	65,30	85,38	88,93	91,94	79,06
94	80,11	97,98	74,73	82,24	79,23	91,93	84,37
95	89,53	84,38	51,90	59,84	84,86	69,57	73,35
96	105,50	55,90	90,17	70,95	76,89	70,94	78,39
97	97,29	87,67	75,23	66,42	74,49	94,59	82,62
98	79,10	57,75	64,01	83,67	68,84	87,93	73,55
99	81,36	88,65	84,26	85,44	85,93	63,94	81,60
100	92,11	83,13	79,46	92,69	90,38	55,09	82,14

ANEXO 4: SIMULACIÓN 2. ACTIVIDAD 2, DETERMINAR PROBABILIDADES QUE IMPLIQUEN EL USO DEL TEOREMA DEL LÍMITE CENTRAL. PREGUNTA 6.A.

Nº de Muestra	Ind.1	Ind.2	Ind.3	Ind.4	Ind.5	Ind.6	Promedio de la muestra
1	90,77	66,51	86,81	75,29	63,51	95,13	79,67
2	73,37	87,51	67,23	72,74	99,49	78,70	79,84
3	87,86	63,66	94,09	71,51	55,71	64,42	72,88
4	86,95	69,62	82,50	79,94	93,79	70,18	80,50
5	84,53	69,15	90,61	88,29	53,14	95,87	80,26
6	76,45	85,51	93,41	55,23	90,78	89,02	81,73
7	86,43	91,55	93,31	92,13	100,03	74,93	89,73
8	93,14	70,96	96,10	44,05	88,36	73,57	77,70
9	102,34	93,25	78,96	76,10	64,83	94,76	85,04
10	59,76	93,16	71,08	79,83	79,25	82,73	77,63
11	80,84	70,35	71,33	78,05	97,37	78,10	79,34
12	57,86	82,52	75,42	70,69	93,14	82,06	76,95
13	69,02	54,24	78,95	88,79	61,62	99,42	75,34
14	83,95	82,88	82,05	65,02	72,88	77,96	77,46
15	62,81	83,64	77,86	51,21	78,74	72,22	71,08
16	93,23	84,22	77,97	52,84	101,60	67,74	79,60
17	68,91	60,11	71,34	76,43	92,53	74,10	73,91
18	61,82	61,84	70,46	93,52	58,78	93,01	73,24
19	59,19	56,38	83,17	63,25	74,12	66,78	67,15
20	89,62	91,11	79,58	60,45	92,87	90,51	84,02
21	64,57	78,65	65,65	64,23	71,59	83,87	71,43
22	100,90	76,16	69,44	84,88	97,02	79,46	84,64
23	74,95	72,67	67,16	80,35	74,42	68,15	72,95
24	46,19	64,28	67,37	73,20	85,60	55,82	65,41
25	86,13	85,82	100,47	82,81	64,62	72,31	82,03
26	87,95	88,70	59,42	82,73	65,55	52,69	72,84
27	68,61	93,18	86,56	62,09	79,27	95,78	80,91
28	81,69	50,47	59,81	74,33	85,92	82,74	72,49
29	60,45	89,81	70,44	50,62	79,86	54,69	67,65
30	88,03	96,07	72,44	57,23	66,10	95,02	79,15
31	68,31	83,96	86,12	81,95	73,37	85,73	79,91
32	54,49	61,78	64,35	68,50	84,87	72,80	67,80
33	91,38	85,57	78,61	83,78	78,05	88,73	84,35
34	65,09	66,04	37,69	103,36	70,52	82,62	70,89
35	90,98	66,06	82,62	80,16	50,13	99,35	78,21
36	72,31	72,70	88,22	74,60	65,15	72,05	74,17
37	69,88	72,21	71,04	106,38	79,93	74,91	79,06

38	61,19	82,30	58,24	85,54	66,15	67,79	70,20
39	56,49	91,48	91,00	71,46	96,77	77,35	80,76
40	87,12	55,05	74,33	46,24	55,66	62,30	63,45
41	84,72	86,91	83,06	78,46	76,80	89,09	83,17
42	73,24	66,60	65,25	79,49	101,98	41,97	71,42
43	84,27	62,13	59,92	84,60	67,84	52,02	68,46
44	78,43	57,14	82,36	93,12	58,53	77,11	74,45
45	91,17	86,58	93,59	94,74	68,71	83,96	86,46
46	80,01	94,27	79,90	73,19	88,45	86,27	83,68
47	80,41	68,41	80,23	87,91	80,53	53,45	75,16
48	73,52	70,39	73,15	56,30	85,97	79,54	73,14
49	79,02	67,87	94,73	73,49	98,82	68,45	80,40
50	42,09	74,29	93,38	82,55	67,94	87,93	74,70
51	56,95	77,24	88,19	80,34	93,02	64,76	76,75
52	88,25	62,14	60,24	57,96	98,04	100,93	77,93
53	68,66	93,73	71,61	87,44	59,49	103,14	80,68
54	58,36	77,65	73,90	78,15	59,04	66,10	68,87
55	68,18	71,31	85,32	57,05	71,08	80,78	72,29
56	86,81	-45,72	63,87	83,42	63,19	62,23	52,30
57	59,41	58,66	61,96	77,43	80,35	83,61	70,24
58	76,81	90,73	105,86	90,95	107,35	64,77	89,41
59	107,38	83,44	97,96	61,97	69,71	80,66	83,52
60	66,23	68,77	75,95	78,40	68,78	84,73	73,81
61	79,05	79,36	93,13	81,09	77,63	78,96	81,54
62	96,96	84,38	86,36	79,35	81,08	79,51	84,61
63	72,95	78,67	64,01	91,33	88,83	62,25	76,34
64	99,38	95,47	94,11	105,73	64,96	68,49	88,02
65	92,05	83,70	89,58	62,04	66,82	71,80	77,66
66	76,22	97,42	66,41	49,30	73,66	77,70	73,45
67	87,28	75,30	75,63	90,87	82,92	81,45	82,24
68	79,44	86,11	72,14	60,56	88,05	83,80	78,35
69	93,89	75,67	67,56	70,89	70,22	85,44	77,28
70	75,56	54,76	48,34	78,82	98,92	61,93	69,72
71	85,81	66,69	65,20	79,30	110,75	103,19	85,16
72	76,19	72,04	68,77	84,38	89,62	71,89	77,15
73	84,85	87,58	82,10	69,79	61,31	90,65	79,38
74	65,58	71,28	79,53	62,24	66,39	73,20	69,70
75	69,66	88,74	73,56	95,23	83,39	76,13	81,12
76	98,01	72,74	104,34	82,59	59,87	83,98	83,59
77	90,25	120,07	85,04	78,50	92,67	86,28	92,13
78	63,61	67,41	103,77	60,96	96,84	64,13	76,12
79	93,65	71,55	101,19	39,32	70,41	98,62	79,12
80	83,55	97,44	81,89	62,57	77,44	79,97	80,48
81	51,68	69,09	80,33	85,88	76,87	68,17	72,01

82	89,44	85,84	88,05	64,48	65,56	70,96	77,39
83	69,60	47,48	72,66	90,62	82,87	64,45	71,28
84	95,83	64,33	80,94	60,17	85,08	69,57	75,99
85	84,82	69,88	93,98	80,94	88,83	81,31	83,29
86	67,47	49,78	71,89	70,35	59,67	98,26	69,57
87	76,81	78,03	94,81	73,56	58,63	58,25	73,35
88	68,63	82,67	83,85	96,13	72,38	90,37	82,34
89	64,44	58,87	77,56	91,06	88,21	85,33	77,58
90	96,83	99,76	81,56	109,21	76,27	86,91	91,76
91	93,42	69,45	91,88	76,86	59,09	75,55	77,71
92	51,54	74,71	91,06	61,80	87,10	79,83	74,34
93	54,86	91,87	86,33	72,77	57,23	115,44	79,75
94	93,93	61,67	107,48	77,39	106,05	80,55	87,85
95	83,23	92,03	102,76	64,25	102,14	76,98	86,90
96	74,43	80,26	81,49	59,24	83,41	78,52	76,23
97	91,20	62,80	74,30	74,11	63,40	82,41	74,70
98	86,16	79,33	67,52	68,01	96,76	102,31	83,35
99	98,41	85,72	83,95	90,20	90,65	81,23	88,36
100	58,11	77,72	61,09	64,80	63,26	68,48	65,57