

Desarrollo de una práctica de Tipos de Distribuciones Probabilísticas con GeoGebra

Salustia Teresa Cano Ibarra

Instituto Tecnológico de Celaya
terecaib@hotmail.com

Ma. Guadalupe Medina Torres

Instituto Tecnológico de Celaya
guadalupe.medina@itcelaya.edu.mx

Ma. Josefina Hernández Patiño

Instituto Tecnológico de Celaya

Aguirre Omar Mandujano

Instituto Tecnológico de Celaya

RESUMEN

El uso de las Tecnologías de información y comunicación (TIC) como apoyo para una mejor comprensión e interpretación de los temas de distribuciones probabilísticas se convierte en algo necesario y conveniente para lograr los objetivos académicos del modelo educativo del Instituto Tecnológico de Celaya basado en competencias.

Este trabajo presenta el desarrollo de una práctica relacionada con los temas de distribuciones probabilísticas discretas y continuas mediante el uso del software GeoGebra que desde la versión 3.2 incorpora una hoja de cálculo, con las mismas posibilidades que las vistas de otros software ampliamente conocidos como Excel y Calc, comandos estadísticos y gráficos y la vista de probabilidades. Es importante mencionar que se hace conveniente contar con los conocimientos teóricos básicos de

estadística descriptiva y distribuciones estadísticas para potencializar el aprendizaje didáctico del software.

PALABRAS CLAVES: Estadística, Distribuciones de Probabilidad Discreta, Distribuciones de Probabilidad Continua, Software, GeoGebra.

1. INTRODUCCIÓN

Una distribución de probabilidad indica toda la gama de valores que pueden representarse como resultado de un experimento si éste se llevase a cabo. Es decir, describe la probabilidad de que un evento se realice en el futuro, constituye una herramienta fundamental para la prospectiva, puesto que se puede diseñar un escenario de acontecimientos futuros considerando las tendencias actuales de diversos fenómenos naturales.

Toda distribución de probabilidad es generada por una variable (porque puede tomar diferentes valores) aleatoria x (porque el valor tomado es totalmente al azar), y puede ser de dos tipos:

VARIABLE ALEATORIA DISCRETA (x).

Porque solo puede tomar valores enteros y un número finito de ellos. Por ejemplo: X Variable que nos define el número de alumnos aprobados en la materia de probabilidad en un grupo de 40 alumnos (0, 1, 2, 3, ..., 40).

VARIABLE ALEATORIA CONTINUA (x).

Porque puede tomar tanto valores enteros como fraccionarios y un número infinito de ellos dentro de un mismo intervalo. Por ejemplo: x es la Variable que define la concentración en gramos de plata de algunas muestras de mineral (14.8, 12.1, 10.0, 42.3, 15.0, 18.4, 19.0, 21.0, 20.8, ..., n).

Los valores de una variable sirven para describir o clasificar individuos o distinguir entre ellos. La mayoría de las personas hacen algo más que simplemente describir, clasificar o distinguir, porque tienen ideas respecto a las frecuencias relativas de los valores de una variable. En Estadística se dice que la variable tiene una función de probabilidad, una función de densidad de probabilidad o simplemente una función de distribución (Badii & Castillo, 2007).

El software GeoGebra ofrece procedimientos usuales en el cálculo de probabilidades para un conjunto bastante amplio de funciones de distribución discretas y continuas que son habituales en el proceso de modelación, estas bondades y otras como ser un software gratuito, libre y de código abierto, multiplataforma, fácil, sencillo de usar y a la vez potente, han permitido que sea seleccionado por la vocalía de Probabilidad y Estadística del Departamento de Ciencias Básicas del Instituto Tecnológico de Celaya para diseñar las prácticas de las asignaturas de Probabilidad y Estadística que se ofrecen en todas las carreras. Es importante mencionar que se cuenta con un manual de Prácticas desarrollado en el programa Excel perteneciente a Microsoft Office. Sin embargo GeoGebra cuenta con una herramienta probabilística donde es posible manejar las distribuciones más frecuentes, así como calcular las medidas de tendencia central y de dispersión además de realizar diferentes test de bondad y ajuste, entre otras. (Del-Pino, 2013).

En el presente documento, se muestra el uso de este software para el desarrollo de una práctica utilizando distribuciones de probabilidad tanto discreta como continua, esperando con ello mostrar la facilidad de uso de esta herramienta y su dinamismo que contribuya al aprendizaje del alumno en estas asignaturas.

2. METODOLOGÍA

A continuación se presentan los pasos a seguir para desarrollar una práctica de distribuciones de probabilidad que se puede realizar dentro y fuera del aula de clases.

2.1 Procedimiento.

1. Para realizar el análisis de las distribuciones discretas y continuas en el GeoGebra, la herramienta a utilizar se identifica como **Cálculo de Probabilidades** y se selecciona del menú del botón ABC (Insertar texto) que aparece en la barra de herramientas principal. Ver Figura 2.1.

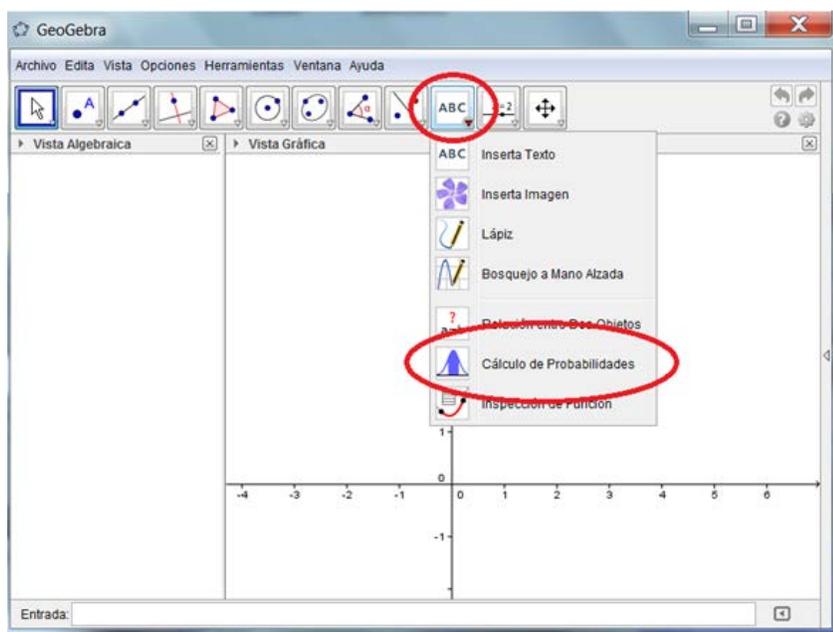


Figura. 2.1 Ventana para ubicar la herramienta cálculo de probabilidades

2. En la ventana **Cálculos de Probabilidad** seleccionar la Distribución de Probabilidad que se adecúe al problema. Dentro de las opciones se tienen en primer lugar las Distribuciones Continuas (Normal, t-Student, Chi Cuadrada, Distribución F, etc.) y posteriormente las Distribuciones Discretas (Binomial, Poisson e Hipergeométrica). Ver figura 2.2

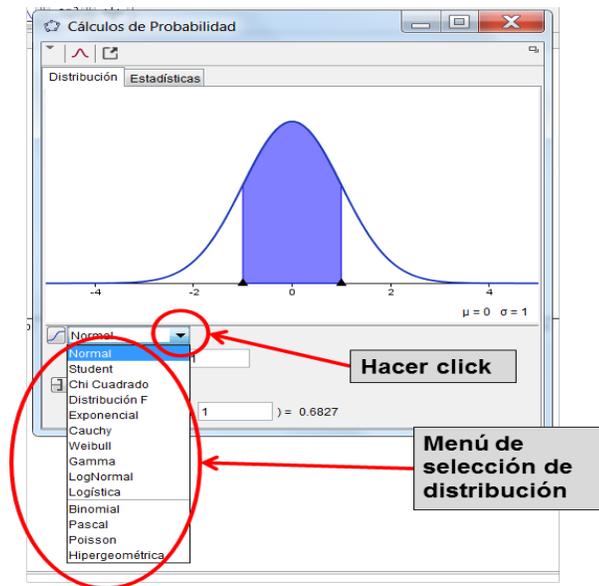


Figura. 2.2 Ventana para ubicar los tipos de distribuciones

DISTRIBUCIONES DISCRETAS

DISTRIBUCIÓN BINOMIAL

Los parámetros requeridos son: **tamaño de muestra (n)** y **probabilidad de éxito (p)**. La probabilidad se puede calcular en intervalos definidos en uno o ambos lados. En el caso de la probabilidad puntual, se calcula utilizando un intervalo cerrado dando el mismo valor en ambos lados de la desigualdad.

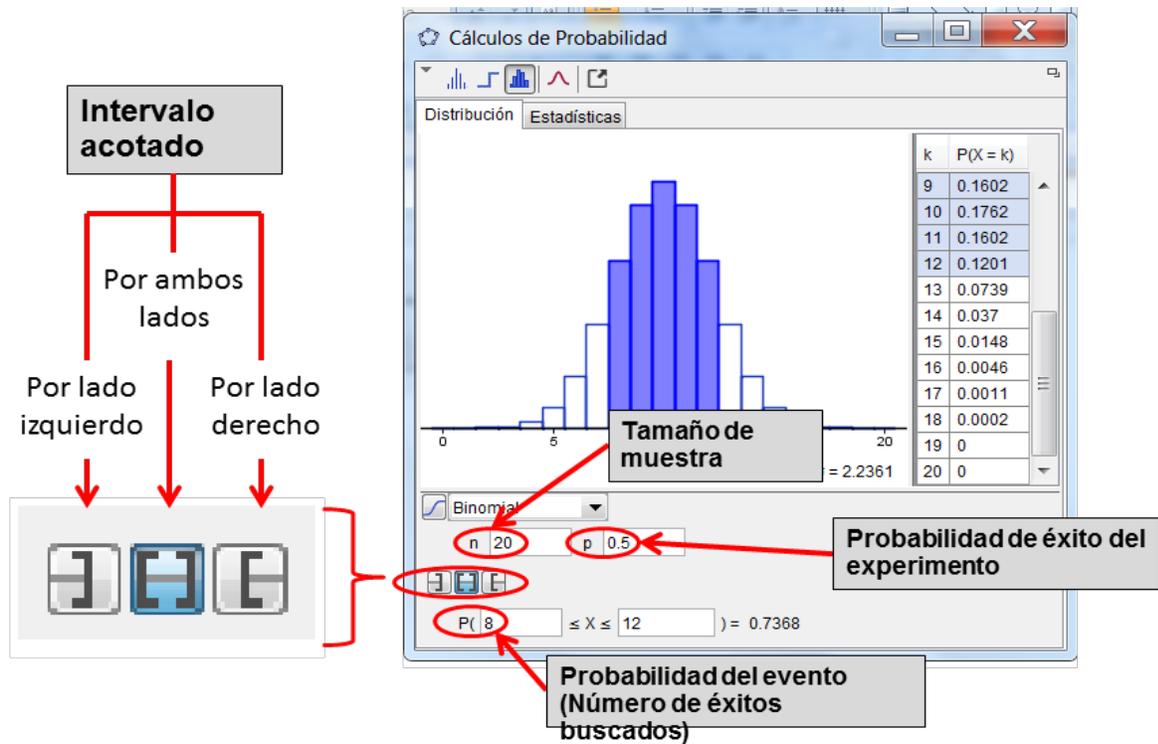


Figura 2.3 Ventana para identificar características de la distribución a utilizar

Como se puede observar en la Figura 2.3, del lado derecho de la gráfica aparece la tabla de distribución de probabilidad del experimento (la parte sombreada de la gráfica y de la tabla corresponde a la probabilidad del intervalo especificado por el usuario).

Ejemplo:

1. Para una muestra de 12 elementos, cuya probabilidad de éxito es de 0.4, calcular la probabilidad de encontrar:
 - a) Exactamente 5 elementos exitosos
 - b) Máximo 5 elementos exitosos, $P(x \leq 5)$
 - c) Más de 5 elementos exitosos, $P(x > 5)$
 - d) Entre 3 y 7 elementos éxitos, incluyendo estos valores, $P(3 \leq x \leq 7)$

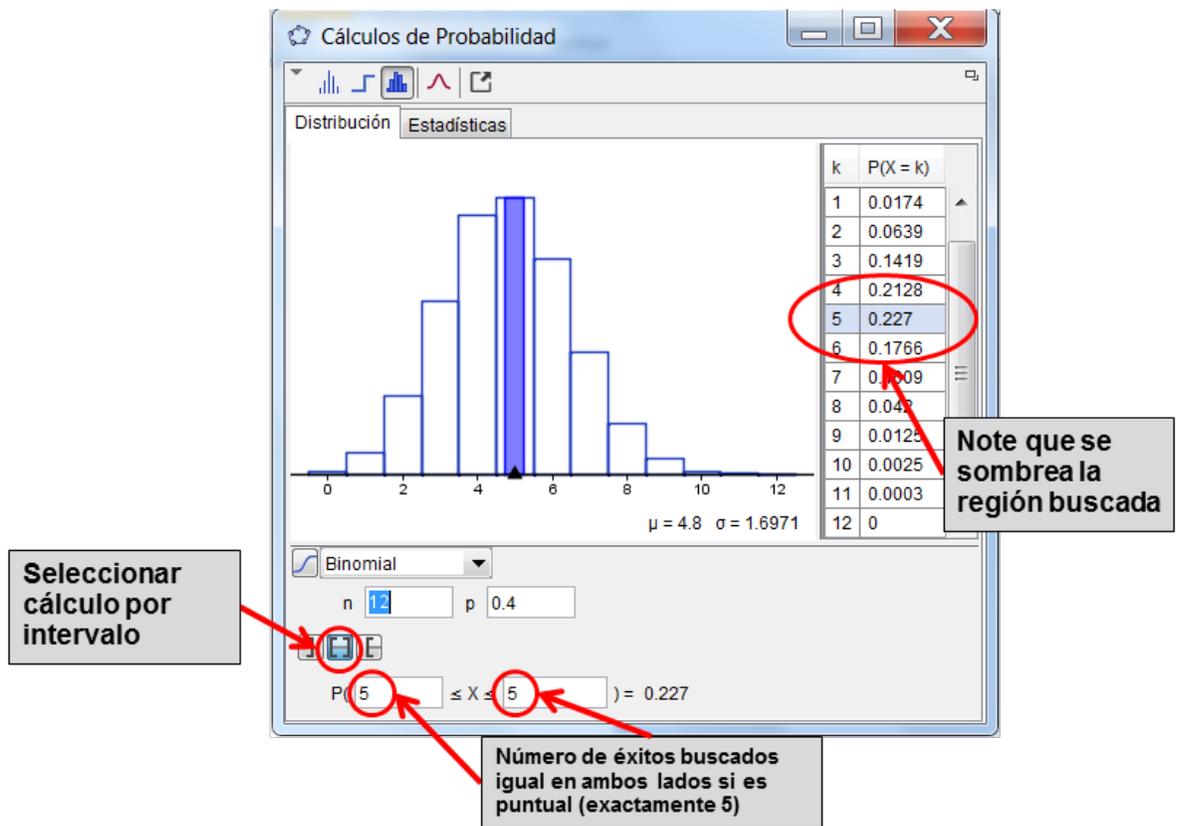


Figura 2.4 Ventana para verificar la probabilidad solicitada

Solución:

De acuerdo a las características del problema, el experimento sigue una distribución Binomial, por lo que el primer paso y que será válido para la solución de todos los incisos es **elegir la distribución Binomial** del menú de distribuciones de probabilidad. Introducir el valor de 12 en **n**, el valor de 0.4 en **p**.

- a) Seleccionar la **opción de intervalo cerrado**. Introducir el valor de 5 en ambos lado del intervalo. Con estos datos obtenemos la probabilidad de encontrar exactamente 5 elementos exitosos que es de 0.227, Figura 2.4.
- b) Seleccionar la **opción de intervalo abierto a la izquierda**. Introducir el valor de 5 en el lado derecho de la desigualdad. Con estos datos obtenemos la probabilidad de encontrar máximo 5 elementos exitosos que es de 0.6652,

- c) Seleccionar la **opción de intervalo abierto a la derecha**. Introducir el valor de 6 en el lado derecho de la desigualdad. Con estos datos obtenemos la probabilidad de encontrar más de 5 elementos exitosos que es de **0.3348**,
- d) Seleccionar la **opción de intervalo cerrado**. Introducir el valor de 3 en el lado izquierdo de la desigualdad y el valor de 7 en el lado derecho de la desigualdad. Con estos datos obtenemos la probabilidad de encontrar entre 3 y 7 elementos exitosos (incluyendo estos valores), que es de **0.5045**,

De manera similar, se procede en el cálculo de las diversas distribuciones que el Software ofrece.

DISTRIBUCIONES CONTINUAS

PROCEDIMIENTO:

1. Se genera una hoja de cálculo, se da clic en **Calculo de Probabilidad**

DISTRIBUCIÓN NORMAL

Los parámetros requeridos son: **media** (μ) y **desviación estándar** (σ). La probabilidad se puede calcular en intervalos definidos en uno o ambos lados. La parte sombreada de la gráfica corresponde a la probabilidad del intervalo especificado por el usuario. Ver figura 4.1.

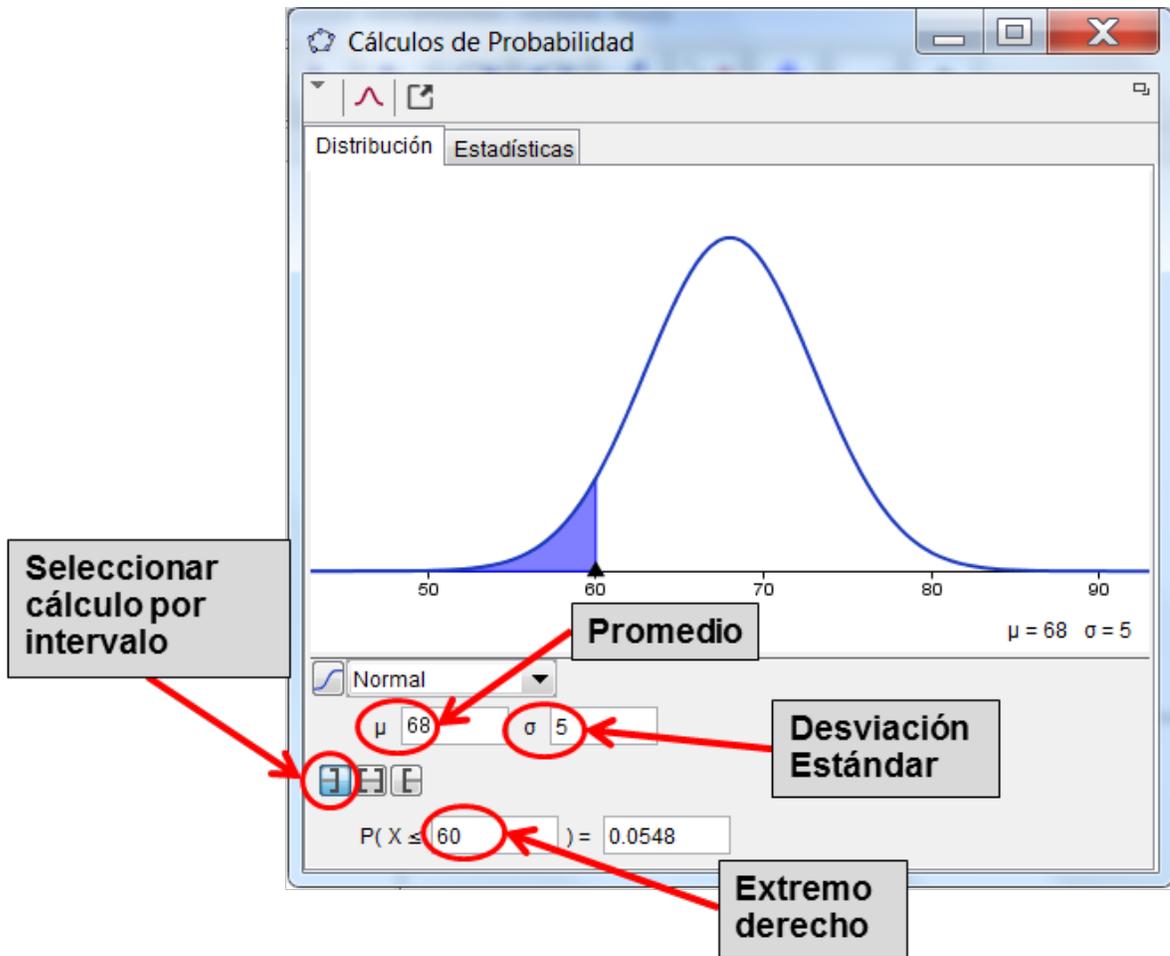


Figura 4.1 Ventana para ingresar los datos para el análisis de distribución normal

Ejemplo:

2. La distribución de edades de los habitantes de un país sigue una distribución normal con una vida media de 68 años y una varianza de 25 años. Se hace un estudio en una pequeña ciudad de 10000 habitantes.
 - a) ¿Cuántas personas se esperaba que vivieran menos de 60 años?

- b) ¿Cuántas personas se esperaría que vivieran más de 60?
- c) ¿Cuántas personas se esperaría que vivieran entre 40 y 50 años?
- d) ¿Por encima de que edad viviría el 10% más longevo?
- e) ¿Por debajo de que edad viviría el 10% de menor edad?
- f) ¿Entre qué edades se esperaría que viviera el 95% de la población de la ciudad?

Solución:

De acuerdo a la descripción del problema, el experimento sigue una distribución Normal, por lo que el primer paso y que será válido para la solución de todos los incisos es **elegir la distribución Normal** del menú de distribuciones de probabilidad. Introducir el valor de 68 en μ y el valor de 5 (raíz cuadrada de la varianza) en σ .

- a) Seleccionar la **opción de intervalo abierto a la izquierda**. Introducir el valor de 60 en el lado derecho de la desigualdad. Con estos datos obtenemos la probabilidad de encontrar personas que vivan menos de 60 años que es de **0.0548**,
- b) Seleccionar la **opción de intervalo abierto a la derecha**. Introducir el valor de 60 en el lado izquierdo de la desigualdad. Con estos datos obtenemos la probabilidad de encontrar personas que vivan más de 60 años, que es de **0.9452**,
- c) Seleccionar la **opción de intervalo cerrado**. Introducir el valor de 40 en el lado izquierdo de la desigualdad y el valor de 50 en el lado derecho de la desigualdad. Con estos datos obtenemos la probabilidad de encontrar personas que vivan entre 40 y 50 años, que es de **0.0002**,
- d) Seleccionar la **opción de intervalo abierto a la derecha**. Introducir el valor de 0.1 en el lado derecho de la igualdad. Con estos datos obtenemos la edad esperada sobre la cual vive el 10% más longevo de la población, que es de **74.4 años**,

- e) Seleccionar la **opción de intervalo abierto a la izquierda**. Introducir el valor de 0.1 en el lado derecho de la igualdad. Con estos datos obtenemos la edad esperada hasta la cual viviría el 10% de la población que fallece a edad más temprana, que es de **61 años**.

Los comandos de GeoGebra permiten resolver problemas que involucren a la distribución normal bajo cuatro opciones diferentes:

- Obtención de probabilidades bajo la curva, partiendo de datos originales (“X”, μ del problema, σ del problema),
- Obtención de probabilidades bajo la curva, partiendo de datos estandarizados (“Z”, $\mu=0$, $\sigma=1$), Obtención del inverso de la distribución normal acumulada partiendo de datos originales,
- Obtención del inverso de la distribución normal acumulada, partiendo de datos de originales (Probabilidad, μ del problema, σ del problema)
- Obtención del inverso de la distribución normal acumulada, partiendo de datos de probabilidad

OPCION 1: OBTENCION DE PROBABILIDADES BAJO LA CURVA, PARTIENDO DE DATOS ORIGINALES (“X”, μ del problema, σ del problema), en este caso, el procedimiento es el que sigue:

1. Se genera una hoja de cálculo y gráficos
2. Después se dará clic en cálculo de probabilidades
3. Se abrirá una ventana con ciertos comandos en los cuales insertaremos los datos de nuestros problemas poniendo en la parte de distribución la opción de normal. Figura 4.2

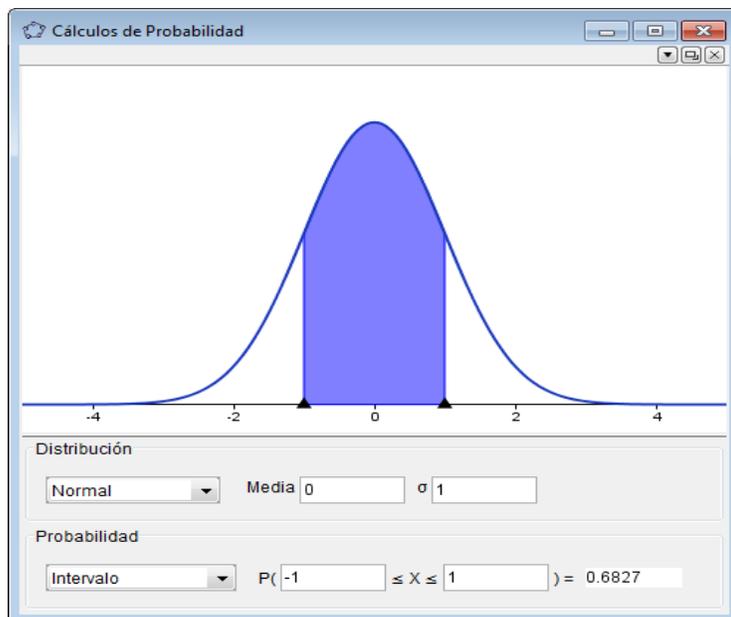


Figura 4.2 ventana para insertar los datos de la distribución normal

Ejemplo 1: La vida media de los habitantes de un país es de 68 años, con una varianza de 25. Se hace un estudio en una pequeña ciudad de 10000 habitantes:

a) ¿Cuántos vivirán menos de 60 años?

Solución usando los comandos de GeoGebra:

A continuación se muestra la información con la cual debes alimentar en la ventana de dialogo:

Distribución: **Normal** (Ya que es la distribución que estamos manejando)

Media: **68** (Media aritmética de la distribución)

σ : 5 (desviación estándar de la población), ver figura 4.3.

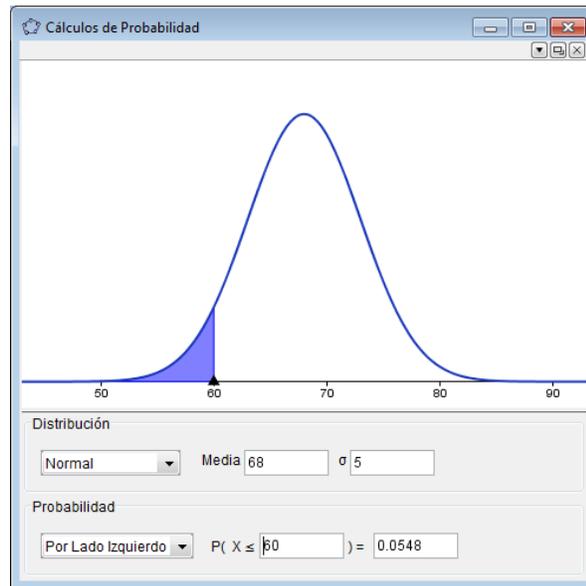


Figura 4.3 ventana para ingresar los datos del ejemplo 1

En la tabla de cálculo de probabilidades quedaría ahí dando como resultado: **0.0548**

OPCION 2: OBTENCION DE PROBABILIDADES BAJO LA CURVA, PARTIENDO DE DATOS ESTANDARIZADOS ("Z", $\mu=0$, $\sigma=1$).

Realizar los tres primeros pasos de la opción 1, la diferencia será en los intervalos.

Ejemplo 2: dada una distribución normal estándar, encuentre el área bajo la curva que está entre $z=-1.91$ y $z=0.86$. Figura 4.4

Esta es nuestra información:

Z: 0.86 (valor estandarizado cuya probabilidad acumulada se desea obtener, para el ejercicio sería $P(-\infty < z < 0.86)$).

Z: -1.97 (valor estandarizado cuya probabilidad acumulada se desea obtener, para el ejercicio sería $P(-\infty < z < -1.97)$).

σ : 5 (desviación estándar).

Intervalo: (por lado izquierdo, proporciona la probabilidad del “ $-\infty$ ” hasta el valor elegido de “ x ”; por el lado derecho, proporciona la probabilidad de $x \pm 5$, útil para aproximar la distribución binomial a la normal, en este caso lo correcto es por el lado izquierdo.

Por lo tanto el área bajo la curva en este intervalo es: 0.7807

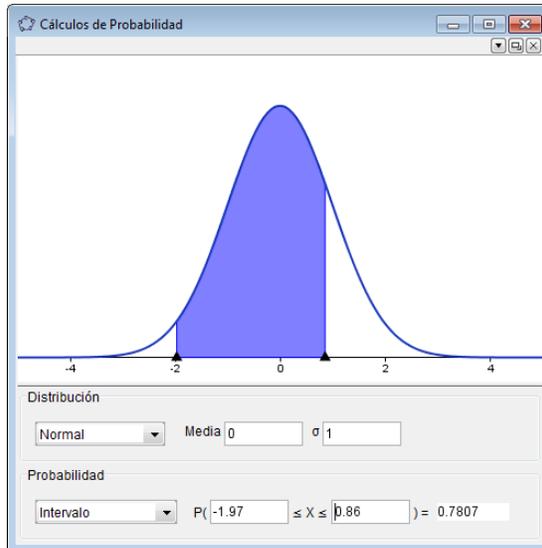


Figura 4.4 ventana de dialogo para el ejemplo 2

OPCION 3. OBTENCION DEL INVERSO DE LA DISTRIBUCION NORMAL ACUMULADA, PARTIENDO DE DATOS ORIGINALES (Probabilidad, μ del problema, σ del problema); el resultado que se obtiene es “ X ”. El procedimiento es el que sigue obteniéndose con respecto al área de la izquierda acumulada:

Ejemplo 3: El consumo medio anual de cerveza de los habitantes de un país es de 58 litros por persona, con una varianza de 36 litros por persona. Se supone que se distribuyen según una distribución normal. Figura 4.5

- a) Si usted presume de ser un buen bebedor, ¿Cuántos litros de cerveza tendría que beber al año para pertenecer al 5% de la población que más bebe? A

continuación se muestra la información con la cual se debe llenar la ventana de dialogo de GeoGebra:

Probabilidad acumulada:	0.95
Media aritmética de la población:	58
Desviación estándar:	6

Los datos ya conocidos se colocan de la misma forma, para determinar el valor de X en este caso se toma el área acumulada de la izquierda y en la parte donde hay un espacio enseguida de un igual es el espacio de la probabilidad donde colocaremos el valor de 0.95 dándonos el resultado final.

El valor obtenido correspondiente al valor de "X". Por lo tanto, tendría usted que beber más de 67,87 litros al año para pertenecer a ese "selecto" club de grandes bebedores de cerveza.

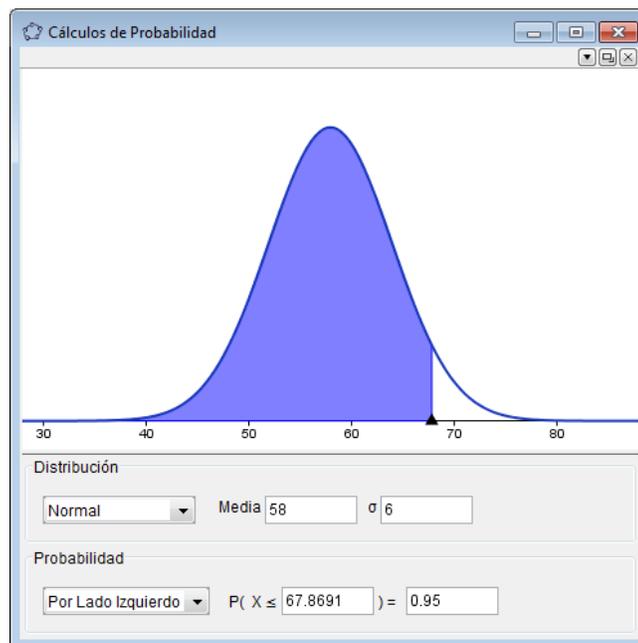


Figura 4.5 Muestra el área de la distribución acumulada y los valores calculados del ejemplo 3

OPCION 4. OBTENCION DEL INVERSO DE LA DISTRIBUCION NORMAL ACUMULADA, PARTIENDO EXCLUSIVAMENTE DE DATOS DE PROBABILIDAD. El resultado que se obtiene es de z, el procedimiento es el que sigue:

Ejemplo 4: Dada una distribución estándar normal, encuentre el valor de k, tal que $P(z < k) = 0.95$. Figura 4.6

A continuación se muestra la información con la cual debes llenar la ventana de dialogo de GeoGebra:

Probabilidad 0.95 (Probabilidad acumulada)

Una vez completada la información con la cual debes alimentar la ventana de dialogo:

Solamente hay que colocar el valor de la probabilidad en el espacio que es de esta variable y después de colocar los datos nos da el resultado de 1.6449

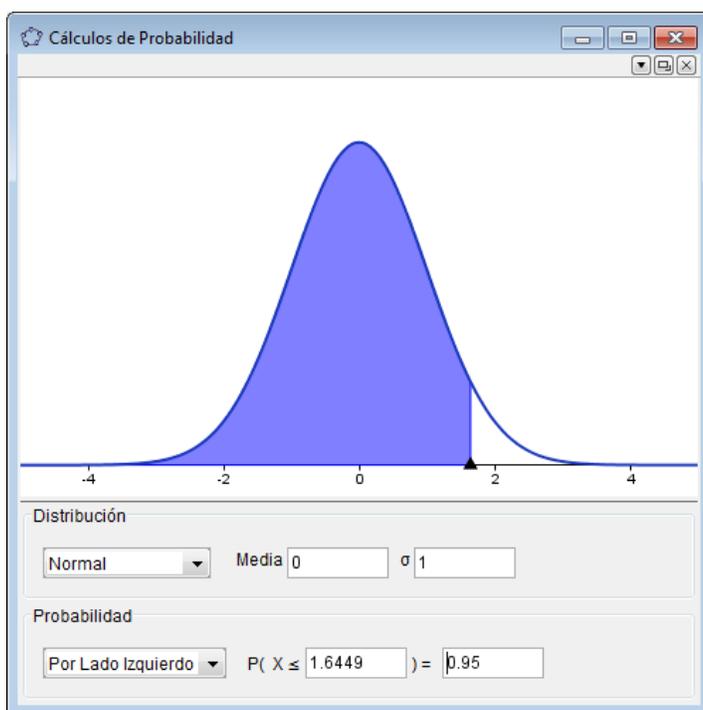


Figura 4.6 Muestra el área de la distribución acumulada y los valores solicitados del ejemplo 4

2.2 Protocolo de construcción

Nº	Nombre	Definición	Valor	Subtítulo
1	Punto A	Punto sobre EjeX	A = (6, 0)	
2	Punto B	Punto sobre EjeX	B = (12, 0)	
3	Número a	Barras[[0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12], [0, 0.02, 0.06, 0.14, 0.21, 0.23, 0.18, 0.1, 0.04, 0.01, 0, 0, 0], 1]	a = 1	
4	Lista lista1	Extrae[[0, 0.02, 0.06, 0.14, 0.21, 0.23, 0.18, 0.1, 0.04, 0.01, 0, 0, 0], x(A) + 1.5, x(B) + 1.5]	lista1 = {0.18, 0.1, 0.04, 0.01, 0, 0, 0}	
5	Lista lista2	Extrae[[0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12], x(A) + 1.5, x(B) + 1.5]	lista2 = {6, 7, 8, 9, 10, 11, 12}	
6	Número b	Barras[[lista2, lista1, 1]	b = 0.33	
7	Punto C	Punto sobre EjeX	C = (42, 0)	
8	Punto D	Punto sobre EjeX	D = (61.59, 0)	
9	Función f		$f(x) = e^x \cdot (-(x - 68)^2 / 50) / (\text{abs}(5) \cdot \text{sqrt}(\pi \cdot 2))$	
10	Número c	Integral de f desde x(C) a x(D)	c = 0.1	

Figura 4.7 Muestra el protocolo de construcción

3. RESULTADOS Y CONCLUSIONES.

En búsqueda de enriquecer el desarrollo de los temas abordados en las asignaturas relacionadas con probabilidad y estadística con el uso de tecnologías de información y comunicación (TIC), la vocalía del Departamento de Ciencias Básicas del ITC, establece un manual de prácticas basadas en el uso del software GeoGebra cuyas bondades son variadas y que permite plantear prácticas para cada unidad de los programas de estudio de la materia de probabilidad.

Se espera que con la implementación de estas prácticas los alumnos de Ingeniería del Instituto desarrollen habilidades tecnológicas y cognitivas que contribuyan a aprender de su propia experiencia y visualizar mejor los conceptos estadísticos.

En el departamento se cuenta con un manual de prácticas desarrollado a través del software de Microsoft Excel, que ha sido de una gran ayuda pues permite un dinamismo a la clase y un gran apoyo tanto para el maestro como a los alumnos, sin embargo, se limita a generar datos y no permite visualizar el comportamiento de los mismos en gráficos y tomar mejores decisiones al resolver un problema dado.

Se tiene la confianza de que con el uso de GeoGebra se tendrá mayor facilidad para realizar los cálculos estadísticos en el análisis e interpretación de los ejercicios expuestos y que serán el complemento de la parte de Estadística descriptiva que se presenta en los primeros capítulos del manual y que es objeto de otra presentación.

BIBLIOGRAFÍA

- [1] Badii, M,H & J. Castillo. (2007) Técnicas Cuantitativas en la investigación UANL, Monterrey.
- [2] Campos, C. A. (2008). Aprendizaje de la estadística a través de casos prácticos. II Jornadas de Innovación Docente, Tecnologías de la información y de la

comunicación e investigación educativa en la universidad de zaragoza 2008.
Zaragoza.

- [3] Del-Pino, J. (2013). El uso de Geogebra como herramientas para el aprendizaje de las medidas de dispersión. *Actas de las Jornadas Virtuales en Didáctica de la Estadística, Probabilidad y Combinatoria*, (pp. 243-250). Granada.
- [4] Gámez Mellado, A., & Marín Trechera, L. M. (2010). Distribuciones estadísticas: Un ejemplo de uso de GeoGebra en enseñanza universitaria. *Epsilon74* , 33-42.
- [5] Gámez Mellado, A., & Marín Trechera, L. M. (2011). Taller de “Diseño de Recursos Didácticos Interactivos para la Enseñanza y el Aprendizaje de la Estadística utilizando GeoGebra. I I Jornadas de Enseñanza y Aprendizaje de la Estadística e Investigación Operativa, (pp. 73-92). Las palmas de Gran Canaria.
- [6] Nickerson, R. S. (1995). Can technology help teach for understanding. In eds. D. N. Perkins, J.L. Schwartz, M. M. West, and M. S. Wiske, *Software goes to school: teaching for understanding with new technologies*, New York: Oxford University Press.
- [7] Rocha Salamanca, P., & Gallego Torres, A. P. (2009). Las prácticas docentes de los profesores de probabilidad y estadística en las Facultades de Ingeniería. 4° Congreso Internacional sobre Formación de Profesores de Ciencias (pp. 197-202). Bogota, Colombia: Tecné, Episteme y Didaxis.