



IX CIEMAC
Congreso Internacional
sobre la Enseñanza de la
Matemática Asistida por Computadora
www.cidse.tec.ac.cr/ciemac

TEC | Tecnológico
de Costa Rica

Estadística descriptiva e inferencial con GeoGebra

Bach. Luis Andrés Ortiz Hernández
UNED, Costa Rica
hoy200319234@gmail.com

Lic. Ronald Andrés Arias Madriz
UCR, Costa Rica
ramartec@gmail.com

Resumen: El presente taller consiste en utilizar el software GeoGebra para resolver problemas de estadística dejando al programa los cálculos, elaboración de tablas y gráficos propios de la estadística y así profundizar en los contenidos y conceptos estocásticos.

Palabras clave: Estadística, GeoGebra, Software libre.

1. Introducción

El GeoGebra es un software de matemática dinámica y estadística que ofrece la facilidad de trabajar con cantidades considerables de datos, estudiar sus comportamientos y representarlos por medio de medidas de tendencia central o en forma de gráficos estadísticos, además posee una calculadora de probabilidades que junto a lo anterior facilita el trabajo estadístico, ahorrando tiempo que puede emplearse en la profundización y apropiación de conceptos propios de la estadística tanto descriptiva como inferencial. Este material fue desarrollado como parte de las evaluaciones del curso Probabilidad y Estadística de la carrera Enseñanza de la Matemática de la Universidad Estatal a Distancia.

2. Aspectos generales

Publico meta: Universitario y profesores de secundaria

Requerimientos del taller: Laboratorio de computadoras con el GeoGebra 5 (o superior) instalado y conexión a Internet además de proyector multimedia.

Conocimientos previos: Nivel básico en el uso de computadoras y conocimiento estadístico.

Objetivo del taller: Utilizar el GeoGebra en la resolución y modelación de problemas de estadística.

Metodología del taller: El taller iniciará con una breve descripción de los contenidos y objetivos, así como una introducción al GeoGebra, luego se trabajará con ejemplos a resolver paso a paso, finalizando con ejercicios propuestos donde se pueda profundizar tanto en el uso del programa como en los contenidos estadísticos.

3. Guías de trabajo

GeoGebra.

Para esta actividad se empleará el software GeoGebra el cual puede descargar de la siguiente dirección:

<https://www.geogebra.org/download>

En dicha página web puede descargar el instalador que requiera según su sistema operativo, y luego debe dar “doble clic” para ejecutar el instalador y seguir los pasos que este le indicará. Una vez instalado el GeoGebra, podrá ejecutarlo desde el Escritorio.

El entorno inicial y con el que se trabajará es el siguiente:

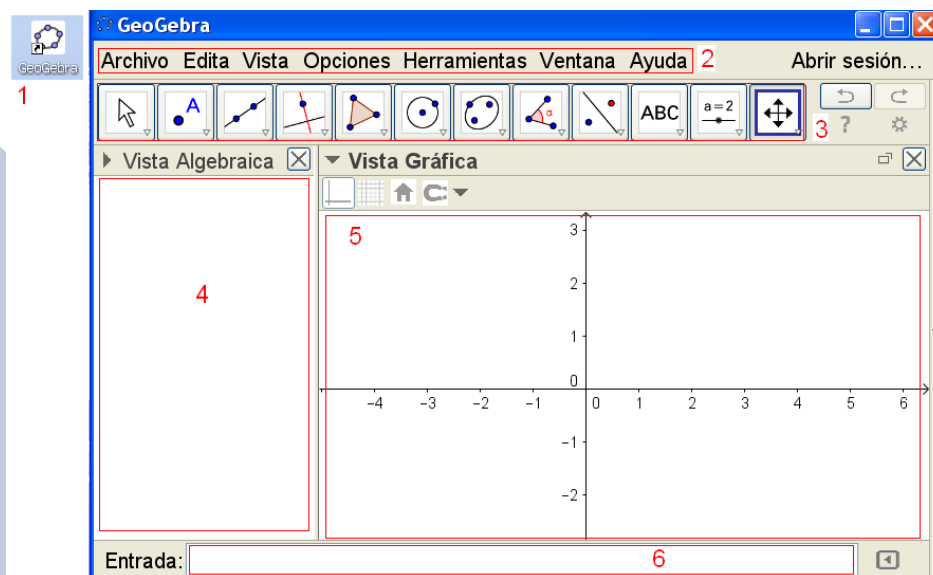


Imagen 1: Pantalla inicial GeoGebra.

1. Es el icono para iniciar el programa.
2. Barra de menús (Archivo y Vista serán los utilizados).
3. Barra de herramientas (cada herramienta en la esquina inferior derecha tiene una flecha que despliega otras herramientas).
4. Vista algebraica, contendrá los resultados numéricos y representaciones algebraicas.
5. Vista gráfica, contendrá los objetos geométricos.
6. Entrada que permite escribir distintos comandos e ingresar funciones.

Actividad 1: Medidas de posición: para calcular las distintas medidas primero debe ingresar al menú **Vista, Hoja de Cálculo**. El entorno es muy similar al de Excel. Al tener activa la Hoja de Cálculo, las demás vistas se mantienen visibles, además puede notar que las herramientas cambian a unas propias de la estadística.

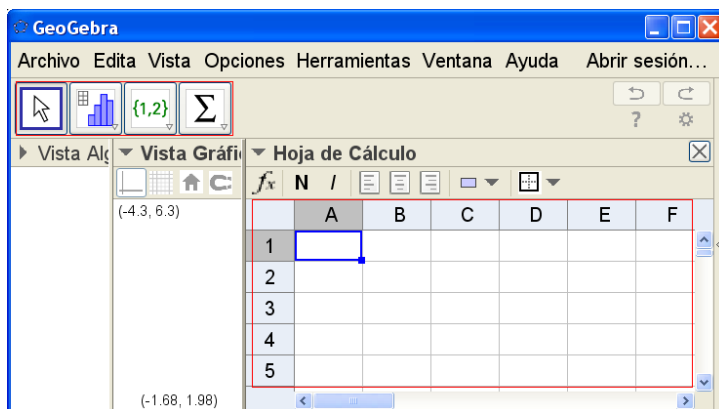


Imagen 2: Hoja de Cálculo y herramientas estadísticas.

Ejemplo 1: Calcular la media, mediana, varianza, desviación estándar, moda, valor máximo y mínimo, así como el primer y tercer cuartil, de los siguientes datos:

2.36	9.25	1.20
4.25	8.52	5.36
4.65	3.00	7.63

Para lo anterior, trabaje en la **Hoja de Cálculo** y copie los datos en la misma, el orden no importa ni la forma de ingresarlo ya que puede ser una sola fila o columna o en forma de tabla. Luego manteniendo presionado el clic izquierdo del ratón, arrastre el puntero hasta abarcar todos los datos; una vez hecho esto, los datos quedarán seleccionados. Sobre dicha selección haga clic derecho y entre las opciones del menú emergente escoja: **Crea, Lista**. En la vista algebraica debe aparecer una lista con los datos.

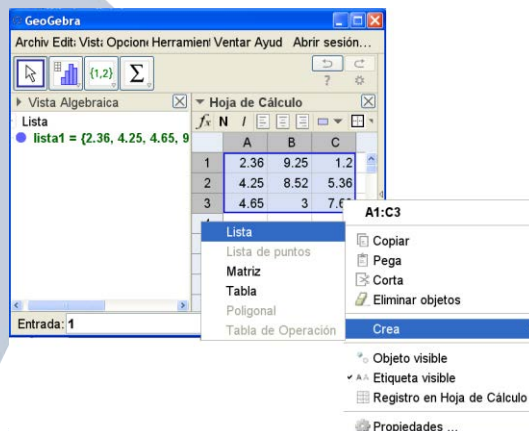


Imagen 3: Datos y creación de lista.

Finalmente en la **Entrada** se digitan los siguientes comandos para cada una de las tareas a realizar:

Comando	Comando
Media[<Lista de datos brutos>]	Mediana[<Lista de números>]
VarianzaMuestral[<Lista de datos brutos>]	Moda[<Lista de números>] (Genera una lista)
Máximo[<Lista>]	Mínimo[<Lista>]
Q1[<Lista de datos brutos>] (Primer cuartil)	Q3[<Lista de datos brutos>] (Tercer cuartil)

Note que los comandos solicitan una lista por tanto, entre los paréntesis cuadrados se debe escribir el nombre de la lista creada anteriormente, por ejemplo: el resultado de *Media[lista1]* aparecerá en la **Vista Algebraica**. Para la Desviación Estándar puede usar el comando **sqrt(<x>)** para determinar la raíz cuadrada de la Varianza.

Renombre cada resultado según la información que le corresponda para que el nombre de cada dato sea significativo.

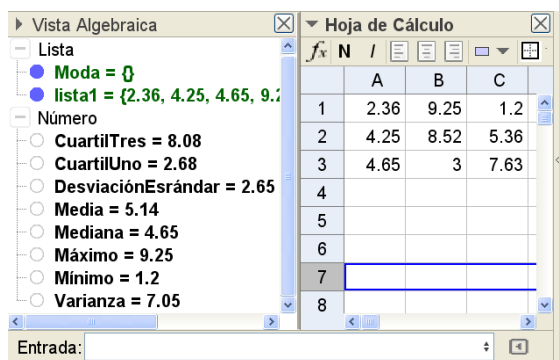


Imagen 4: Valores del ejemplo.

Ejercicio 1: Calcular la media, mediana, varianza, desviación estándar, moda, valor máximo y mínimo, así como el primer y tercer cuartil, de los siguientes datos:

9.8	9.2	8.2	7.1	8.1
7.9	9.1	9.5	8.9	0.1

Además:

- Responda ¿La media y/o la mediana brindan una idea correcta de dónde está el centro de los datos? ¿Por qué? ¿Cuál es más significativa?
- Modifique los valores de la hoja de cálculo para que exista moda, luego modifique de nuevo para que la moda no sea única.

Actividad 2: Representación gráfica de datos: dichas representaciones se realizan con ayuda de la herramienta **Análisis de una variable** (esta herramienta aparece cuando está activa la **Hoja de Cálculo**).

Ejemplo 2: Recuerde los datos del primer ejemplo y realice los gráficos de tallo-hoja, una gráfica de puntos y diagrama de cajas.

Para lo anterior seleccione con ayuda del puntero los datos, una vez seleccionado escoja la herramienta **Análisis de una variable**. En la ventana emergente se puede modificar los datos seleccionados. Si la selección es la correcta, presione el botón **Analiza**. Por defecto aparecerá el histograma.

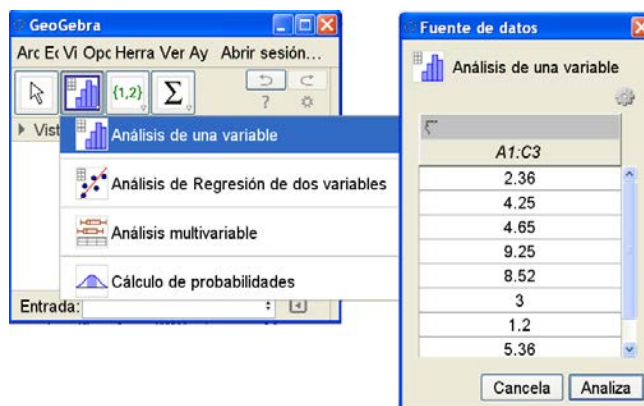


Imagen 5: Selección de herramienta y pantalla de fuente de datos.

La ventana del histograma en su parte superior muestra un menú desplegable donde puede seleccionar los distintos diagramas o gráficos requeridos.

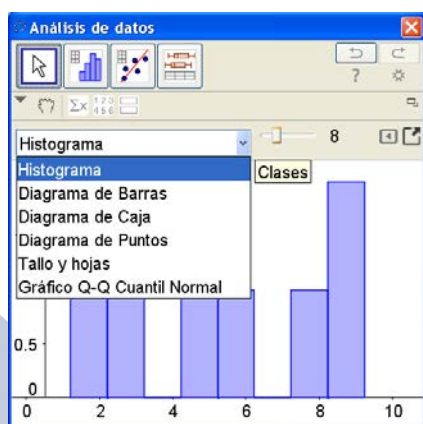


Imagen 6: Escoger entre distintos gráficos.

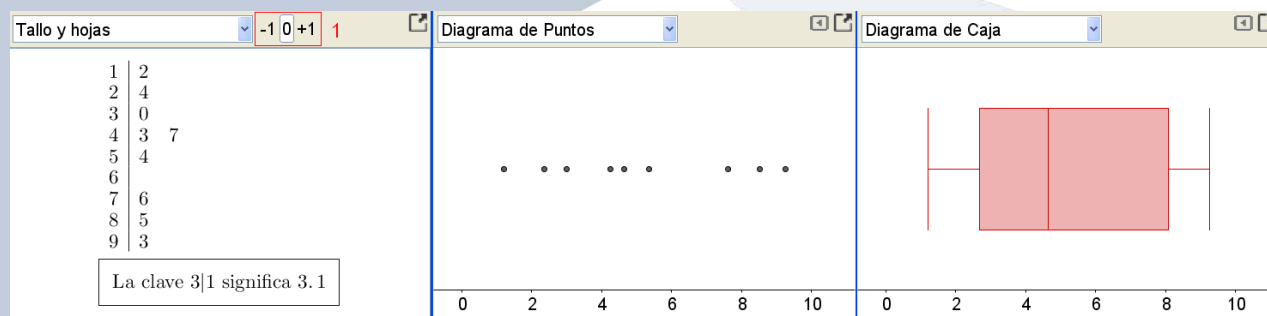


Imagen 7: Gráficos solicitados (1: Permite modificar la forma del tallo y hoja).

Ejercicio 2: Continúe el **ejercicio 1** construyendo; los gráficos de tallo-hoja, una gráfica de puntos y diagrama de cajas.

Además:

- Modifique los valores de la tabla del ejercicio anterior con valores en distintos intervalos como: [100-150], [10 000, 10 500] y [12.0 y 15.0]. Construya el diagrama de tallo-hoja en cada caso y describa qué se toma por tallo y qué por hoja.

Actividad 3: Un poco de historia:

“El duque de Toscana le preguntó un día a Galileo: ¿por qué cuando se lanzan 3 dados se obtiene más veces la suma 10 que la suma 9, aunque se obtenga de 6 maneras diferentes cada una?”

Ejercicio 3: Descargue el y abra en el GeoGebra siguiente archivo que simula la situación anterior:

<https://dl.dropboxusercontent.com/u/48466722/dados.ggb> (problema y archivo tomado de: <http://docentes.educacion.navarra.es/msadaall/geogebra/azar.htm>)

Con ayuda de la simulación, concluya si lo observado por el duque es verdadero o es falso y responda a la pregunta realizada por el mismo.

Además:

Justifique sus respuestas definiendo el espacio muestral y calculando las probabilidades de ambos casos ($P(10)$ y $P(9)$).

Actividad 4: Distribuciones continuas de probabilidad: dichas funciones se representan generalmente mediante funciones reales a trozos, por tanto podemos graficarlas con ayuda del GeoGebra y el comando **Si[<Condición>, <Entonces>]**, que se escribe en la barra de **Entrada**.

Ejemplo 3: El número total de horas, medidas en unidades de 100 horas, que una familia utiliza la aspiradora en un periodo de un año es una variable aleatoria continua X que tiene la función de densidad:

$$f(x) = \begin{cases} x, & 0 < x < 1 \\ 2 - x, & 1 \leq x < 2 \\ 0, & \text{en cualquier otro caso} \end{cases}$$

Encuentre la probabilidad de que en un periodo de un año, una familia utilice la aspiradora:

- a) Menos de 120 horas
- b) Entre 50 y 100 horas
- c) Exactamente 150 horas

Para resolver el ejemplo anterior escriba en la **Entrada** el siguiente comando que define la función a trozos requerida: **Si[$0 < x < 1, x$, Si[$1 \leq x < 2, 2 - x, 0$]]**

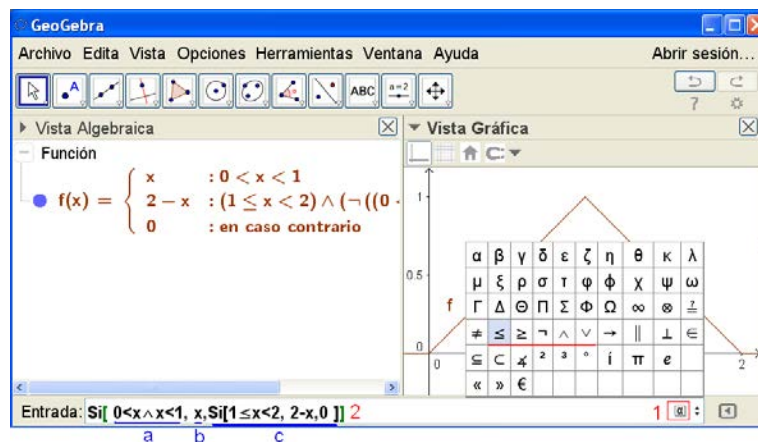


Imagen 8: Gráfica de función, comando y caracteres especiales.

Nota:

1. El símbolo del alfa (α) aparece al escribir en el Entrada y permite ingresar a una tabla con caracteres especiales.
2. El comando se explica brevemente de la siguiente manera (ver entrada algebraica de la imagen anterior):

- a. Primera condición.
- b. Lo que se grafica en caso de cumplir la primer condición.
- c. Lo que se grafica en caso de no cumplir la primer condición, note que se usa otro comando: **Si[<Condición>, <Entonces>]** anidado.

Lista la función se calculan las probabilidades deseadas, recuerde que la probabilidad es un área bajo la curva es decir se usa el comando:

Integral[<Función>, <Extremo inferior del intervalo>, <Extremo superior del intervalo>]

- a) Menos de 120 horas

Es decir $P(X < 1.2) = \int_0^{1.2} f(x) dx$ y se calcula con ayuda del comando: **Integral[f,0,1.2]**

- b) Entre 50 y 100 horas

Es decir $P(0.5 < X < 1) = \int_{0.5}^1 f(x) dx$ y se calcula con ayuda del comando: **Integral[f,0.5,1]**

- c) Exactamente 150 horas

Es decir $P(X = 1.5) = f(1.5)$ y se calcula con ayuda del comando: **f(1.5)**

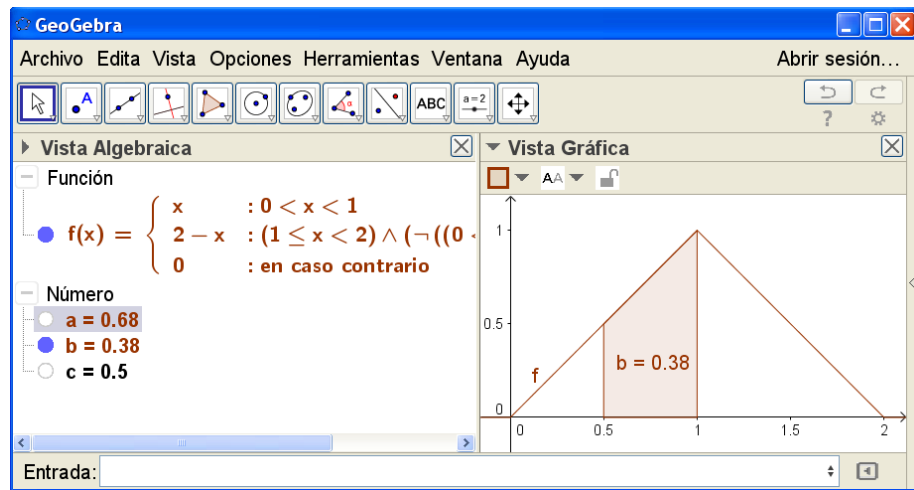


Imagen 9: Respuestas y área bajo curva de la pregunta b).

Ejercicio 4: La vida útil, en días, para frascos de cierta medicina de prescripción es una variable aleatoria que tiene la función de densidad:

$$f(x) = \begin{cases} \frac{20000}{(x+100)^3}, & x > 0 \\ 0, & \text{en cualquier otro caso} \end{cases}$$

Encuentre la probabilidad de que un frasco de esta medicina tenga una vida útil de:

- a) Al menos 200 días
- b) Cualquier lapso entre 80 y 120 días.

Actividad 5: Esperanza matemática: se puede emplear la **Hoja de Calculo** para construir tablas y calcular la esperanza matemática.

Ejemplo 4: La distribución de probabilidad de X, con X el número de imperfecciones por cada 10 metros de una tela sintética, está dada por:

x	0	1	2	3	4
f(x)	0.41	0.37	0.16	0.05	0.01

Encuentre el valor esperado de imperfecciones.

Para resolver lo anterior se construye la tabla en la **Hoja de Calculo**:

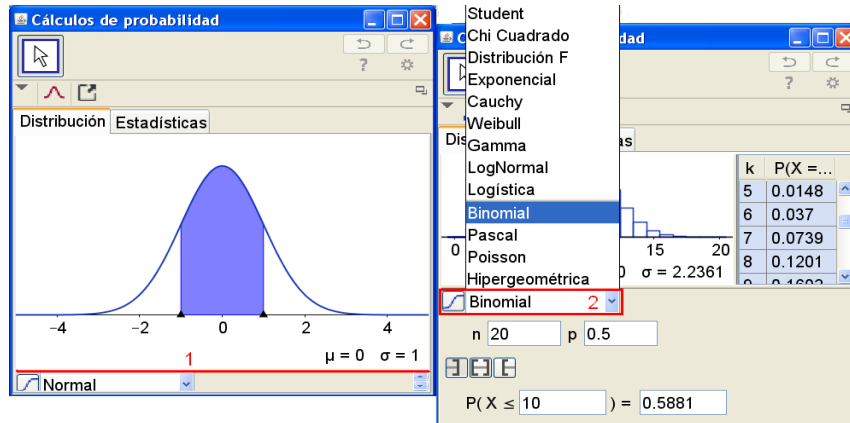


Imagen 13. Muestras de la calculadora de probabilidades.

1. Al iniciar puede que la parte inferior no este completa, coloque el puntero sobre la línea marcada en rojo y haciendo clic izquierdo, arrastre hacia arriba para aumentar esa zona como se muestra en la figura de la derecha.
2. En el menú desplegable seleccione **Binomial**.

Ejemplo 5: De acuerdo a una empresa, aproximadamente el 30% de todas las fallas de operación en las tuberías de plantas químicas son ocasionadas por errores del operador.

¿Cuál es la probabilidad de que de las siguientes 20 fallas en las tuberías al menos 10 se deban a un error del operador?

¿Cuál es la probabilidad de que no más de 4 de 20 fallas se deban al error del operador?

Para resolver lo anterior se utiliza la distribución binomial de la calculadora de probabilidades como se mostró anteriormente. Ingresando los datos dados y presionando “**enter**”.

¿Cuál es la probabilidad de que de las siguientes 20 fallas en las tuberías al menos 10 se deban a un error del operador?

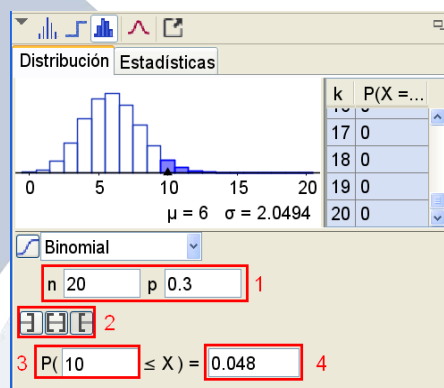



Imagen 14. Datos y resultados de la primera pregunta.

1. Datos del ejemplo.

2. Botones para determinar el intervalo deseado, en este caso () por ser área a la derecha.
3. Dato del ejercicio.
4. Resultado del ejercicio.

¿Cuál es la probabilidad de que no más de 4 de 20 fallas se deban al error del operador?

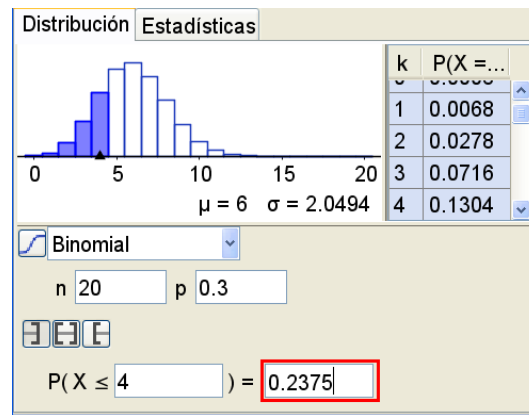


Imagen 15. Datos y resultados de la segunda pregunta.

Ejercicio 6: Al probar cierta clase de neumáticos, se encuentra que 25% presentaban pinchaduras. De los siguientes 15 neumáticos, encuentre con ayuda del programa la probabilidad de que:

- a) de 3 a 6 tengan pinchaduras.
- b) menos de 4 tengan pinchaduras.
- c) más de 5 tengan pinchaduras.

Actividad 7: Distribución normal: El GeoGebra cuenta con una calculadora de probabilidades, puede acceder a la misma mediante el menú **Vista, Cálculos de Probabilidad**.

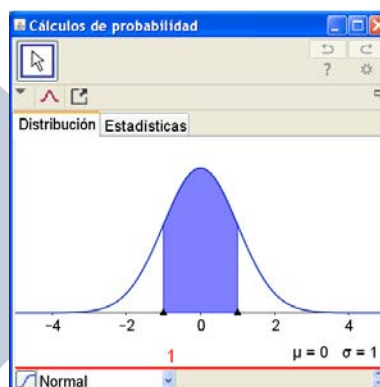


Imagen 16. Muestras de la calculadora de probabilidades.

3. Al iniciar puede que la parte inferior no esté completa, coloque el puntero sobre la línea marcada en rojo y aumente esa zona tal y como se realizó en la actividad 6.

Ejemplo 6: Una máquina expendedora de bebidas gaseosas se regula para que sirva un promedio de 200 ml por vaso. Si la cantidad de bebida se distribuye normalmente con una desviación estándar igual a 15 ml,

- ¿qué fracción de los vasos contendrá más de 224 ml?
- ¿cuál es la probabilidad de que un vaso contenga entre 191 y 209 ml?

Para resolver el ejemplo anterior, en la **calculadora de probabilidades** se ingresan los datos $\mu = 200$ y $\sigma = 15$:

a)

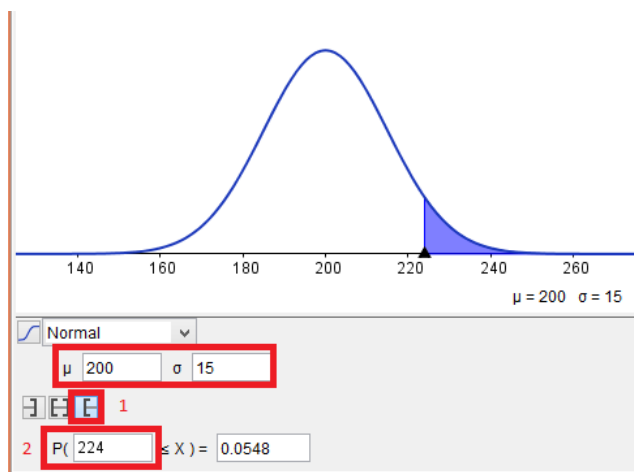


Imagen 17. Datos y resultado del ejemplo 1.a.

- Se presiona dicho botón dado que es un área a la derecha (dichas áreas no se calcularán de esa forma en otras instancias del curso, solamente se calculan áreas a la izquierda).
- Cantidad de vasos planteada, la respuesta es 0.0548.

b)

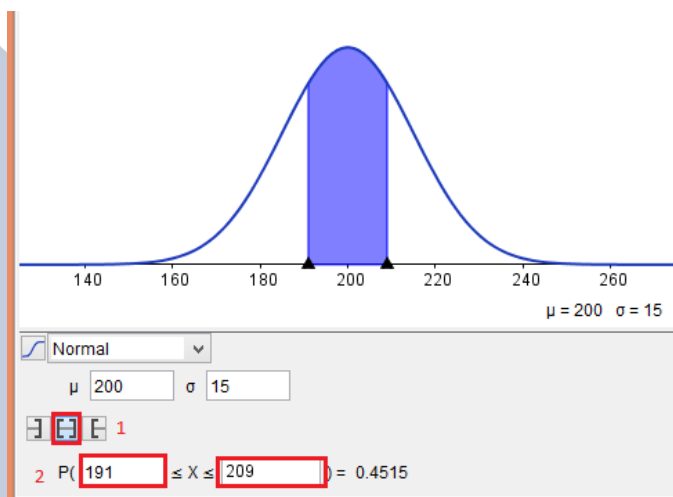


Imagen 18. Datos y resultado del ejemplo 1.b.

- Se presiona dicho botón dado que es un área "entre".
- Cantidad de vasos planteada, la respuesta es 0.4515.

Las preguntas anteriores se realizan utilizando la distribución normal (no estándar). Se repetirá la parte **b** del ejemplo y se resolverá de esa forma.

b) ¿Cuál es la probabilidad de que un vaso contenga entre 191 y 209 ml?

Se calcula la nueva variable: $Z_1 = \frac{X_1 - \mu}{\sigma} = \frac{191 - 200}{15} = -0.6$ y $Z_2 = \frac{X_2 - \mu}{\sigma} = \frac{209 - 200}{15} = 0.6$ por tanto:

$P(191 \leq x \leq 209) = P(-0.6 \leq z \leq 0.6)$ Ahora con la **calculadora de probabilidades** utilizamos la normal estándar, es decir con $\mu = 0$ y $\sigma = 1$.

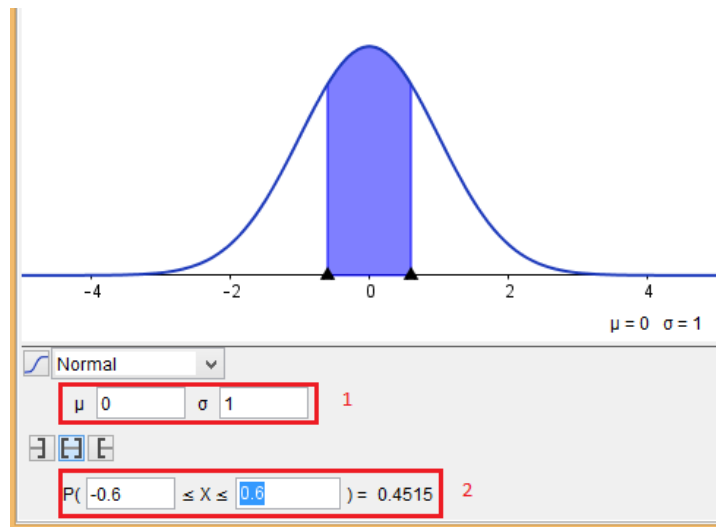


Imagen 19. Datos y resultado del ejemplo 1.b (“estandarizada”).

1. Note los datos ingresados.
2. La respuesta es la misma 0.4515.

Ejercicio 7: Dada la variable X normalmente distribuida con media 18 y desviación estándar 2.5 encuentre:


- a) $P(X < 15)$
- b) El valor de k tal que $P(X < k) = 0.2236$.

Resuelva lo anterior, primero para la distribución normal y luego para la normal estándar, compare sus resultados.

Actividad 8: Recordando la distribución binomial:

La **calculadora de probabilidades** ofrece la distribución binomial, recordemos un ejercicio anterior:

Ejercicio 8: Al probar cierta clase de neumáticos, se encuentra que 25% presentaban pinchaduras. De los siguiente 15 neumáticos, encuentre con ayuda del programa la probabilidad de que...

- Inserte los datos del problema, en la parte superior encontrará el botón  presiónelo para ver la normal sobre la binomial y responda:

¿La normal presente es la normal estándar? Justifique

- Aumente el número n de neumáticos a 20, luego a 50, luego a 100, observe en detalle la gráfica en cada caso, en base a lo anterior y a su estudio responda:

¿Por qué podemos hablar de una “Aproximación normal a la binomial”?

¿Para que es útil dicha aproximación si podemos calcular la misma binomial? Justifique pensando en el caso $n = 100$ y la necesidad de calcular $P(45 < x < 50)$.

Actividad 9: Estimación de la media:

La **calculadora de probabilidades** además nos permite calcular intervalos de confianza usando los estadísticos **z** o **t**, según corresponda, para ello presione el botón **Estadística**.

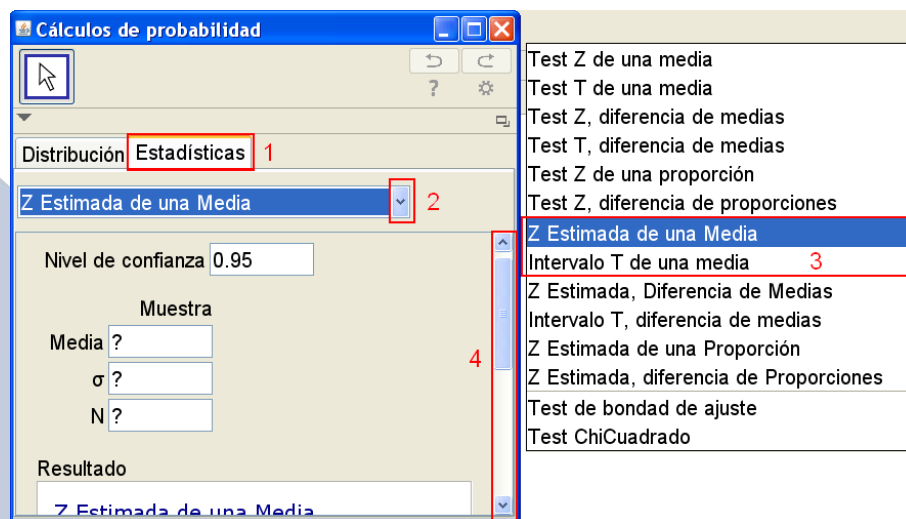


Imagen 20. Calculadora de probabilidad para estadística inferencial.

1. Botón **Estadísticas** para acceder al cálculo de estadística inferencial.
2. Menú desplegable para seleccionar la inferencia deseada.
3. Opciones que se utilizarán en este apartado.
4. Barra que permite acceder a la información (resultados) que nos son de interés.

Ejemplo 7: Una muestra de 100 automóviles recorren en promedio 23 500 kilómetros por año con una desviación estándar de 3900 kilómetros. Suponga que la distribución es aproximadamente normal.

- a) Construya un intervalo de confianza de 99% para el número promedio de kilómetros que recorre un automóvil.

Para resolver el ejemplo se usa la calculadora de probabilidades. En la sección **Estadísticas** se selecciona del menú la opción **Z Estimada de una media**, se colocan los datos y el resultado es inmediato.

Cálculos de probabilidad

Distribución: Estadísticas

Z Estimada de una Media

Nivel de confianza 0.99

Muestra

Media 23500

σ 3900

N 100

Resultado

Z Estimada de una Media

Media	23500
σ	3900
ES	390
N	100
Límite por la derecha	22495.4266
Límite por la izquierda	24504.5734
Intervalo	23500 ± 1004.5734

Imagen 21. Datos y resultado del ejemplo 3.

1. Datos del problema.
2. Intervalo solicitado: $22495.436 < \mu < 24504.57$.

Ejemplo 8: Una muestra de 10 barras de chocolate tiene en promedio 230 calorías con una desviación estándar de 15 calorías. Construya un intervalo de confianza de 99% para el contenido medio de calorías. Suponga que la distribución es aproximadamente normal.

Para resolver el ejemplo se usa la calculadora de probabilidades. En la sección **Estadísticas** se selecciona del menú la opción **Intervalo T de una media**, se colocan los datos y el resultado es inmediato.

Cálculos de probabilidad

Distribución: Estadísticas

Intervalo T de una media

Nivel de confianza 0.99

Muestra

Media 230

s 15

N 10

Resultado

Intervalo T de una media

Media	230
s	15
ES	4.7434
N	10
glib	9
Límite por la derecha	214.5847
Límite por la izquierda	245.4153
Intervalo	230 ± 15.4153

Imagen 22. Datos y resultado del ejemplo 4.

1. Datos del problema.
2. Intervalo solicitado: $214.58 < \mu < 245.42$.

Ejercicio 9: Lea en detalle los ejemplos anteriores, note que en uno se usa el estadístico **Z** y en otro el estadístico **T** y responda ¿En qué caso se utiliza uno o el otro?

Además: Resuelva el siguiente problema:

Una máquina produce piezas de forma cilíndrica, una muestra de los diámetros arroja los siguientes datos: 1.01, 0.97, 1.03, 1.04, 0.99, 0.98, 0.99, 1.01 y 1.03 centímetros. Encuentre un intervalo de confianza de 99% para el diámetro medio de las piezas. Suponga que la distribución es aproximadamente normal.

4. Referencias bibliográficas

MEP (2012) Programas de estudio en matemáticas, MEP, Costa Rica.

Myers, Myers, Ye. (2007) Probabilidad y estadística para ingeniería y ciencias, octava ed. Walpole, Pearson-Prentice Hall.

Hernández L, Probabilidad y Estadística Matemática, Guía de Estudio, UNED.

Manual de GeoGebra 5.0. <https://wiki.geogebra.org/es/Manual>