Tema: Distribuciones probabilísticas.

Alumno: José Luis Arroyo Núñez. *N.U.A: 390893.*

Alumno: Bryan Ricardo Cervantes Mancera. *N.U.A:146809*

Fecha: 15 de diciembre de 2020.

Descripción:

Basados en los conocimientos adquiridos en la U.D.A probabilidad y estadística elaborar una interfaz por medio de Matlab “GUIDE” en la cual se le apoye al usuario para poder graficar y conocer los valores para las distribuciones: Binomial, Poisson, Exponencial, Normal.

Consideraciones: de correr el programa, considere que fue elaborado en “Matlab R2016a” para versiones Windows y MacOS, por lo que las imágenes pueden ajustarse de manera forzada.

Elaborar una interfaz gráfica para introducir distintos valores dependientes al tipo de distribución ya sea discreta o continua, donde también se especifique el tipo de distribución a usar.

Introducción:

Como se mencionó anteriormente en este proyecto se estará realizando la implementación de los conocimientos adquiridos en el curso de probabilidad y estadística, en particular de los temas de distribuciones, en continuas, la distribución Exponencial y la Norma, por otro lado, en las discretas las distribuciones Binomial y la Poisson.

Para poder hacer uso de estas se optó por el utilizar ventanas separadas para cada una de las formas en las que se pueden tener los datos, es decir, discretos o continuos, para poder elegir entre estos se realizó una GUI donde se podrá realizar la selección dependiendo de los casos que se quiere utilizar.

Desarrollo:

<u>Discretas</u>

Binomial:

Se contemplo para este tipo de distribución el de tener una probabilidad, esta probabilidad deberá de ser continua, es decir, que no deberá de ser cambiante dependiendo de los valores de X, este se encuentra denotado por la letra “p”.

Donde debemos de recordar que p debe de ser:

$$0 < p < 1$$

Otro de los puntos que se tomaron en cuenta es que debe de tenerse un valor entero de las veces que se estará repitiendo el evento con esa misma posibilidad, denotado por la letra “n”.

Donde debemos de recordar que n debe de ser:

$$n > 1$$

Para la aplicación de esta distribución es la es la aplicación de la siguiente formulas:

Para la determinación de la probabilidad individual, función densidad:

$$f_X(x) = \binom{n}{x}(p)^x(1-p)^{n-x}$$

Para la determinación de la probabilidad en intervalos y las que son dependiendo de mayor a o menor a, función acumulativa:

$$F_X(x) = \sum_{i=0}^{n}\binom{n}{i}(p)^i(1-p)^{n-i}$$

Para la determinación de la esperanza matemática y la varianza se contemplaron las siguientes formulas:

Esperanza matemática:

$$E[X] = np$$

Varianza:

$$Var[X] = np(1-p)$$

Poisson:

Este tipo de distribuciones que se utiliza en los casos en donde se quiera saber la probabilidad de un suceso ya sea en el espacio o en el tiempo, de esta forma es posible que tengamos que realizar la repetición del evento un número de veces muy grande y de la misma forma la probabilidad de estos es demasiado pequeña, de la misma forma actúa como una forma de simplificar el realizar los cálculos de la distribución binomial, esto es posible si se cumplen las siguientes condiciones:

$$E[X] < 5$$

Esto también se tomó en cuenta para la realización de la interfaz que se desarrolló.

La diferencia de esta distribución con respecto a la distribución binomial es que solamente requiere un parámetro el cual esta denominado como λ, este valor tiene la particularidad de que con este ya se tiene la esperanza matemática y la varianza por lo que quedaría de la siguiente forma expresada:

$$E[X] = Var[X] = \lambda$$

De la misma forma debido a que es posible el realizar el cambio de distribución con la binomial este valor podrá ser obtenido mediante la fórmula de la esperanza matemática de la binomial la cual es la siguiente:

$$E[X] = np$$

Por lo que, si se pasa de la binomial a esta y se tiene estos campos, directamente podremos obtener el valor de λ, Para la determinación de la probabilidad individual, función densidad:

$$f_X(x) = e^{-\lambda}(\frac{\lambda^x}{x!})$$

Para la determinación de la probabilidad en intervalos y las que son dependiendo de mayor a o menor a, función acumulativa:

$$F_X(x) = \sum_{i=0}^{x} e^{-\lambda}(\frac{\lambda^i}{i!})$$

Debido a que nos encontramos con valores discretos la determinación de las probabilidades que podemos obtener seria de la siguiente forma:

$$P(X = x) = f_X(x)$$

$$P(X < x) = F_X(x^-)$$

$$P(X \leq x) = F_X(x)$$

$$P(X > x) = 1 - F_X(x)$$

$$P(X \geq x) = 1 - F_X(x^-)$$

$$P(a \leq X \leq b) = F_X(b) - F_X(a^-)$$

$$P(a < X < b) = F_X(b^-) - F_X(a)$$

$$P(a \leq X < b) = F_X(b^-) - F_X(a^-)$$

$$P(a < X \leq b) = F_X(b) - F_X(a)$$

Donde la apostrofe que se tiene significa que el valor que se tienen se tendrá que realizar la resta de 1 el valor de a, b, o x.

De la misma forma si los intervalos que se tienen contemplado se encuentran fuera del parámetro que se tiene, es decir, desde 0 hasta el numero de repeticiones, el resultado de este es 0, así como de el valor especifico que se tiene en ese valor.

De la misma forma de los valores que se tiene en los intervalos si el inferior se encuentra fuera y dentro del intervalo que se tiene deberá de solamente tomar en cuenta el valor del extremo inferior hasta el valore que se quiere, en caso de que se esta realizando el intervalo se supere por el intervalo de b a el valor donde se encuentra el valor máximo, en este caso solo se tomara en cuenta el valor dado hasta donde se encuentra adelante.

<u>Continuas</u>

Exponencial:

Para la distribución exponencial en cuanto a su programación contemple, únicamente el uso de una x, dado que el cálculo con intervalos puede hacerse fácilmente, de forma manual, restando el límite superior del límite inferior, por lo que no es necesario la implementación adicional de más espacios, en cuanto al valor de c, tome su referencia como tal, por lo que el usuario en el uso del GUIDE, debe considerar que, de tener una media, el valor de c deberá de ser $\frac{1}{media}$; para que de esta forma la formula pueda aplicarse sin problema alguno.

Se siguieron las siguientes fórmulas para obtener los valores mostrados al usuario:

Para la densidad exponencial:

$$fx(x) = \begin{cases} ce^{-cx} & x>0 \\ 0 & x < 0 \end{cases}$$

Para la acumulación exponencial:

$$Fx(x) = \begin{cases} 1 - e^{-cx} & x > 0 \\ 0 & x < 0 \end{cases}$$

Esperanza matemática exponencial:

$$E[x] = \frac{1}{c}$$

Varianza exponencial:

$$Var[x] = \frac{1}{c^2}$$

Normal:

Para la distribución normal fue un poco más sofisticado, pues considere tres situaciones clave para su resolución;

A) x<=X: En este caso solo se considera un valor para x, donde su densidad como acumulación serán afectadas, dado que, al ser por debajo de un valor desconocido, gráficamente se tendrá que restar un 0.5 a la cantidad resultante de dicho valor, pues la media es superior al valor de X.

B) x>=X: Este es el caso contrario, pues se debe considerar que la media es inferior al valor de x deseado, por lo que se debe sumar el valor de 0.5 para así obtener el valor real de la probabilidad.

C x1 <= 2 <= x2: en el caso de los intervalos, fue el más complicado dado que mediante un "linspace" se evalúan a ambos intervalos en relación con el valor de x que introduzca el usuario, por lo tanto, se da el funcionamiento de una integral, solo que, a modo de sumatoria, por lo que es algo más rudimentario, la forma de programarlo se podría decir

Ahora bien, para la aplicación de estas fórmulas se debe considerar puntos importantes dado que, aunque se utilicen las mismas formulas se aplican de diferentes maneras, por lo menos para el caso de los intervalos, dado que en los casos donde:

$x \leq X \; o \; x \geq X$ se aplican de la siguiente manera:

Densidad normal estandarizada:

$z = \frac{x-\mu}{\sigma}$ donde z pasa a calcularse como un error función, mediante:

$$\mathrm{erf}(z) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{\varsigma^2}{2}} d\varsigma$$

Acumulación normal: en este caso no se estandariza dado que se puede calcular de forma normal, por lo tanto, se tiene:

$$Fx(x) = \int_{-\infty}^{x} e^{-\frac{(\varsigma-\mu)^2}{2\sigma^2}} d\varsigma$$

Esperanza:

$$E[x] = \mu$$

Varianza:

$$Var[x] = \sigma^2$$

Finalmente, para la consideración de los intervalos tenemos que la única variación se da en la densidad y en la acumulación dado que las fórmulas en vez de normalizarse se aplican fácilmente con el valor que el usuario le asigna a ambas x, por lo tanto:

Para X entre intervalos, cálculo de la densidad:

$$fx(x) = \frac{1}{\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-\frac{\varsigma^2}{2}} d\varsigma$$

Mientras para la distribución acumulada, se tiene la siguiente modificación a la formula base:

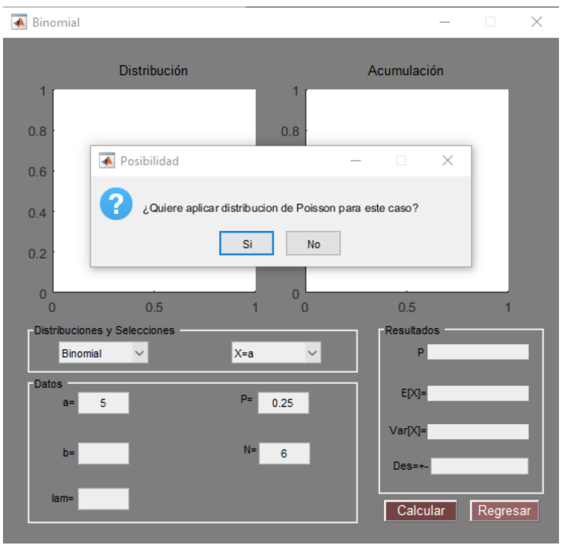
$$Fx(x) = \int_{x_1}^{x_2} e^{-\frac{(\varsigma-\mu)^2}{2\sigma^2}} d\varsigma$$

Resultados:

Binomial:

Para el desarrollo de este se optó por tener la opción de que, si los parámetros que se tienen satisfacen la posibilidad de usar una distribución diferente, en este caso la de Poisson, las condiciones que se deben de cumplir son las siguientes:

$\mu<5$; $n>1$; $p<1$, para esto se muestra la siguiente ventana



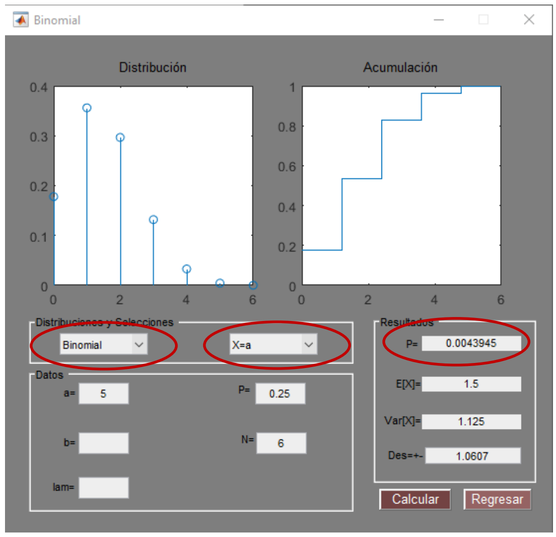
Pruebas del funcionamiento:

Ejemplo:

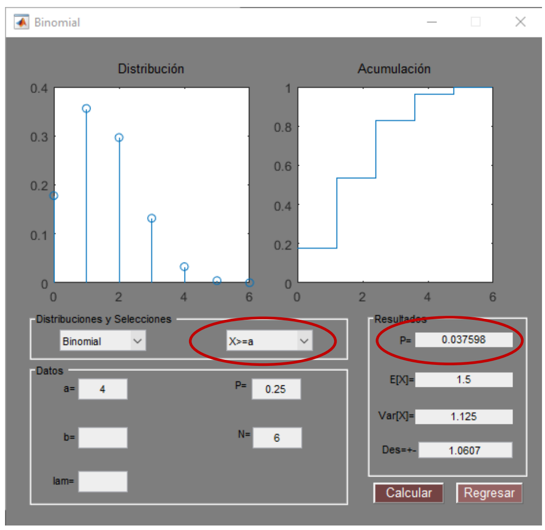
Se tienen 6 urnas para una rifa y en cada una de ellas se tienen 4 pelotas de diferentes colores, determinar las siguientes probabilidades:

N=6 P=1/4=0.25

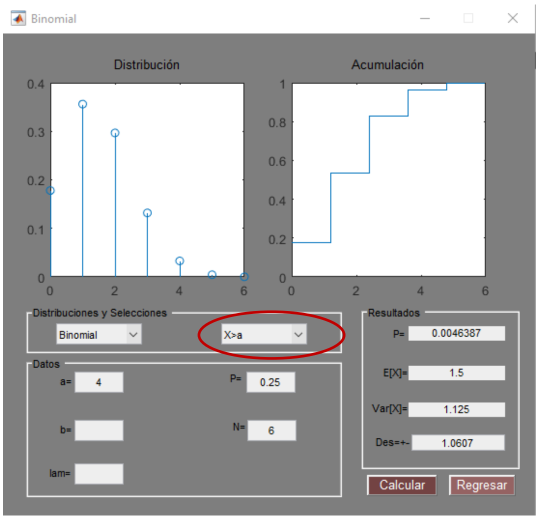
P(X=5)=0.004



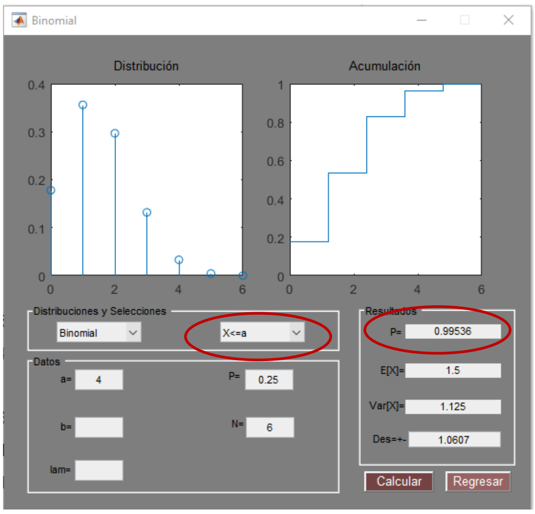
P(X>=4)=0.0375



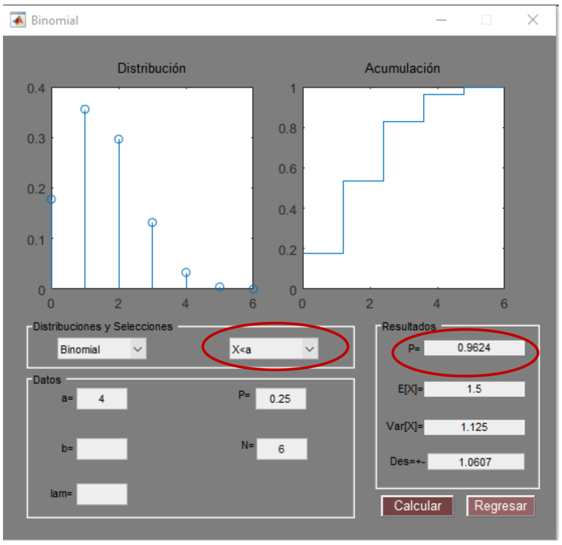
P(X>4)=0.00463



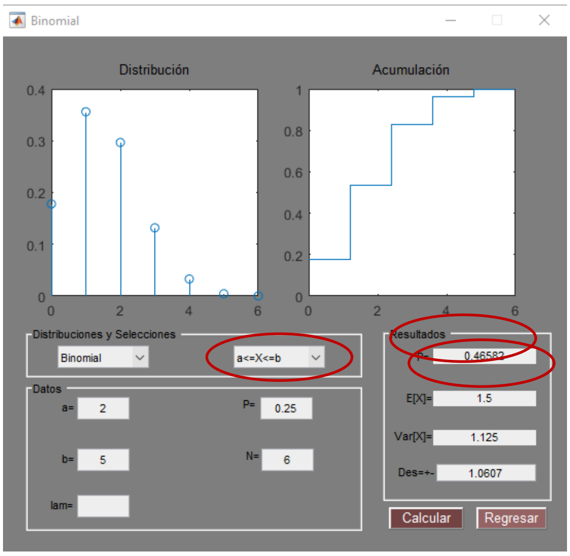
P(X<=4)=0.99536



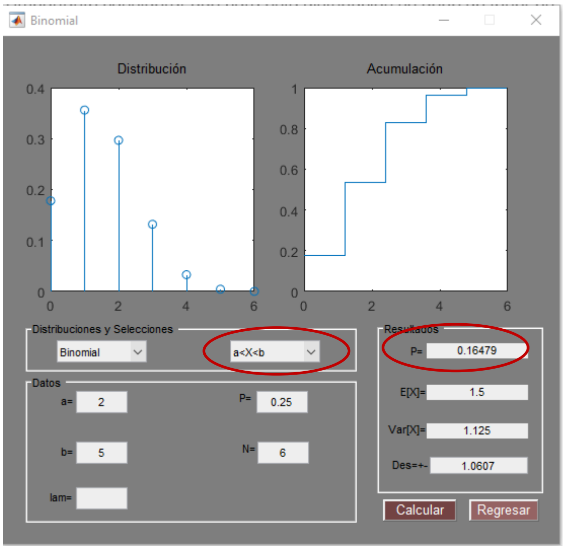
P(X<4)=0.96240



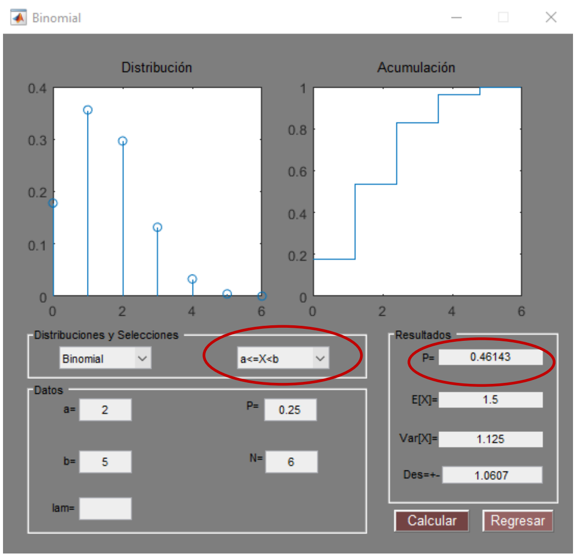
P(2<=X<=5)=0.46581



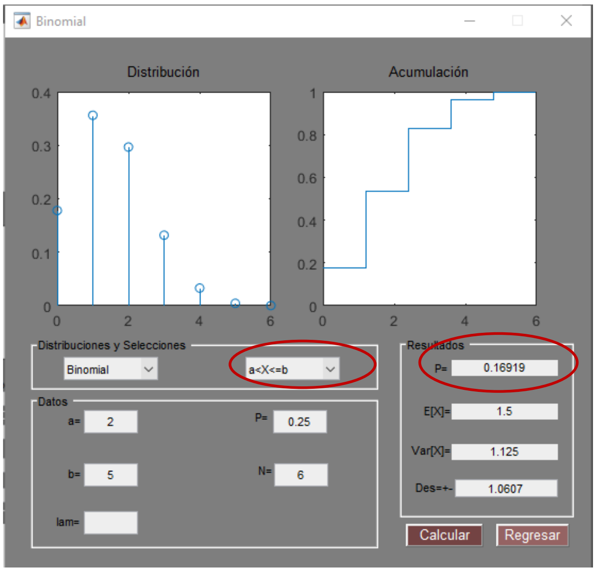
P(2<X<5)=0.164793

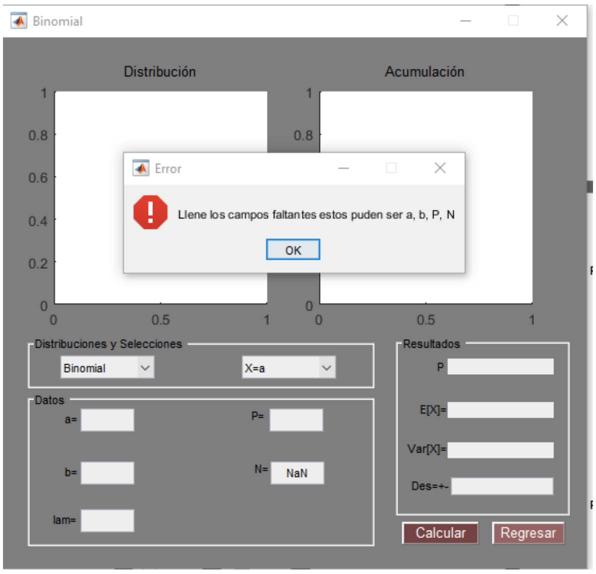


P(2<=X<5)=0.461425

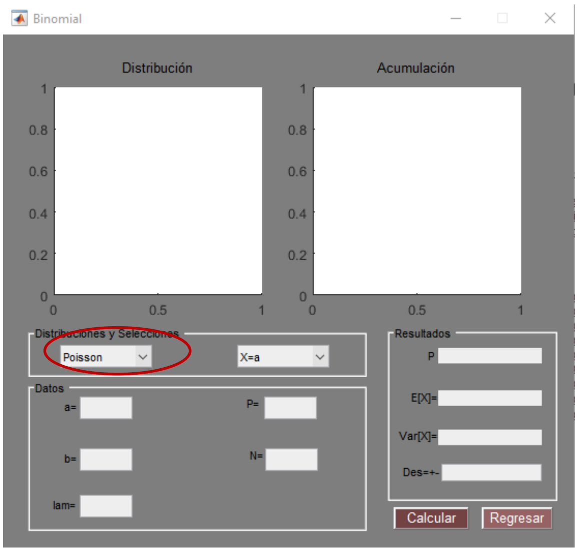


P(2<X<=5)=0.169183





Poisson:



Para poder seleccionar un cambio de la distribución se seleccionará en el menú pop-up

Pruebas del funcionamiento:

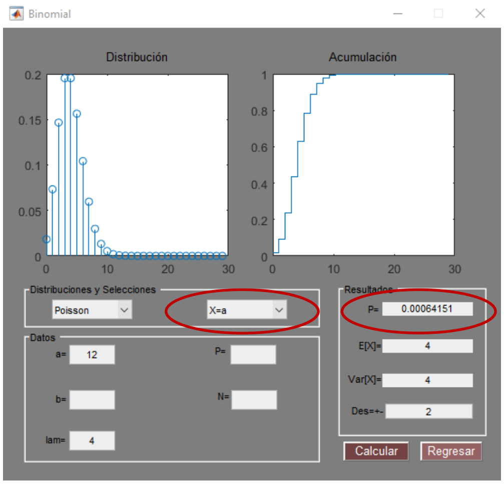
Ejemplo:

Un laboratorio se encarga de realizar un medicamento en especifico el cual en 1 año cada persona tiene una posibilidad entre mil de necesitarlo, de la misma forma este laboratorio produce en promedio 4 mil unidades de este producto por año.

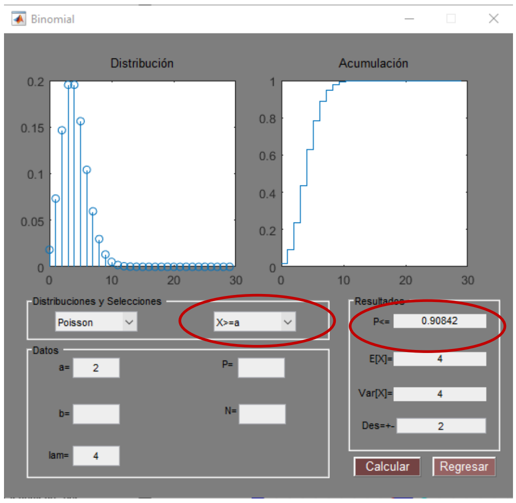
Calcular las probabilidades correspondientes a los datos que se proporcionan:

P=1/1000=0.001 probabilidad de uso de medicamento μ=4000 unidades por año λ=0.001(4000)=4

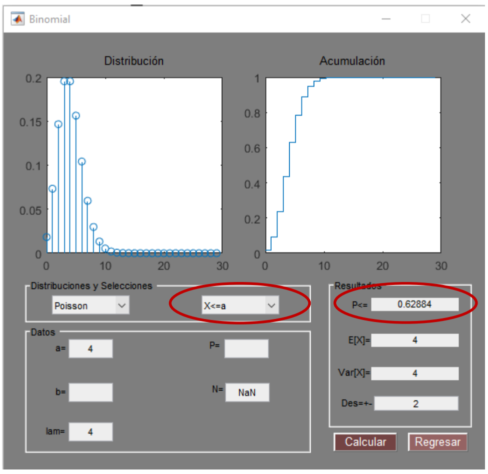
P(X=12)=0.0006



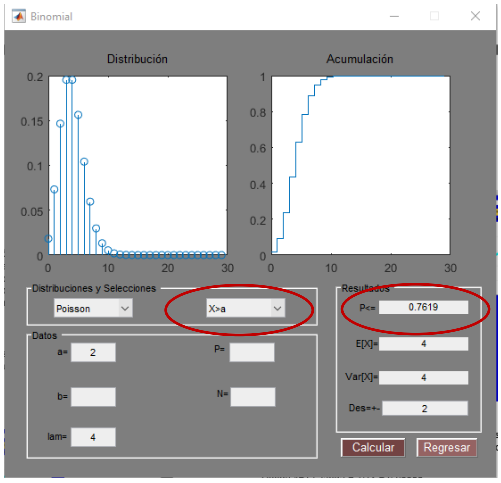
P(X>=2)=0.908421



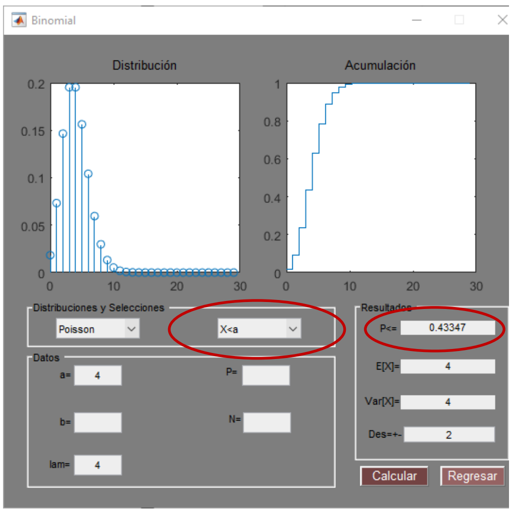
P(X<=4)=0.6288



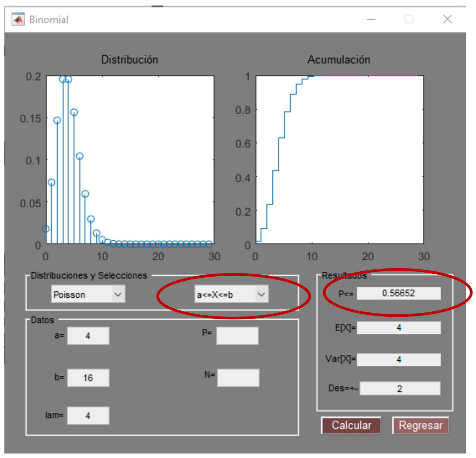
P(X>2)=0.761896



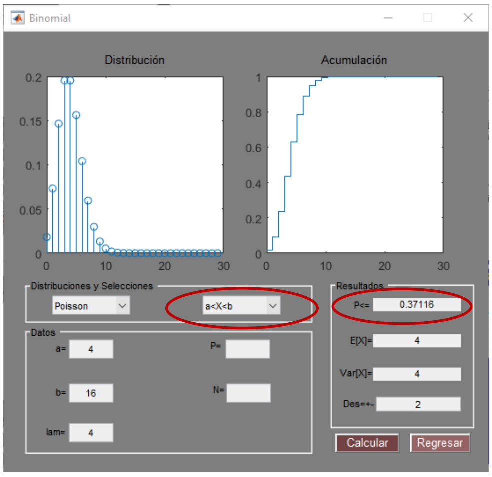
P(X<4)=0.43347

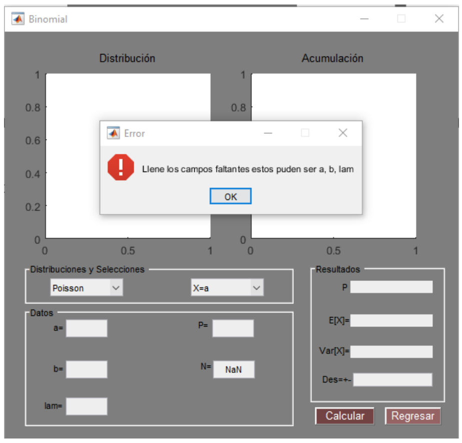


P(4<=X<=16)=0.56651

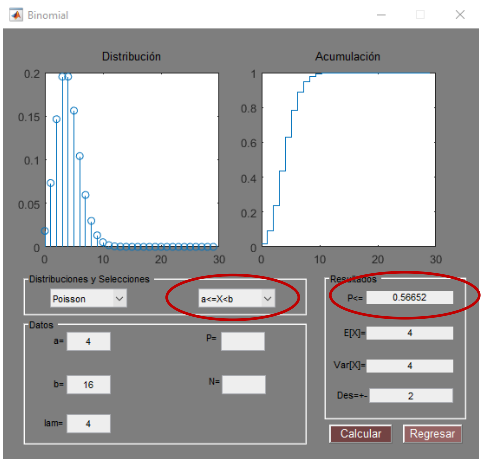


P(4<X<16)=0.371153

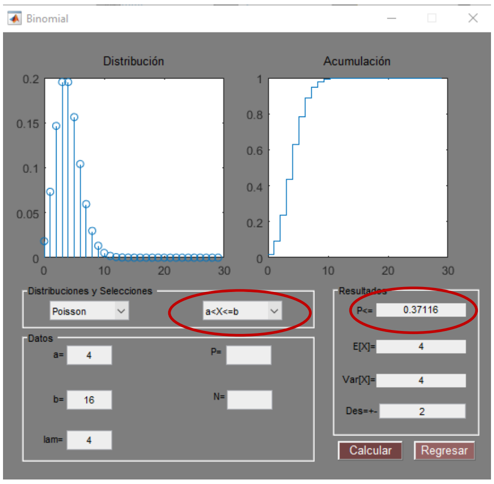




P(4<=X<16)=0.566519



P(4<X<=16)=0.371153



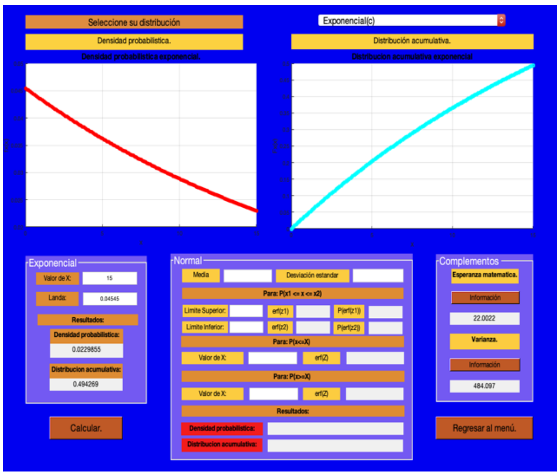
Exponencial: Es importante considerar que para esta distribución se debe de tener en cuenta cual es el resultado que se va a obtener, es decir, para que nos sirve aplicar esta distribución, pues su principal función es la de crear un modelo para representar el tiempo de uso, funcionamiento o rendimiento de un objeto o un fenómeno, por ejemplo, la duración de una batería de X amperaje en un automóvil de alto, mediano y bajo rendimiento.

Ejemplos prácticos:

1) El tiempo de reparación de reparación de unas máquinas de coser tienen una media de 22 minutos aproximadamente, cual es la probabilidad de que el tiempo de reparación sea menor o igual a 15 minutos.

Donde x=15, landa = 1/22 = 0.04545

Prueba de su funcionamiento:

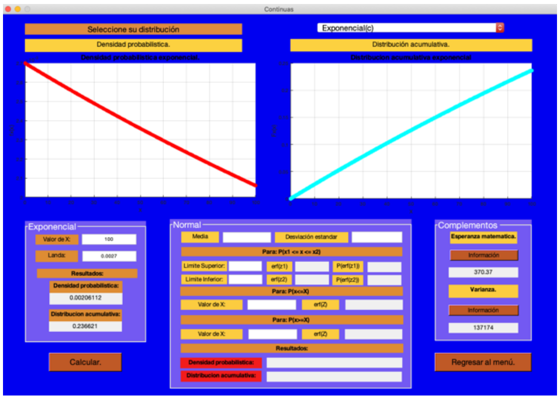


2) El tiempo durante el cual cierta marca de batería trabaja en forma efectiva hasta que falle se distribuye de forma exponencial, con un tiempo promedio de fallas de 365 días, encuentre la probabilidad de que la batería falle cuando se tienen en uso de entre 300 y 400 días.

Dado que se tiene un intervalo [300, 400] se dice que;

x = Límite superior – Límite inferior, por lo que x = 100

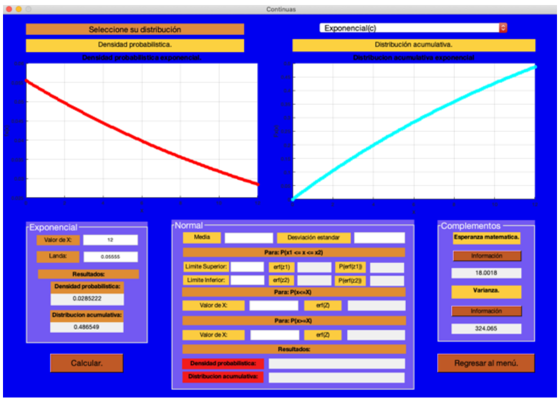
Pruebas:



3) La vida media de una plancha es de 18 meses, siendo esta una variable aleatoria distribuida exponencialmente, ¿Cuál es la probabilidad de que falle antes de los 12 meses?

Donde x=12, landa = 1/18 = 0.05555

Prueba de su funcionamiento:



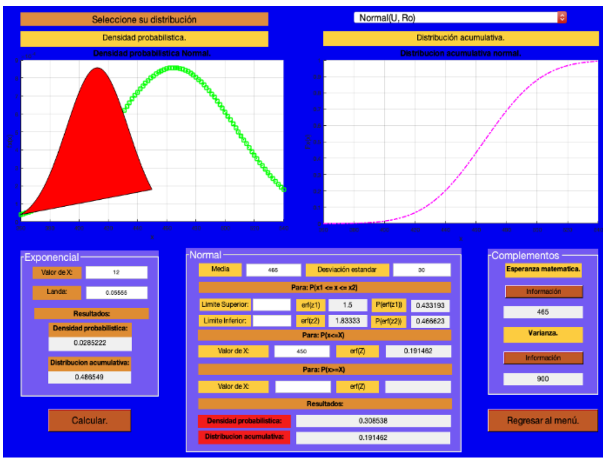
Normal: esta es una distribución que genera un modelo lo más cercano a la realidad, siempre y cuando se tenga una situación ideal, donde se posea una valor promedio, como también una desviación estándar que se podría decir que es la amplitud de la media para las variaciones positivas o negativas que puede tener un suceso, esto se expresa como la división de la campana gaussiana usando la media como referencia donde va desde el 0 a la media, para las variaciones negativas y de la media a un valor x final para las variaciones positivas.

Por lo tanto, una distribución normal se puede usar cuando se requiera una representación real ideal de una situación donde se tengan valores contables, pero demasiado grandes para poder conocer una probabilidad de sucesos exacta de manera individual.

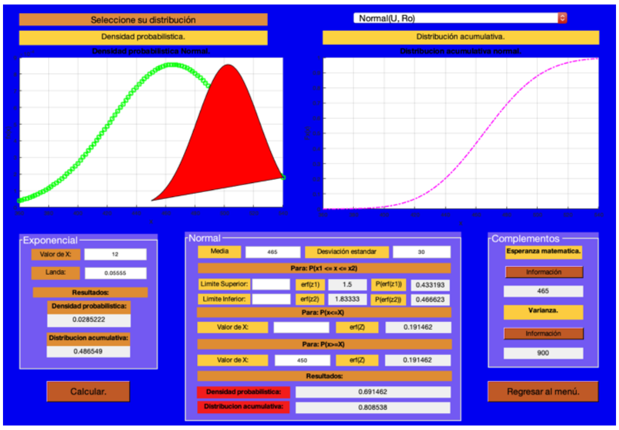
Nota: para usar los casos de los límites, menor que o mayor que, es necesario borrar las cantidades de los otros campos, es decir para poder usar el caso de menor que, los campos de edición en los límites y en el caso mayor que, deben estar vacíos.

Ejemplo: media = 465, desviación estándar = 30

1) Cuando se tiene x<=450:

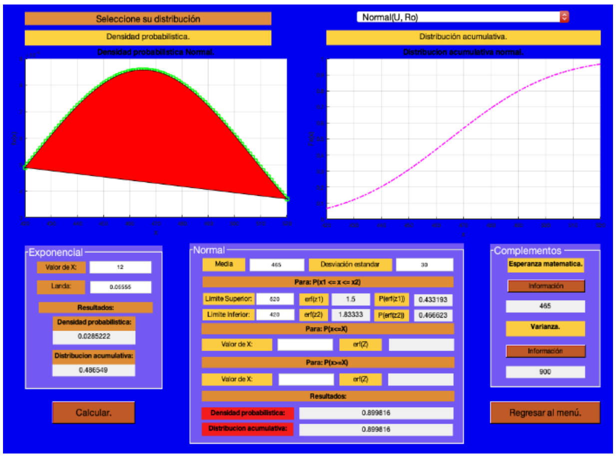


2) Cuando se tiene x>=450:



3) Cuando se tiene a x entre un rango de intervalos:

Donde: límite superior = 520, límite inferior = 420.

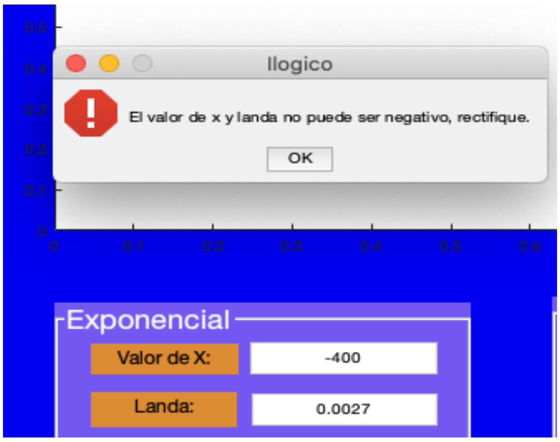


Casos de prueba o falla que el usuario pudiera provocar.

Exponencial:

1) Si x o landa está vacío:

2) Si x o landa es negativo:

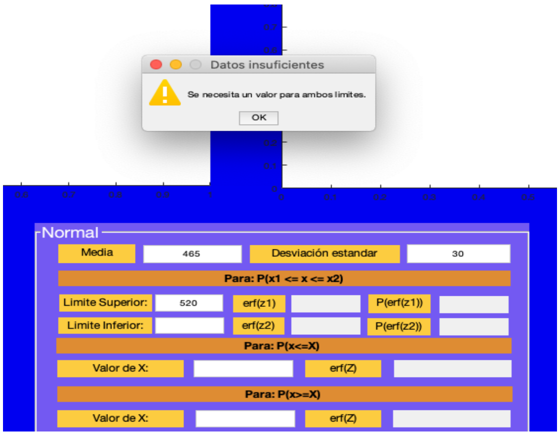


Normal:

1) Si la media o la desviación estándar está vacía:

2) Si los cuatro campos están vacíos:

3) Si uno de los campos para los limites superiores están vacíos:



Conclusiones personales:

Bryan Ricardo Cervantes Mancera:

En particular la forma en la que se realizó la programación sobre las distribuciones discretas, lo mayormente pesado fue el de estar considerando los diversos casos en los que pudieran presentarse al momento en el que el usuario o los diferentes formas en las que los límites que se estén tomando en cuenta para la realización del calculo de la probabilidad que se quiere.

Otra de las observaciones que puedo observar es que en ocasiones con los mismos valores, pero realizando en otra distribución los valores de la probabilidad que se está calculando cambia significativamente, y adicional a esto lo interesante es la forma en la que se comporta la gráfica tanto de la densidad como en la de la acumulativa, ya que conforme se aumente el número de repetición del experimento, estas toman la forma de una área completa, por lo que se ve que en cierto punto estas distribuciones pueden llegar a tener semejanzas con las de las distribuciones continuas.

José Luis Arroyo Núñez: En mi opinión sobre las distribuciones continuas creo que son bastante interesantes en cuanto a su programación pues, se deben considerar varios puntos importantes donde los criterios o valores dados puedan generar un error, además de considerar diferentes rangos de valores para X al momento de generar una u otra gráfica, por lo que considero que programar distribuciones es un ejercicio práctico para entender los puntos clave de la programación, respecto a las múltiples restricciones que puede tener un método implementado de forma computacional

En cuanto a cada una de las distribuciones me pareció interesante el ver cómo es que se graficaba la densidad y la acumulación, pues las mismas formulas ayudan a restringir los errores por parte del usuario, de esta forma haciendo un poco más sencillo el manipular y programar los datos como también la función en si para trabajar con dichos datos.

Conclusiones generales:

Fue una práctica bastante grata e interesante sobre todo por el hecho de usar Matlab pero mediante el comando GUIDE, que nos ayuda a introducir los datos de una manera más sencilla y tener un programa con respuesta así como también un entorno gráfico, para facilitar la entrada y salida de los datos que debemos calcular, a su vez me pareció interesante que la estructura por medio de la GUI varié tanto con respecto a la forma convencional de usar Matlab mediante un workspace y un command window, por lo que fue bueno entender cómo es que funciona una GUIDE, dado que como ingenieros en sistemas puede necesitemos este conocimiento.

Referencias y bibliografías:

Libros:

- Guerrero, D. O. B. (2008). Manual de interfaz gráfica de usuario en Matlab. Parte I–www. matpic. com, 16.
- Guerrero, D. O. B. MANUAL DE INTERFAZ GRÁFICA DE USUARIO EN MATLAB.

Páginas de internet:

- https://la.mathworks.com/help/matlab/creating_guis/adding-components-to-the-gui.html
- https://la.mathworks.com/help/matlab/creating_guis/modal-dialog-box-in-a-guide-gui.html
- https://la.mathworks.com/help/matlab/matlab_env/color-settings.html
- https://la.mathworks.com/matlabcentral/answers/395348-how-can-i-change-the-font-size-of-msgbox
- https://la.mathworks.com/matlabcentral/answers/24674-multiple-lines-in-msgbox
- https://la.mathworks.com/matlabcentral/answers/64095-how-do-i-change-the-msgbox-window-size
- https://la.mathworks.com/help/matlab/learn_matlab/data-analysis.html#bt42juq-1
- https://la.mathworks.com/help/matlab/learn_matlab/data-analysis.html
- https://la.mathworks.com/matlabcentral/answers/80530-running-an-m-file-by-clicking-on-a-pushbutton-in-gui
- https://www.mathworks.com/matlabcentral/answers/36117-check-edittext-empty-or-not
- https://www.mathworks.com/matlabcentral/answers/75688-calling-multiple-gui-s
- https://www.mathworks.com/matlabcentral/answers/80530-running-an-m-file-by-clicking-on-a-pushbutton-in-gui
- https://www.lawebdelprogramador.com/foros/Matlab/1064870-Cerrar-un-gui.html
- https://www.mathworks.com/matlabcentral/answers/63644-how-to-set-the-value-of-popup-menu-in-another-callback
- https://la.mathworks.com/help/matlab/ref/questdlg.html
- https://metodosnumericoscem.weebly.com/uploads/2/5/9/7/25971049/guia_simple_guide_matlab.pdf

Audios visuales:

- https://www.youtube.com/watch?v=YMBQ8_sj24Q
- https://www.youtube.com/watch?v=pJxCDNVGNS4
- https://www.youtube.com/watch?v=NjMixKjFWME
- https://www.youtube.com/watch?v=nak3IY_rvLI

- https://www.youtube.com/watch?v=pN_68no9gOo
- https://www.youtube.com/watch?v=etW5ZRmohh8
- https://www.youtube.com/watch?v=dJdBiavW7KA
- https://www.youtube.com/watch?v=E-hR_OFCNNE
- https://www.youtube.com/watch?v=u1WiazG6Z5s
- https://www.youtube.com/watch?v=ufTSN2s7pqc