respectivamente. Sean W y Z el total de ventas diarias en ambas sucursales y la diferencia de ventas entre sucursales, respectivamente. ¿Cuáles son las desviaciones estándares de W y Z?

9) Se estudió el monto de los préstamos en euros solicitados en un banco en un período de varios años, determinándose que tienen una esperanza de € 2500 y un desvío estándar de € 500. Si se expresaran esos montos en dólares, ¿cuáles serían su esperanza y varianza? Considere un tipo de cambio de U$S 1,25 por euro.

10) En un casino uno de los juegos consiste en lanzar dos dados. Si la suma de los dados es 7 el jugador gana $50, si la suma es 11 gana $100 y si la suma es 2 gana $200. Con cualquier otro resultado no gana nada. Si el costo de jugar es de $2, ¿cuál es la ganancia esperada del jugador?

11) Una compañía de seguros emite una póliza de 1 año, por 1000 dólares, que ofrece protección contra el evento A que ocurre a 2 de cada 100 personas que la compran. Los costos administrativos son de 15 dólares por póliza. ¿Cuánto debería cobrar anualmente la compañía por póliza si se requiere que la ganancia esperada que produzca sea de 50 dólares? (Suponga que el evento A no puede suceder más de una vez en un año).

12) Sean $X_1, \ldots, X_n$ independientes e idénticamente distribuidas como X, donde $E(X) = \mu$ y $V(X) = \sigma^2 > 0$. Indique en cada uno de los siguientes ítems si lo que se presenta se trata de (i) constante no nula, (ii) cero o (iii) variable aleatoria. Justifique convenientemente su elección en cada ítem.

a) $\bar{X} - \mu$ b) $V(\bar{X} - \sigma^2/n)$ c) $E(X_1 - \bar{X})$ donde $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

13) Cuando se apuesta $1 a un número de dos cifras en primera en el juego de la Quiniela se puede obtener un premio de $70. Calcule la ganancia esperada para un jugador que apuesta $10. ¿Cómo se interpreta este valor? Calcule la varianza de la ganancia.

14) Suponga que X es una variable aleatoria con: $E(X)=10$, $P(X \leq 7)=0,2$ y $P(X \geq 13)=0,3$. Demostrar que $V(X) \geq 9/2$.

15) Un servicio voluntario de ambulancias recibe de 0 a 5 llamadas en cualquier día. La distribución de probabilidad del número de llamadas en el servicio se muestra en la siguiente tabla:

Número de llamadas	0	1	2	3	4	5
Probabilidad	0,10	0,15	0,30	0,20	0,15	0,10

a) ¿Cuál es número esperado de llamadas al servicio?
b) ¿Cuál es la varianza de la cantidad de llamadas al servicio? ¿Cuál es la desviación estándar?

16) Suponga que X e Y tienen la siguiente función de probabilidad conjunta:

f(x,y)		***x*** = **2**	***x*** = **4**
y	**1**	0,2	0,2
	3	0,15	0,1
	5	0,15	0,2

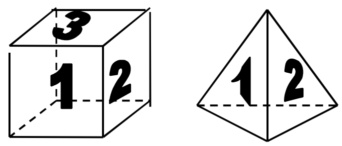
a) Calcule $E(X^2Y)$.
b) Determine $E(X)$, $E(3X\text{-}5)$, $V(5X)$, $V(Y\text{-}2)$ y $Cov(X,Y)$.

c) ¿Son X e Y variables independientes? Justifique.

17) Se arrojan un dado, con sus caras numeradas del 1 al 6, y un tetraedro, con sus caras numeradas del 1 al 4, y se observan los números que aparecen en la cara superior del dado y en la cara del tetraedro que queda apoyada en la mesa.

X: "cantidad de números pares obtenidos"

Y: "diferencia entre la cantidad de números pares y la cantidad de impares obtenidos"



a) Determine la función de probabilidad conjunta.
b) ¿Son independientes las variables X e Y?
c) Calcule $E(X)$ y $E(Y)$, usando las distribuciones marginales y la probabilidad conjunta.
d) Calcule $E(X^2Y)$.
e) ¿Es cierto que $V(X + Y) = V(X) + V(Y)$? Justifique.

18) Dos cápsulas se seleccionan al azar de un frasco que contiene tres aspirinas, dos sedantes y cuatro laxantes. Si X e Y, son respectivamente, los números de cápsulas de aspirinas y sedantes incluidas entre las dos cápsulas que se sacaron del frasco:
a) Encuentre la función de distribución conjunta de X e Y.
b) Halle las distribuciones marginales de X e Y.
c) Calcule $E(X)$, $E(Y)$, $V(X)$, $V(Y)$ usando las distribuciones marginales y la probabilidad conjunta.

19) Considere la variable aleatoria X que registra el número de veces en una semana que un estudiante de un grupo ¨A¨ llega tarde a clases. Suponga que la distribución de probabilidad de esta variable es:

x	0	1	2	3	4
$p_X(x)$	0,15	0,15	0,50	0,15	0,05

a) Encuentre el número esperado de veces que un estudiante llega tarde a clases en una semana.
b) ¿Cuál es la varianza y la desviación estándar de X?
c) Si se define $g(X) = 5.X + 10$, calcule la esperanza y la varianza de $g(X)$.

20) Suponga que X e Y son variables aleatorias discretas con distribución de probabilidad conjunta dada por:

$f(x,y)$		x		
		-1	0	1
y	-1	1/8	0	1/8
	0	1/8	0	1/8
	1	1/8	1/4	1/4

a) Calcule $E(X.Y)$, $E(X)$, $E(Y)$, $E(X+2)$, $V(3Y)$ y $Cov(X,Y)$.
b) ¿Son X e Y independientes? Justifique.

21) Un fabricante produce cierto tipo de aceite lubricante que pierde alguno de sus atributos especiales si no se usa dentro de un período de tiempo. Sea X el volumen de aceite pedido al fabricante durante un año. Suponga que la función de densidad de X es $f(x) = \begin{cases} x-2 & 2 < x < 3 \\ 4-x & 3 \leq x < 4 \\ 0 & \text{en cualquier otro punto} \end{cases}$, donde una unidad es igual a 1.000 litros.

a) Calcule el volumen promedio de aceite pedido al fabricante durante un año.
b) Por cada unidad vendida se obtiene una utilidad de $300, mientras que cada unidad no vendida durante un año determinado produce una pérdida de $100, porque tendrá que ser descartada. Si en un año se fabrican 3.200 litros, ¿cuál es la utilidad esperada?

22) En una industria farmacéutica una máquina produce 100 cápsulas por minuto. Se sabe que en promedio la máquina produce un 5% de las cápsulas sin la cantidad suficiente del medicamento y por lo tanto se consideran defectuosas. Haciendo los supuestos necesarios responda:
a) Si las cápsulas son envasadas en tabletas de 5 unidades, ¿cuál es la probabilidad de que una tableta contenga como máximo 2 cápsulas defectuosas?
b) ¿Cuál es el número esperado de cápsulas con defecto por tableta?
c) ¿Cuál es el número esperado de cápsulas con defecto por minuto de producción?

EJERCICIOS SUGERIDOS DEL LIBRO "PROBABILIDAD Y ESTADÍSTICA PARA INGENIERÍA Y CIENCIAS" de WALPOLE, MYERS y MYERS. (9º Ed.)

Capítulo 4:

Sección 4.1, página 117: ejercicios: 4.9, 4.19, 4.23.
Sección 4.2, página 127: ejercicios: 4.49, 4.50.
Secciones 4.3 y 4.4, páginas 137: ejercicios: 4.62, 4.63, 4.78.

CAPÍTULO 5: DISTRIBUCIONES DISCRETAS DE PROBABILIDAD

Distribuciones discretas de probabilidad: distribución Bernoulli, distribución discreta uniforme, distribución binomial, distribución hipergeométrica, distribución geométrica y distribución de Poisson.

Objetivos:

El alumno debe ser capaz de:

- Conocer la expresión matemática de la función de masa de las siguientes variables aleatorias: uniforme, Bernoulli, binomial, hipergeométrica, geométrica y Poisson.
- Para una distribución dada de una variable aleatoria discreta calcular su esperanza y varianza.
- Dado un problema que involucre variables aleatorias discretas identificar la distribución de probabilidad que describa su comportamiento.
- Calcular probabilidades bajo las distribuciones binomial y Poisson y utilizar estos valores para la resolución de problemas.

Resumen

D1. *Distribución Bernoulli*: La variable aleatoria *X* se dice que tiene distribución Bernoulli con parámetro π,$(X \sim B(\pi))$si y sólo si

$$f(x;\pi) = \pi^x (1-\pi)^{1-x}, x = 0,1.$$

D2. *Distribución Uniforme Discreta*: La variable aleatoria *X* se dice que tiene distribución Uniforme discreta si y sólo si

$$f_X(x;k) = \frac{1}{k}, x = x_1, x_2, \ldots, x_k.$$

D3. *Distribución Binomial*: La variable aleatoria *X* se dice que tiene distribución Binomial con parámetros *n* y π,$(X \sim b(n,\pi))$si y sólo si

$$f_X(x;n,\pi) = \binom{n}{x} \pi^x (1-\pi)^{n-x}, x = 0,1,\ldots,n.$$

D4. *Distribución Geométrica*: La variable aleatoria *X* se dice que tiene distribución Geométrica con parámetro π, $(X \sim Geom(\pi))$, si y sólo si

$$f_X(x;\pi) = \pi(1-\pi)^{x-1}, x = 1,2,\ldots$$

D5. *Proceso de Bernoulli*: Se dice que un proceso es de Bernoulli cuando presenta las siguientes características:

(1) El proceso consiste en repeticiones (intentos) de un mismo procedimiento.

(2) Los resultados de cada uno de los intentos pueden clasificarse como un éxito o un fracaso.

(3) La probabilidad de éxito, representada porπ, permanece constante para todos los intentos.

(4) Los intentos repetidos son independientes.

T1. En un proceso de Bernoulli con *n* intentos y probabilidad de éxito π, la variable aleatoria que cuenta el número de éxitos en las *n* repeticiones tiene distribución Binomial con parámetros *n* yπ.

D6. *Distribución Hipergeométrica*: La variable aleatoria *X* se dice que tiene distribución Hipergeométrica con parámetros *N*, *n* y *k* $\left(\boldsymbol{X \sim Hip(N, n, k)}\right)$ si y sólo si

$$\boldsymbol{f_X(x; N, n, k) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}, x = 0, 1, \ldots min(n, k).}$$

T2. La variable aleatoria que cuenta el número de éxitos en una muestra aleatoria de tamaño *n* seleccionada, sin reposición, de un conjunto constituido por *N* elementos, de los cuales *k* son considerados éxitos y *N-k* fracasos, tiene distribución Hipergeométrica con parámetros *N*, *n* y *k*.

D7. *Distribución Poisson*: La variable aleatoria *X* se dice que tiene distribución Poisson con parámetro β $(\boldsymbol{X \sim P(\beta)})$si y sólo si

$$\boldsymbol{f_x(x, \beta) = e^{-\beta}\frac{\beta^x}{x!}, x \in N_0}$$

D8. *Proceso de Poisson*: Se dice que un proceso es de Poisson cuando consiste en observar la ocurrencia de un fenómeno concreto en un proceso físico particular y presenta las siguientes características:

(1) Hipótesis de independencia: el número de ocurrencias en dos intervalos de tiempo o regiones disjuntos son independientes.

(2) Hipótesis de proporcionalidad y homogeneidad: la probabilidad de exactamente una ocurrencia en un intervalo de tiempo muy pequeño, o una región muy pequeña, es proporcional a la longitud del intervalo, o al tamaño de la región, y no depende del intervalo en particular.

(3) Hipótesis de regularidad: la probabilidad de tener más de una ocurrencia en un intervalo de tiempo particular muy pequeño, o región muy pequeña, es despreciable.

T3. En un proceso de Poisson, la variable aleatoria que cuenta el número de ocurrencias en un intervalo de tiempo fijo $[0,t]$ tiene una distribución Poisson con parámetro λt, donde λ es el número medio de ocurrencias por unidad de tiempo.

T4. Sea *X* una variable aleatoria Binomial con parámetros *n* yπ. Cuando $n \to \infty, \pi \to 0$, y $n\pi$ permanece constante, la distribución Binomial se puede aproximar por una distribución Poisson con media$n\pi$, es decir

$$b(x; n, \pi) - p(x; n\pi) \underset{n\to\infty}{\to} 0,$$

donde$b(x; n, \pi)$ y $p(x; n\pi)$ denotan las funciones de masa de la Binomial y la Poisson respectivamente.

Observación: La aproximación de la Binomial a la Poisson se considera razonable en la práctica, para un *n* finito, cuando:$\boldsymbol{\pi < 0,1}$**y**$\boldsymbol{n\pi > 1.}$

Esperanza y varianza de algunas distribuciones discretas

Distribución	Función de masa	Parámetros	Esperanza	Varianza
$B(\pi)$	$\pi^x(1-\pi)^{1-x}, x=0,1$	π	π	$\pi(1-\pi)$
$Unif(k)$	$\frac{1}{k}, x=x_1 \ldots x_k$	k	$\sum_{i=1}^{k} x_i/k$	$\frac{\sum_{i=1}^{k} x_i^2}{k}-\frac{\left(\sum_{i=1}^{k} x_i\right)^2}{k^2}$
$b(n,\pi)$	$\binom{n}{x}\pi^x(1-\pi)^{n-x}, x=0 \ldots n$	n,π	$n\pi$	$n\pi(1-\pi)$
$P(\beta)$	$\frac{e^{-\beta}\beta^x}{x!}, x=0,1,\ldots$	β	β	β
$Hip(N,n,k)$	$\frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}$	N,n,k	$\frac{nk}{N}$	$\frac{N-n}{N-1}n\frac{k}{N}\left(1-\frac{k}{N}\right)$
Geom (π)	$\pi(1-\pi)^{x-1}, x=1,2,\ldots$	π	$\frac{1}{\pi}$	$\frac{1-\pi}{\pi^2}$

Ejemplos

Ejemplo 1: Una empresa vitivinícola produce vinos finos y ha solicitado catadores expertos capaces de discernir entre un vino fino y uno ordinario el 90% de las veces, con solo degustar un sorbo de cada tipo. Todos los aspirantes al cargo deben probar nueve tipos de vino y decidir si se trata de vino fino u ordinario. La empresa ha determinado que quienes acierten por lo menos en seis de los nueve ensayos serán contratados.

a) Determinar la probabilidad de que un individuo que no conoce nada de vinos y sólo está adivinando logre pasar la prueba y ser contratado.
b) Calcular la probabilidad de que un catador experto (que en efecto es capaz de acertar el 90% de las veces) no logre pasar la prueba.
c) Calcular la media y el desvío estándar del número de aciertos para un catador experto y para otro que adivina.

Resolución:

El proceso que se realiza es
ε: degustar un sorbo de 9 marcas de vino y clasificarlas como vino fino y ordinario.

Este proceso es de Bernoulli porque:

i) Consiste en repeticiones de un mismo experimento (degustación y clasificación).
ii) Los resultados de cada intento se pueden clasificar en éxito o fracaso (acierta o no la calidad del vino).
iii) La probabilidad de éxito permanece constante para todos los intentos (se puede pensar que la probabilidad de acertar es inherente a cada catador y que no depende del tipo de vino que se está experimentando).
iv) Se puede suponer que los intentos son independientes (el hecho de que acierte al calificar un vino en una degustación no modifica su chance de calificar correctamente a otro vino).

a) Si un individuo no conoce nada de vinos (novato) y está adivinando, la probabilidad de que acierte es $\pi = \frac{1}{2}$. Para que un novato sea contratado, debe acertar en la clasificación de por lo menos 6 vinos.

Se define la variable aleatoria

X: número de clasificaciones correctas en las 9 degustaciones.

Se puede suponer que la variable tiene una distribución binomial con parámetros $n = 9$ y $\pi = \frac{1}{2}$, puesto que cuenta el número de éxitos en 9 intentos de un proceso Bernoulli.

$$P(X \geq 6) = \sum_{x=6}^{9} f\left(x; 9, \frac{1}{2}\right) = 1 - P(X \leq 5)$$

$$= 1 - \sum_{x=0}^{5} f\left(x; 9, \frac{1}{2}\right) = 1 - 0{,}7461 = 0{,}2539$$

Por lo tanto la probabilidad de que un novato pase la prueba es de 0,2539.

b) Si el catador es experto, $\pi = 0{,}9$, no logrará pasar la prueba si clasifica correctamente menos de 6 vinos.

$$P(X < 6) = \sum_{x=0}^{5} f(x; 9{,}0{,}9) = 0{,}0083.$$

c) Para el catador experto:

$$E(X) = n\pi = 9 \times 0{,}9 = 8{,}1$$
$$V(X) = n\pi(1-\pi) = 9 \times 0{,}9 \times 0{,}1 = 0{,}81$$
$$\sigma = \sqrt{V(X)} = \sqrt{0{,}81} = 0{,}9.$$

Para el catador que adivina:

$$E(X) = n\pi = 9 \times 1/2 = 4{,}5$$
$$V(X) = n\pi(1-\pi) = 9 \times 1/2 \times 1/2 = 2{,}25$$
$$\sigma = \sqrt{V(X)} = \sqrt{2{,}25} = 1{,}5.$$

Ejemplo 2: De un lote de discos usados, que contienen trabajos de alumnos para revisión, se estima que aproximadamente el 60% contiene virus. Si los discos se someten al detector de virus de McAfee, uno por uno, y el detector es perfecto (detecta toda vez que el disco tenga un virus).

a) Calcular la probabilidad de que al examinar el séptimo disco se detecte virus por primera vez.

b) ¿Tuvo que agregar algún supuesto para poder resolver el ejercicio?

Resolución:

a) El experimento aleatorio es

ε: se examina una serie de discos con el antivirus McAfee hasta que se detecta virus por primera vez.

Se considera éxito (E) si se detecta virus y fracaso (F) si no se lo detecta.

En el contexto de variable aleatoria, se considera éxito a lo que se está buscando observar, aunque esté asociado a algo no deseable, como lo es el caso de que un disco tenga virus. En consecuencia, no se debe confundir éxito, con lo deseable.

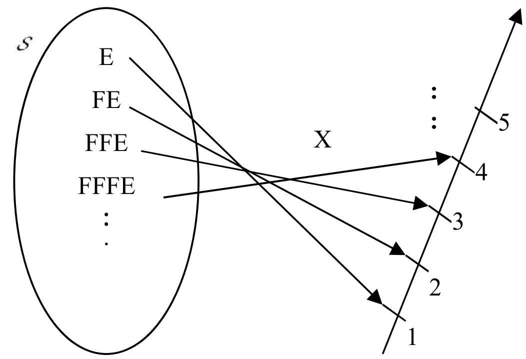
El espacio muestral de este experimento es

$\mathcal{S}$ = {E, FE, FFE, FFFE, FFFFE,...}.

La variable aleatoria X asocia a cada resultado el número de discos observados hasta que se detecta virus por primera vez, por lo tanto

Rec(X)= {1,2,3,4,...}.

Consecuentemente, esta variable es discreta, porque tiene un recorrido infinito numerable.



Si los sucesivos ensayos son independientes, la distribución de probabilidad de esta variable se puede identificar con una distribución *geométrica* y su función de masa es

$$g(x;\pi) = (1-\pi)^{x-1}.\pi \qquad x = 1,2,3,...$$

donde π denota la probabilidad de éxito. De esta manera la probabilidad de que al examinar el séptimo disco se detecte virus por primera vez se calcula de la siguiente manera:

$$P(X=7) = g(7;0{,}6) = (1-0{,}6)^{7-1} \times 0{,}6 = 0{,}00246.$$

b) Se utilizó el supuesto de independencia de los ensayos para poder clasificar a la variable como geométrica.
La función de masa de la variable geométrica es $f(x) = (1-\pi)^{x-1}\pi$, puesto que, si recién en el tiro x se obtuvo un éxito, se infiere que los primeros x-1 intentos fueron fracasos, consecuentemente, la probabilidad de tener x-1 fracasos y un éxito es $(1-\pi)^{x-1}\pi$.

Ejemplo 3: El departamento de reservas de una aerolínea habilitó un centro de atención telefónica. Sea X la variable aleatoria que cuenta el número de llamadas que reciben en un intervalo de tiempo [0, t].

a) ¿Con qué distribución de probabilidad describiría Ud. la variable X?¿Por qué?
b) Suponiendo que en promedio entran 2 llamadas cada 5 minutos, calcular la probabilidad de que al observar 3 períodos disjuntos de 10 minutos, en ninguno de ellos lleguen llamadas.
c) Por cada diez minutos, el mantenimiento del centro de atención tiene un costo fijo de 2 pesos, y un costo adicional de 1 peso si se reciben más de 10 llamadas. Suponiendo que en media entran 2 llamadas cada 5 minutos, calcular el costo esperado por hora de funcionamiento del centro de atención.
d) Resuelva los apartados b) y c) empleando Microsoft Office Excel y la aplicación Probability Distributions[2].

Resolución:

a) X: número de llamadas telefónicas que entran a un conmutador en [0, t].
X es una variable aleatoria Poisson y su función de masa es

$$f(x;\lambda) = e^{-\lambda t}(\lambda t)^x/x!, \quad x = 0,1,2,3,....$$

donde λ representa el número medio de llamadas por unidad de tiempo.

[2] Aplicación desarrollada por Matthew Bognar de la universidad de Iowa, que calcula y grafica probabilidades de variables aleatorias con diferentes distribuciones, según los parámetros que se le indiquen. Disponible en forma gratuita en Google play.

Se afirma que sigue una distribución Poisson, puesto que puede considerarse que el proceso que consta en observar las llamadas a la central posee las siguientes características:

(1) Hipótesis de independencia: el número de llamadas en un intervalo de tiempo $[\boldsymbol{t_1}, \boldsymbol{t_2}]$ es independiente del número de llamadas en otro intervalo de tiempo $[\boldsymbol{t_3}, \boldsymbol{t_4}]$, con $\boldsymbol{t_2} < \boldsymbol{t_3}$.

(2) Hipótesis de proporcionalidad y homogeneidad: la probabilidad de que ocurra exactamente una llamada en un intervalo de tiempo muy pequeño es proporcional a la longitud del intervalo de tiempo y no depende del particular intervalo.

(3) Hipótesis de regularidad: la probabilidad de tener más de una llamada en un intervalo de tiempo muy pequeño es despreciable.

b) Si se considera un lapso de tiempo de 10 minutos, $\lambda t = \frac{2}{5} \times 10 = 4$, y la probabilidad de que no lleguen llamadas en un período de 10 minutos es:

$$P(X = 0) = e^{-4} \times \frac{4^0}{0!} = 0{,}018.$$

Por la hipótesis de independencia del proceso de Poisson, el número de ocurrencias en dos intervalos de tiempos disjuntos son independientes, luego la probabilidad de seleccionar 3 períodos disjuntos de 10 minutos en los cuales no se registren llamadas es

$$P(X = 0) \times P(X = 0) \times P(X = 0) = 0{,}018^3 = 0{,}00000614.$$

c) Sea *X* el costo del mantenimiento del centro por cada 10 minutos. La función costo, que se denota con *C*, puede considerarse de la siguiente manera:

$$C = \begin{cases} 2 & si\ \ X \leq 10 \\ 3 & si\ \ X > 10 \end{cases}$$

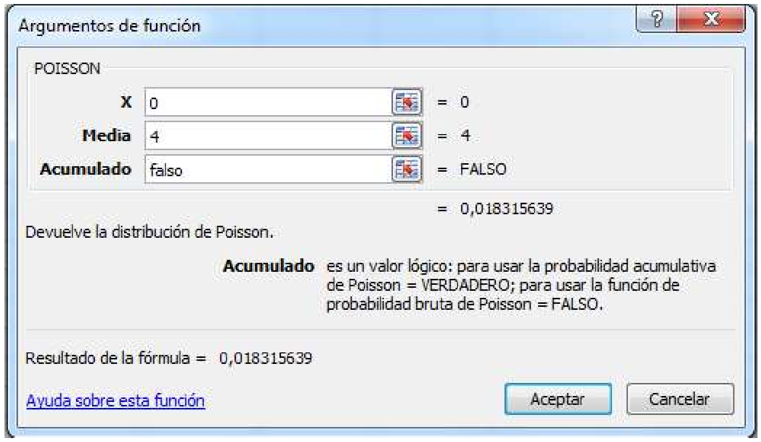
Por cada periodo de diez minutos el costo medio de mantenimiento es:

$$E(C) = 2 \times P(X \leq 10) + 3 \times P(X > 10)$$
$$= 2 \times 0{,}997 + 3 \times 0{,}028 = 2{,}003.$$

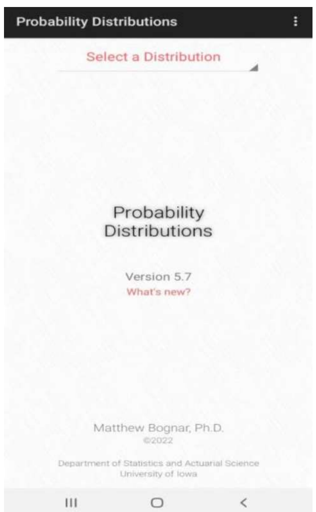
Luego el costo esperado por hora de funcionamiento es igual a $6 \times 2{,}003 = 12{,}02.$ Puesto que es el costo esperado de los primeros diez minutos, más el de los segundos 10 minutos, y así sucesivamente hasta sumar el costo esperado de los últimos 10 minutos de la hora.

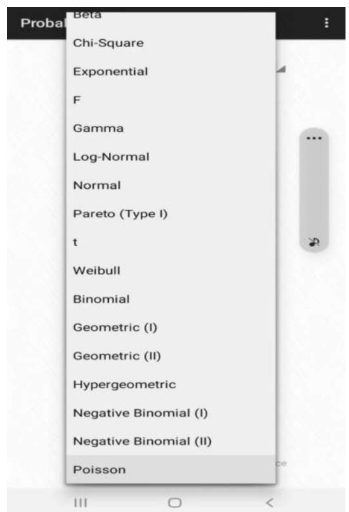
d-b) Para encontrar la probabilidad de que en ninguno de los 3 períodos disjuntos de 10 minutos lleguen llamadas, primero se debe calcular el número medio de llamadas λ en un lapso de tiempo de 10 minutos, $\lambda t = \frac{2}{5} \times 10 = 4$y, para encontrar la probabilidad, al usar Excel se selecciona la opción Fórmulas, Más Fórmulas, Estadísticas, POISSON.DIST y se procede a llenar el cuadro como se muestra en la imagen siguiente:

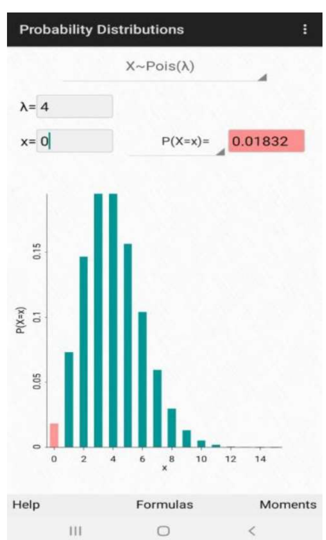
Por lo tanto, la probabilidad de que no lleguen llamadas en un lapso de tiempo de 10 minutos es *P*(*X*=0)=0,018.



También es posible calcular esta probabilidad utilizando la aplicación *Probility Distributions*. Luego de instalar la aplicación seleccionamos la distribución bajo la cual queremos calcular la probabilidad, en este caso Poisson después especificamos su parámetro y el valor particular del recorrido para el cual deseamos calcular la probabilidad, como se indica en las imágenes, así obtenemos el resultado.







Luego la probabilidad de seleccionar 3 períodos disjuntos de 10 minutos en los cuales no se registren llamadas es:

$$P(X=0) \times P(X=0) \times P(X=0) = 0.018^3$$

d-c) La función de costo es:

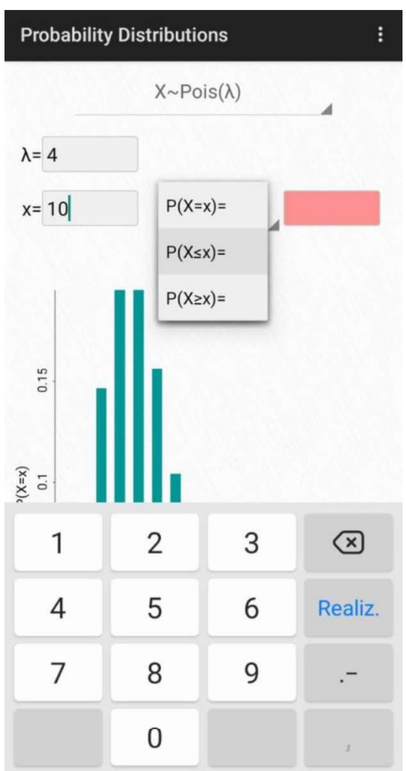
$$f(x) = \begin{cases} 2 & x \leq 10 \\ 3 & x > 10 \end{cases}$$

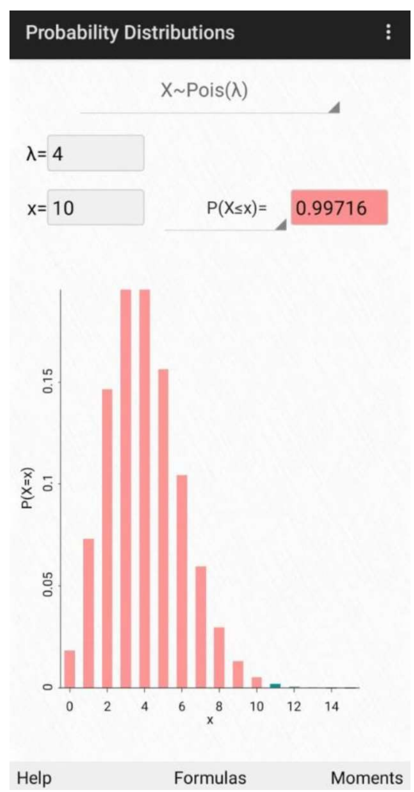
Por cada 10 minutos el costo de mantenimiento es:

$$E(C) = 2 \times P(X \leq 10) + 3 \times P(X > 10)$$

De manera que se emplea la fórmula POISSON.DIS en Excel para encontrar P(X ≤ 10) :

También puede utilizar la aplicación Probility Distributions como se indica en la siguiente imagen:





Así la P(X > 10) = 1-P(X ≤ 10) = 1-0,99716 = 0,028 y reemplazando en la función de costos y tomando esperanza de la variable:

$$E(C) = 2 \times 0{,}997 + 3 \times 0{,}028 = 2{,}003.$$

Luego el costo esperado por hora de funcionamiento es $6 \times 2{,}003 = 12{,}02$.

TRABAJO PRÁCTICO Nº 5

1) Sea X una variable aleatoria con distribución binomial: $X \sim b(n,\pi)$. Obtenga las siguientes probabilidades:
   a) $P(X = 8 \mid n = 25; \pi = 0{,}60)$ c) $P(X \geq 5 \mid n = 15; \pi = 0{,}40)$

   b) $P(X < 1{,}5 \mid n = 10; \pi = 0{,}30)$ d) $P(1 < X \leq 4 \mid n = 10; \pi = 0{,}20)$

2) El 30% de los estudiantes de una universidad privada poseen tarjeta de crédito a su nombre.
   a) Si un experimento consiste en seleccionar al azar tres estudiantes y se registra si poseen o no tarjeta de crédito. Describa las condiciones bajo las cuales puede tratarse como proceso Bernoulli y realice un diagrama de árbol para describir los resultados obtenidos.
   b) Encuentre la distribución de probabilidad para la variable Y que cuenta el número de alumnos con tarjeta de crédito entre los tres alumnos seleccionados.
   c) ¿Cuál es la probabilidad de que por lo menos dos de los tres alumnos seleccionados tengan su propia tarjeta de crédito?
   d) Calcule la esperanza y la varianza de Y.
   e) Al seleccionar diez alumnos aleatoriamente, ¿cuál es la probabilidad de que ninguno tenga su propia tarjeta de crédito?

3) Una máquina produce piezas metálicas, de las cuales el 10% son de calidad excelente.
   a) ¿Cuál es el número mínimo de piezas que deberían producirse para que la probabilidad de que haya por lo menos una pieza de calidad excelente sea mayor a ½?
   b) ¿Cuál es la probabilidad de que la primera pieza fabricada sea de calidad excelente?
   c) ¿Cuál es la probabilidad de que la tercera pieza fabricada sea la primera pieza de calidad excelente?
   d) ¿Cuál es el número esperado de piezas que debe fabricarse hasta obtener la primera pieza de calidad excelente?

4) De los 10 postulantes a un cargo gerencial en una empresa, 6 son contadores y 4 licenciados en administración de empresas, igualmente capacitados. Los entrevistadores eligen 3 de ellos para ser tomados a prueba por un mes, antes de decidir quien obtiene el cargo.
   a) ¿Con qué variable aleatoria modelaría el número de contadores elegidos para ser puestos a prueba por un mes?
   b) Calcule la probabilidad de que se elijan:
      I) Exactamente un contador.
      II) Por lo menos dos licenciados en administración de empresas.
      III) Ningún contador.
   c) Calcule el valor esperado y la varianza del número de contadores seleccionados y el número esperado de licenciados en administración de empresas.

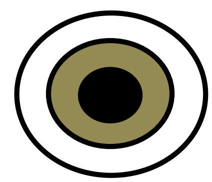
5) Sea X una variable aleatoria con distribución Poisson: $X \sim P(\beta)$. Obtenga las siguientes probabilidades:
   a) $P(X = 3;\ \beta = 4)$
   b) $P(X \leq 8;\ \beta = 3)$
   c) $P(X > 5;\ \beta = 6{,}5)$

d) $P(0{,}25 \leq X < 3;\ \beta = 0{,}5)$

6) Los pacientes llegan al azar y en forma independiente a un centro médico para ser atendidos. La frecuencia promedio de llegada es de 4 personas cada hora.
a) ¿Con qué distribución de probabilidad modelaría la cantidad de pacientes que llegan al centro médico para ser atendidos? Indique los supuestos necesarios.
b) ¿Cuál es la probabilidad de que no llegue ninguna persona en una hora?
c) ¿Cuál es la probabilidad de que lleguen a lo sumo 6 personas en una hora?
d) ¿Cuál es la probabilidad de que lleguen por lo menos 3 personas en una hora?
e) ¿Cuál es la varianza del número de personas que llegan en una hora?
f) Si se observan tres intervalos disjuntos y elegidos aleatoriamente de una hora cada uno, ¿cuál es la probabilidad de que no llegue ninguna persona en cada período de una hora?, ¿y de que en al menos uno de ellos lleguen 5 personas?

7) Una sucursal bancaria provee un seguro de celular que cubre contra robo y hurto a 3.500 de sus clientes con tarjetas de crédito. Asuma que la probabilidad de sufrir el robo de un celular es de 0,025.
a) Calcule la probabilidad exacta y usando la aproximación a la distribución Poisson de que la sucursal bancaria pague 10 pólizas durante ese año.
b) ¿Bajo qué condiciones se puede usar a la distribución Poisson como una aproximación de la binomial?

8) Una persona arroja dardos a un blanco circular que consta de tres sectores concéntricos: negro, gris y blanco, como se muestra en la figura. Esta persona acierta en cada sector con la misma probabilidad. Nunca pega con el dardo fuera del círculo externo. Si acierta en el sector negro gana 20 puntos, en el gris 10 puntos y en el blanco 5 puntos.



a) Determine la función de masa o probabilidad y la función de distribución acumulada de la variable aleatoria *X*: puntaje obtenido y represéntalas gráficamente. ¿Qué tipo de variable es *X*?
b) Calcule la esperanza y la varianza de *X*.
c) Si arroja 5 veces el dardo, ¿cuál es la probabilidad de que acierte en el círculo central al menos 3 veces? Indique si realizó algún supuesto para poder responder.

9) ¿Preferirías un examen de opción múltiple o uno en el que tuvieras que recordar todo? Si un alumno desconoce por completo el material, obtendrá cero como calificación en un examen en el que necesite memorizar. Sin embargo, si le dan cinco opciones para cada pregunta, tiene por lo menos una oportunidad en cinco de responder correctamente si elige al azar una respuesta. Si un examen de opción múltiple contiene 100 preguntas, cada una con cinco posibles respuestas y solo 1 de las opciones es correcta:
a) ¿Cuál es la calificación esperada para un alumno que contesta al azar cada pregunta, si todas las respuestas correctas tienen el mismo puntaje?
b) Use el Teorema de Chebyshev para acotar la proporción de las calificaciones que están a menos de tres desviaciones estándares de la media suponiendo que los alumnos responden al azar. ¿Aprobarán quienes están en este intervalo? Concluya si adivinar le ayudará a aprobar al alumno.

1) Enumere tres distribuciones de variables aleatorias miembros de la familia binomial.

2) Enuncie las características de un experimento para que la variable aleatoria que cuenta el número de éxitos ocurridos durante la realización del experimento sea binomial.

3) Sea $\boldsymbol{X} \sim \boldsymbol{b}(\boldsymbol{n}, \boldsymbol{\pi})$, para cada uno de los siguientes casos encontrar el valor de x tal que se verifiquen las siguientes igualdades:
   a) $P(X = x \mid 7;\ 0{,}7)=0{,}3177$.   b) $P(x < X \leq 5 \mid 9;\ 0{,}6)=0{,}4924$.

4) Sea X una variable aleatoria binomial $b(4, \pi)$. Encuentre por lo menos un valor de π para el cual la variable aleatoria binomial tenga una distribución asimétrica negativa.

5) Un examen consta de 10 preguntas de opción múltiple, con cuatro posibles respuestas a cada pregunta y sólo una es la respuesta correcta en cada caso. Si un estudiante responde todo el examen completamente al azar, calcule la probabilidad de que logre acertar por lo menos en 60% de las respuestas correctas, (contesta todas las preguntas).

6) El promedio de llamadas telefónicas que entran en un conmutador es de dos cada tres minutos, y se supone que el flujo de llamadas sigue un proceso de Poisson.
   a) ¿Cuál es la probabilidad de que entren precisamente dos llamadas durante los próximos tres minutos?
   b) Determine el valor medio y la desviación estándar del número de llamadas que entran en tres minutos.
   c) ¿Cuál es la probabilidad de que entren exactamente tres llamadas en los próximos nueve minutos?

7) Las averías de máquinas en un taller siguen una distribución de Poisson de media 2 averías/semana.
   a) Encuentre el percentil 99 de la distribución.
   b) Calcule la probabilidad de:
      i) Que no se presente ninguna avería en una semana.
      ii) Que se presenten menos de cinco averías en una semana.
      iii) Que se presenten por lo menos seis averías en un mes (cuatro semanas).

8) Una caja con 24 calculadoras contiene cuatro que están defectuosas. Si se eligen cuatro al azar sin reemplazo de esa caja, ¿cuál es la probabilidad de que:
   a) tres estén defectuosas?
   b) a lo sumo una resulte defectuosa?
   c) las cuatro sean defectuosas?

9) Suponga que la proporción de personas daltónicas en cierta población es 0,005 ¿Cuál es la probabilidad aproximada de que no haya más de 5 personas daltónicas en un grupo de 600 personas seleccionadas aleatoriamente?

10) El ochenta por ciento de los fusibles fabricados en una línea de producción pasan exitosamente la prueba de control de calidad. De 15 fusibles elegidos al azar:
   a) Obtenga el número más probable de fusibles que pasarán la prueba de control de calidad.
   b) Calcule cuál es el valor esperado y el desvío estándar de la cantidad de fusibles que pasan la prueba de control de calidad entre los 15 elegidos.
   c) ¿Tuvo que agregar algún supuesto para resolver el problema?
   d) Hay un lote de 100 fusibles de los cuales 10 son defectuosos ¿Cuál es la probabilidad de que al tomar una muestra de 15 solo 1 sea defectuoso?

11) a) Si la probabilidad de que una computadora se infecte con un virus informático es de 0,001, determine la probabilidad de que, de un total de 2.000 computadoras de una empresa, exactamente tres se infecten.

b) Indique bajo qué condiciones se puede aproximar la distribución Binomial por la distribución Poisson.

12) Se recibieron 30 facturas en el mes, de las cuales cinco tienen equivocado el valor facturado. ¿Cuál es la probabilidad de que al seleccionar sin reemplazo 10 facturas, aparezcan las cinco facturas que tienen error? ¿Y la probabilidad de que aparezcan exactamente k facturas con error? ¿Qué distribución de probabilidad es adecuada para describir el número de facturas con error entre las 10 facturas seleccionadas en este experimento?

13) Un examen de opción múltiple consta de ocho preguntas y tres respuestas a cada pregunta, de las cuales sólo una es correcta. Si un estudiante responde a cada pregunta tirando un dado equilibrado y marca la primera respuesta si obtiene 1 o 2, la segunda si obtiene 3 o 4, y la tercera si obtiene 5 o 6. ¿Cuál es la probabilidad de que obtenga cuatro respuestas correctas?

14) Un producto industrial se envía en lotes de 20 unidades. Efectuar pruebas para determinar si cada uno de los artículos tienen defectos es costoso, así que el fabricante toma muestras de su producción en lugar de probar el 100% de la misma. Un plan de muestreo elaborado para reducir al mínimo el número de artículos defectuosos que se envían a los consumidores requiere que se muestreen 5 artículos de cada lote y se rechace el lote completo si se encuentra más de un artículo defectuoso. Si un lote contiene 4 artículos defectuosos, ¿cuál es la probabilidad de que sea rechazado?

15) Una industria recibe pedidos de sus proveedores a través de fax, teléfono o internet. Suponga que el número de pedidos que llegan por cualquier medio (en horario comercial) es una variable aleatoria discreta con distribución Poisson con un promedio de 2 pedidos por hora.
   a) Calcule la probabilidad de tener más de dos pedidos entre las 9:00 y 10:00 de la mañana.
   b) En una mañana de trabajo (4 horas), ¿cuál es la probabilidad de tener más de 12 pedidos?
   c) ¿Cuál es el número esperado de pedidos en una mañana de trabajo (4hs.)?
   d) ¿Considera razonable suponer que el número de pedidos sigue una distribución Poisson? Justifique claramente su opinión.

16) Para verificar la exactitud de sus estados financieros, las empresas a menudo emplean auditores que verifiquen sus ingresos. Los empleados de la empresa se equivocan al registrar los ingresos el 5% de las veces. Suponga que un auditor revisa aleatoriamente dos ingresos, y que cada vez que hay un error él lo puede detectar.
   a) Encuentre la distribución de probabilidad para la variable Y que cuenta el número de errores detectados por el auditor.
   b) Encuentre la probabilidad de que el auditor detecte más de un error.
   c) ¿Tuvo que agregar algún supuesto para poder resolver el punto (a) y (b)?
   d) Responda el apartado a) si la probabilidad que el auditor detecte un error, cuando de hecho lo hay es del 0,9, y no detecta falsos errores.

17) Un comercio de insumos de informática recibe un envío de veinte computadoras. Antes de aceptar el envío se revisan cuidadosamente dos computadoras, si por lo menos una de las computadoras presenta algún inconveniente en el funcionamiento se rechaza el envío. Si en un lote hay dos computadoras con falla, ¿cuál es la probabilidad de que el comerciante no acepte este lote?

18) El promedio de visitas que recibe en su oficina el gerente de una empresa es de seis al día. Suponiendo distribución Poisson para el número de visitas, calcule la probabilidad de que en un día cualquiera, dicho gerente reciba en su oficina:

a) A lo sumo cuatro visitas.
b) Entre cinco y diez visitas inclusive.
c) Por lo menos 25 visitas en tres días seguidos.
d) Solo una visita un día y al día siguiente ninguna visita.
e) ¿Es razonable suponer que la distribución es Poisson? Justifique su respuesta evaluando cada supuesto.

19) Encuentre la probabilidad de que una persona que lanza una moneda obtenga:
a) la tercera cara en el séptimo lanzamiento.
b) la primera cara en el cuarto lanzamiento.
c) ¿qué distribución empleó en los apartados a) y b)?

20) Si X es una variable aleatoria uniforme discreta en$\{x_1, x_2, x_3, \ldots, x_n\}$, calcule su esperanza y su varianza.

21) Cada año (365 días) ocurren en una ciudad del litoral 365 muertes por arma de fuego.
a) ¿Cuál es la cantidad promedio de muertes por arma de fuego en una semana normal?
b) ¿Cuál es la probabilidad de que no haya muertes por armas de fuego en una semana?
c) ¿Cuál es la probabilidad de que haya dos o más muertes por armas de fuego en un día?

22) De los treinta alumnos que asisten a clase de Estadística, veinte son alumnos de la carrera de Administración y diez de Economía. Si el profesor elije un grupo de 6 estudiantes para presentar un trabajo la semana siguiente.
a) ¿Con qué distribución modelaría la variable que cuenta el número de estudiantes de Administración en el grupo elegido?
b) Calcule la probabilidad de que:
   i) se elija exactamente a dos alumnos de Administración.
   ii) se elija por lo menos dos alumnos de Economía.
   iii) ningún alumno de Administración sea elegido.
c) Calcule el número esperado de estudiantes de Administración que se eligen para presentar el trabajo y el número esperado de estudiantes de Economía.

23) Los pasajeros de una línea aérea de bajo costo llegan al azar e independientemente al mostrador de la compañía en un aeropuerto internacional. La frecuencia promedio de llegada es de diez pasajeros por minuto.
a) ¿Cuál es la probabilidad de que no lleguen pasajeros en un periodo de un minuto?
b) ¿Cuál es la probabilidad de que lleguen como máximo tres pasajeros en un periodo de un minuto?
c) ¿Cuál es la probabilidad de que no lleguen pasajeros en treinta segundos?
d) ¿Cuál es la probabilidad de que llegue al menos un pasajero en treinta segundos?

23) Suponga que el número de defectos en un rollo de tela fabricado con un cierto proceso tiene una distribución de Poisson con media de 0,4 defectos por rollo. Se inspeccionan cinco rollos de tela elegidos aleatoriamente, ¿cuál es la probabilidad de que el número total de defectos en los cinco rollos sea al menos seis? ¿Cuál es la probabilidad que haya un defecto en cada rollo?

24) Las personas que sacan número en una sucursal de un banco para ser atendidos en boxes llegan al azar y en forma independiente. La frecuencia promedio de llegada es de 6 personas cada media hora.

a) ¿Con qué distribución de probabilidad modelaría la cantidad de personas que llegan a la sucursal para ser atendidos en boxes? Indique los supuestos necesarios.

b) ¿Cuál es la probabilidad de que no llegue ninguna persona en media hora?
c) ¿Cuál es la probabilidad de que lleguen a lo sumo 5 personas en media hora?
d) ¿Cuál es la varianza del número de personas que llegan en media hora?
e) ¿Cuál es la probabilidad de que lleguen por lo menos 7 personas en una hora?
f) Si se observan tres intervalos disjuntos y elegidos aleatoriamente de media hora cada uno, ¿cuál es la probabilidad de que no llegue ninguna persona en cada período de media hora?, ¿y de que en al menos uno de ellos lleguen 5 personas?

25) Una compañía de seguros de vida asegura a 5.000 hombres de 42 años de edad. Si los estudios actuariales muestran que la probabilidad de que un hombre de 42 años muera es de 0,001, calcule la probabilidad exacta y usando la aproximación a la distribución Poisson de que la compañía pague 4 pólizas durante ese año. ¿Bajo qué condiciones se puede usar a la distribución Poisson como una aproximación de la binomial?

EJERCICIOS SUGERIDOS DEL LIBRO "PROBABILIDAD Y ESTADÍSTICA PARA INGENIERÍA Y CIENCIAS" de WALPOLE, MYERS y MYERS. (9º Ed.)

Capítulo 5:

Sección 5.1 y 5.2, página 150: ejercicio 5.7
Sección 5.3 página 157: ejercicio 5.32
Sección 5.4 a 5.5 página164: ejercicios 5.51, 5.54, 5.62, 5.68.

CAPÍTULO 6: DISTRIBUCIONES CONTINUAS DE PROBABILIDAD

Distribuciones continuas de probabilidad: distribución Normal, distribución uniforme continua, distribución Exponencial. Aproximación de probabilidades binomiales usando la distribución Normal. Corrección por continuidad.

Objetivos:

El alumno debe ser capaz de:

- Reconocer la función de densidad y parámetros de las distribuciones estudiadas.
- Representar gráficamente las funciones de densidad de diversas variables aleatorias continuas, identificando los cambios que se producen en los gráficos al variar los parámetros.
- Calcular probabilidades, esperanza y varianza conociendo la función de densidad de la variable aleatoria continua involucrada.
- Calcular valores de probabilidad de variables aleatorias binomiales usando su aproximación a la distribución normal e identificando la corrección de continuidad que debe realizarse en cada caso.
- Resolver problemas de aplicación que involucren distintas distribuciones de probabilidad.
- Valorar la importancia del conocimiento de las distintas distribuciones de probabilidad en la medida que permiten modelar y resolver problemas de la vida real.

Resumen

D1. *Distribución Uniforme Continua*: La variable aleatoria *X* se dice que tiene distribución Uniforme Continua con parámetros α y β $(X \sim U(\alpha, \beta))$si y sólo si

$$f_X(x;\alpha,\beta) = \begin{cases} \frac{1}{\beta-\alpha} \text{si} x \in [\alpha, \beta] \\ 0 \text{casocontrario} \end{cases}.$$

D2. *Distribución Exponencial*: La variable aleatoria *X* se dice que tiene distribución exponencial con parámetro β, $(X \sim Exp(\beta))$ si y sólo si

$$f_X(x;\beta) = \begin{cases} \frac{1}{\beta} e^{-x/\beta} x > 0 \\ 0 \text{otro caso} \end{cases}.$$

D3. *Distribución Normal Estándar*: La variable aleatoria *Z* se dice que tiene distribución Normal Estándar $(Z \sim N(0,1))$ si y sólo si

$$\boldsymbol{f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, z \in \mathbb{R}}..$$

D4. *Distribución Normal*: La variable aleatoria *X* se dice que tiene distribución Normal con parámetros μ y $\sigma (\sigma > 0)$, $(X \sim N(\mu, \sigma^2))$ si y sólo si

$$\boldsymbol{f_X(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, x \in \mathbb{R}}.$$

Propiedades de la distribución normal:

(1) Si $X \sim N(\mu, \sigma^2)$, entonces $\frac{X-\mu}{\sigma} \sim N(0,1)$.

(2) Si $X \sim N(\mu, \sigma^2)$, entonces $\boldsymbol{E(X) = \mu y V(X) = \sigma^2}$.

Notación: * La función de distribución acumulada de una variable aleatoria Normal Estándar (Z) se denota con $\boldsymbol{\Phi}(\blacksquare)$: $\boldsymbol{\Phi}(\boldsymbol{z}) = \boldsymbol{P}(\boldsymbol{Z} \leq \boldsymbol{z})$.

* z_α denota el valor del recorrido de la normal estándar que deja a su derecha una región de probabilidadα, es decir $\alpha = P(Z > z_\alpha)$.

T1. Si $X \sim b(n,\pi)$ entonces la distribución límite de $\frac{X-n\pi}{\sqrt{n\pi(1-\pi)}}$ cuando $n \to \infty$ es la distribución $N(0,1)$.

T2. Si $X \sim b(n,\pi)$ entonces la proporción muestral ($\hat{\pi} = X/n$) es tal que la distribución límite de $\frac{\hat{\pi}-\pi}{\sqrt{\pi(1-\pi)/n}}$ cuando $n \to \infty$ es la distribución $N(0,1)$.

Esperanza y varianza de algunas distribuciones continuas

Distribución	Parámetros	Esperanza	Varianza
$U(\alpha,\beta)$	α,β	$\frac{\alpha+\beta}{2}$	$\frac{(\beta-\alpha)^2}{12}$
$Exp(\beta)$	β	β	β^2
$N(\mu,\sigma^2)$	μ,σ	μ	σ^2

Correcciones por continuidad: Se quiere calcular probabilidades relacionadas con X, una variable aleatoria discreta que asume únicamente valores enteros, y se sabe que una transformación de ella, por algún motivo, tiene una distribución que puede aproximarse con la distribución de una variable aleatoria Y continua. En el intento de mejorar dicha aproximación, antes del cálculo de probabilidades sobre X se realizan correcciones como las que se presentan a continuación.

(1)$P(X = a) = P(a - 1/2 \leq X \leq a + 1/2)$ (2)$P(X < a) = P(X \leq a - 1/2)$
(3)$P(X \leq a) = P(X \leq a + 1/2)$ (4)$P(X > a) = P(X \geq a + 1/2)$
(5)$P(X \geq a) = P(X \geq a - 1/2)$

Aunque, por tratarse de una variable aleatoria discreta, estas correcciones no modifican las probabilidades calculadas sobre X, si mejoran la aproximación al calcular la probabilidad basada en la distribución de Y.

***Obs.*:** En situaciones donde la variable de interés es continua (Y) pero se mide por algún procedimiento que la discretiza, observando en realidad una variable X discreta, para calcular la probabilidad de observar algún evento (probabilidades sobre X) conociendo únicamente la distribución de Y, se debe corregir por continuidad siendo que la cantidad con la que se corrige (que en el caso anterior era ½) ahora debe ser la mitad de la menor unidad que se puede observar (saltos en la variable X) .

Ejemplos

Ejemplo 1: Una máquina produce esferas de metal, cuyos diámetros siguen una distribución normal con media μ = 5 cm y desviación típica σ= 0,2 cm. Para los usos que tiene destinados, la esfera se considerará inservible si su diámetro cae fuera del intervalo [4,8; 5,2] (en cm).

a) ¿Qué porcentaje de esferas inservibles produce la máquina?
b) ¿Cuál es la probabilidad de que entre 10 esferas elegidas al azar ninguna sea inservible?
c) Resolver los apartados anteriores empleando Microsoft Excel y/o la aplicación Probability Distributions.

Resolución:

Sea X el diámetro de una esfera de metal producida por la máquina. Por hipótesis del problema, X tiene distribución Normal con parámetros μ = 5 cm; σ^2 = 0,04 cm^2.
Si se realiza la transformación $Z = \frac{X-\mu}{\sigma}$ se obtiene una variable normal estándar, cuyos valores de probabilidad están tabulados.

a) Se considera inservible la esfera, cuando su diámetro X es mayor que 5,2 cm, o menor que 4,8 cm,
$P(X > 5{,}2 \text{ ó } X < 4{,}8) = P(X<4{,}8) + P(X > 5{,}2) =$
$= P\left(\frac{X-5}{0{,}2} < \frac{4{,}8-5}{0{,}2}\right) + P\left(\frac{X-5}{0{,}2} > \frac{5{,}2-5}{0{,}2}\right) = P(Z < -1) +$
$P(Z > 1)$
$P(X > 5{,}2 \vee X < 4{,}8) = \Phi(-1) + [1 - \Phi(1)] = 2\Phi(-1) = 2.0{,}15866 = 0{,}3173$
Donde $\Phi(a)$ es el valor de la probabilidad acumulada hasta el valor a bajo la curva de la distribución Normal Estándar, recordemos que por la simetría de la distribución Normal Estandar alrededor de 0, $\Phi(-a) = 1 - \Phi(a)$. y que por ser la distribución Normal continua $P(Z\leq a)=P(Z<a)=\Phi(a)$.Por lo tanto, el 31,73% de las esferas son inservibles.

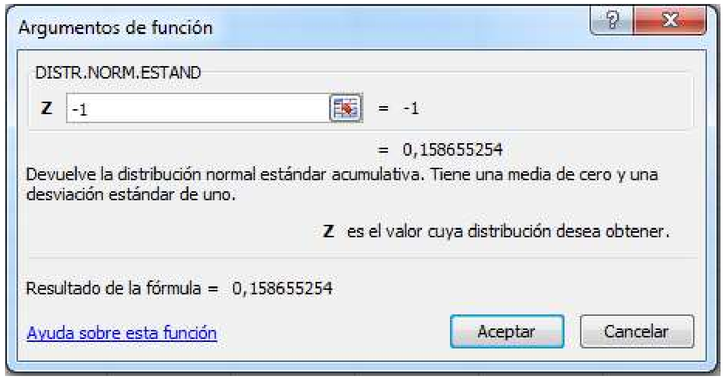
b) Sea el experimento
ε: se seleccionan 10 esferas al azar y se observa si son o no inservibles.
Se puede suponer que el proceso de seleccionar aleatoriamente esferas y observar si son o no inservibles cumple las condiciones de un proceso de Bernoulli, puesto que (i) la selección es aleatoria, lo que asegura la independencia entre los resultados de una selección y otra, (ii) siempre se observa un éxito (la esfera es inservible) o fracaso (la esfera no es inservible), y (iii) la probabilidad de éxito es constante (π= 0,3173). Consecuentemente si se define la variable aleatoria
Y: Número de esferas inservibles entre las 10 observadas,
entonces Y tendrá distribución binomial con parámetros n = 10 y π = 0,3173 y así la probabilidad de que ninguna esfera sea inservible entre las 10 seleccionadas será la probabilidad de que Y=0,

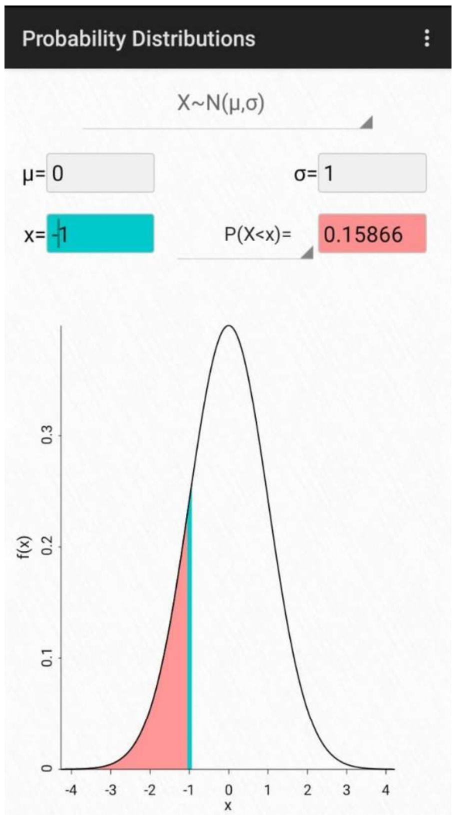
$$P(Y = 0) = b(0; 10; 0{,}3173) = \binom{10}{0} 0{,}3173^0 \times 0{,}6827^{10} = 0{,}022.$$

c-a) Para conocer cuál es el porcentaje de esferas inservibles que produce la máquina se procede a transformar X mediante un proceso de estandarización en una variable Normal Estándar Z:

$$P(X > 5{,}26 \text{ ó } X < 4{,}8) = 2.\Phi(-1)$$

Para conocer las probabilidades con Excel se recurre al Menú Fórmulas, Más Fórmulas y se usa la función DISTR.NORM.ESTAND.N:





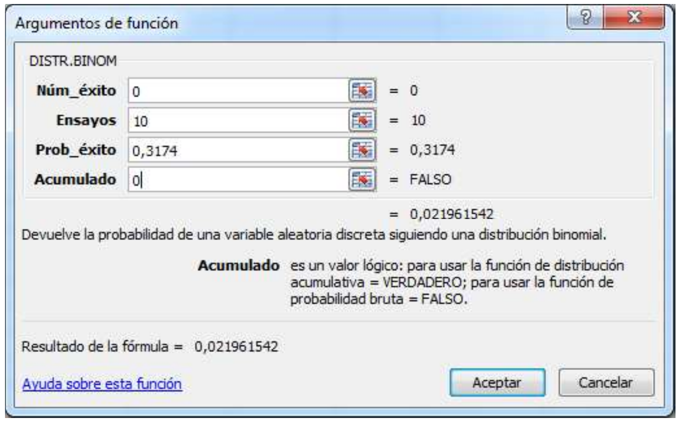
Y utilizando la aplicación Probability Distributions:

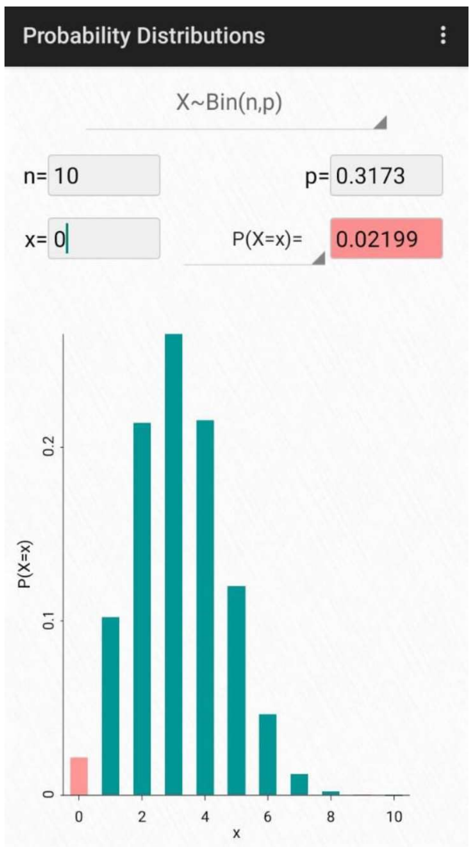
Al conocer las probabilidades, se puede calcular el porcentaje de esferas inservibles:

$P(X > 5,2 \vee X < 4,8) =$

$$= 2\Phi(-1) = 2.0,15866$$
$$= 0,3173$$

c-b) Se define a Y como el número de esferas inservibles entre las 10 observadas, y empleando Excel, el Menú Fórmulas, Más Fórmulas, DISTR.BINOM.N se puede calcular la probabilidad de que ninguna de las 10 esferas elegidas al azar sea inservible al rellenar el cuadro de la siguiente forma:





Con la aplicación

Por lo que $P(Y = 0) = b(0; 10; 0{,}3173) = 0{,}022$.

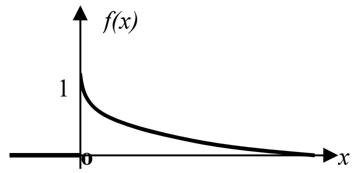
Ejemplo 2: Sea la variable aleatoria continua X, con distribución exponencial con parámetro $\beta = 1$. Representar gráficamente su función de densidad, calcular su mediana y la probabilidad de que asuma valores mayores que 1.

Resolución:

Si $\beta = 1$, la función de densidad de la variable será

$$f(x) = \begin{cases} e^{-x}, \text{ si } x > 0 \\ 0, \text{ en cualquier otro caso.} \end{cases}$$

- El gráfico de f es:



- La mediana es el valor hasta el cual se acumula la mitad de la probabilidad.

$$P(X \geq Me) = P(X \leq Me) = \frac{1}{2}.$$

De esta manera,

$$P(X \leq Me) = \int_{-\infty}^{Me} f(x)dx = \int_{0}^{Me} e^{-x}dx = -e^{-x}\Big|_0^{Me} = -e^{-Me} + 1 = \frac{1}{2}$$

$$e^{-Me} = \frac{1}{2}$$

$$-Me = ln\frac{1}{2}$$

$$Me = -\,ln\tfrac{1}{2} = 0{,}69.$$

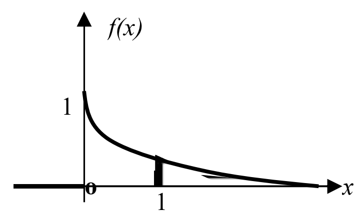
Por otro lado, $E(X) = \beta = 1 > Me = 0{,}69$. Con esto se observa que, a diferencia de lo que ocurre en la normal, no hay coincidencia entre la esperanza y la mediana porque la distribución exponencial es asimétrica positiva.

- Para calcular la probabilidad de que *X* asuma valores mayores que 1, se procede de la siguiente manera:

$$P(X > 1) = \int_1^{\infty} f(x)dx = \int_1^{\infty} e^{-x}dx = \lim_{t\to\infty} \int_1^{t} e^{-x}dx$$

$$= \lim_{t\to\infty} - e^{-x}\Big|_1^t = \lim_{t\to\infty}(-e^{-t} + e^{-1}) = e^{-1} = 0{,}37$$

Esta probabilidad representa el área comprendida entre el eje horizontal y la función de densidad, a la derecha de la recta *x* =1.



TRABAJO PRÁCTICO Nº 6

1) El tiempo de vuelo, en minutos, entre Tucumán y Buenos Aires es una variable aleatoria uniforme en [105, 115].
   a) Determine la función de densidad y la función de distribución acumulada del tiempo de vuelo. Grafíquelas.
   b) Calcule el valor esperado y la varianza.
   c) Si se elige un vuelo al azar, ¿cuál es la probabilidad de que su duración no supere los 112 minutos? Interprete gráficamente.
   d) Una aerolínea anunció que sus vuelos tienen una duración de 110 minutos. ¿Cuál es la probabilidad de que entre los próximos cuatro vuelos de esta aerolínea exactamente dos lleguen en el tiempo anunciado? Indique los supuestos realizados para poder calcular la probabilidad.

2) En un mismo gráfico represente las densidades de los grupos de distribuciones normales $X_1, X_2 y X_3$definidas en cada apartado:
   a) $X_1, X_2 y X_3$ tienen la misma esperanza μ= 4, pero desviaciones típicas distintas: σ_1=0,5 σ_2=1, σ_3=2.

b) $X_1, X_2 y X_3$ tienen la misma esperanza μ= -4, pero desviaciones típicas distintas: σ_1=0,5 σ_2=1, σ_3=2.
c) $X_1, X_2 y X_3$ tienen el mismo desvío estándar σ= 2, pero sus esperanzas son μ_1=-2, μ_2=0 y μ_3=10.

3) Sea Z una variable aleatoria normal estándar. Determine los valores de k tales que:
a) $P(Z > 2{,}65) = 1 - k$
b) $P(Z > k) = 0{,}0548$
c) $P(Z \leq k) = 0{,}9265$
d) $P(Z < k) = 0{,}95$
e) $P(|Z| < k) = 0{,}0357$
f) $P(|Z| > k) = 0{,}6892$
g) $P(k < Z < 2) = 0{,}25$

4) Un Municipio compra lámparas para el alumbrado público que duran en promedio 7.500 horas encendidas. Se sabe que la duración de las lámparas sigue una distribución normal con desvío estándar igual a 1.000.
   a. ¿Cuál es la probabilidad de que una lámpara dure menos de 6.000 horas?
   b. ¿Cuál es la probabilidad de que una lámpara dure más de 8.000 horas?
   c. Encuentre el primer y tercer cuartil de la duración de las lámparas.

5) Una industria multinacional estima que el 65% de sus empleados demoran menos de 20 minutos en ensamblar el producto que fabrican.
   a) Si se toma una muestra aleatoria de 300 empleados de esta industria, ¿sería raro encontrar que al menos 210 ensamblan el producto en menos de 20 minutos? Realice el cálculo exacto y la aproximación usando la distribución normal.
   b) ¿Bajo qué condiciones se puede usar a la distribución normal como una aproximación de la binomial?

6) El tiempo que una persona debe esperar un taxi en determinado punto de la ciudad a las 22 horas, es una variable aleatoria exponencial con un promedio de 5 minutos.
   d. Indique la función de densidad y la función de distribución acumulada de esta variable y grafíquelas.
   e. ¿Cuál es la probabilidad de que deba esperar el taxi entre 4 y 6 minutos? , ¿y más de 4 minutos?
   f. Indique la varianza y el desvío estándar de la variable.
   g. Si una persona esperó el taxi en ese lugar y horario de lunes a viernes de una semana, ¿cuál es la probabilidad de que en al menos dos días el taxi haya demorado entre 4 y 6 minutos? Indique si tuvo que realizar algún supuesto para resolver.

7) El rendimiento promedio del cultivo de maíz en Argentina es 7,68 toneladas por hectárea. Suponga que el rendimiento se distribuye normalmente y que su desviación estándar es 1,46 toneladas por hectárea.
   a) ¿Qué porcentaje de los cultivos tienen un rendimiento superior a 9 toneladas por hectárea?
   b) ¿Cuál es el rendimiento máximo del 15% con menor rendimiento?
   c) ¿Cuál es el valor del rendimiento del cultivo tal que solo el 10% tenga un rendimiento superior?

8) Una máquina automática hace pernos que deben tener 3 pulgadas de longitud. Si en realidad las longitudes de los pernos de 3 de pulgadas se distribuyen uniformemente en el intervalo (2,5; 3,5) pulgadas.

a) ¿Cuál es la probabilidad de que uno de los pernos, elegido al azar, de un lote determinado, tenga una longitud que:
   i. esté entre 2,75 y 3,25 pulgadas?
   ii. sea mayor que 3,25 pulgadas?
   iii. sea exactamente igual a 3 pulgadas?
b) Entre los próximos cinco pernos que se fabriquen, ¿cuál es la probabilidad de que al menos cuatro tengan una longitud mayor que 3,25 pulgadas? ¿Qué supuesto agregó para resolver este apartado? ¿Es razonable hacer este supuesto?

9) La vida útil, en meses, de una lámpara es una variable aleatoria continua con la siguiente distribución exponencial

$$f(x) = \begin{cases} \dfrac{1}{50}\, e^{-\frac{x}{50}} \; para\ x \geq 0 \\ 0\ en\ cualquier\ otro\ punto \end{cases}$$

a) ¿Cuál es la vida promedio del dispositivo?
b) ¿Cuál es la probabilidad de que el dispositivo falle en los primeros treinta meses de funcionamiento?
c) ¿Cuál es la probabilidad de que el dispositivo funcione por lo menos 75 meses antes de que falle?
d) ¿Por cuántos meses de vida debe el fabricante garantizar sus lámparas, si desea que la probabilidad de que una lámpara, elegida al azar, sobreviva el periodo de garantía sea de 0,80?

Ejercicios Adicionales

1) Calcule la probabilidad de que una v.a. normal asuma valores dentro de k desvíos estándares de la media, para k =1,2 y 3. Compare con la cota dada por el Teorema de Chebyshev.

2) El porcentaje de grasa de cierto queso marca X es una variable aleatoria que sigue una distribución normal con media μ = 4,35% y desviación estándar σ= 0,59%. Para una porción de este queso, elegida al azar, calcule la probabilidad de que el contenido de grasa sea:
   a) de por lo menos 5,5%.
   b) entre 4 y 5%.

3) Sea X el número de respuestas "*si*" obtenidas en 50 encuestas, en las que una respuesta "*si*" es igualmente probable que una respuesta "*no*".
   a) ¿Cuál es la distribución de probabilidad exacta de X? ¿Cuál es la distribución de probabilidad aproximada de X? ¿Cuál es la distribución de probabilidad de la proporción de respuestas afirmativas?
   b) Calcule la esperanza y varianza de cada una de las distribuciones encontradas en el punto a).
   c) Calcule la probabilidad aproximada de que X sea a lo sumo 30.

4) El sesenta por ciento de las familias de cierta ciudad tiene vivienda propia (casa o departamento). Si se eligen al azar 500 familias de esa ciudad, ¿cuál es la probabilidad de que:

a) menos de 315 tengan vivienda propia?
b) entre 290 y 310, inclusive, posean vivienda?
c) a lo sumo 320 tengan vivienda propia?

5) Una de las primeras aplicaciones de la curva normal fue debida al astrónomo F. W. Bessel en 1818 que comprobó que los errores de medida de 300 medidas astronómicas coincidían con bastante aproximación con lo previsto por Gauss con la curva normal. Suponiendo que la esperanza de estos errores es cero y la desviación estándar 4 grados, calcular:
a) La probabilidad de que un error no sea mayor que 6 grados, ya sea por exceso o por defecto.
b) La probabilidad de que los errores por defecto superen los 8 grados.
c) Si llamamos pequeños a los errores menores que 7 grados ya sea por exceso o por defecto, calcular el número esperado de errores pequeños en 300 observaciones.

6) Un mecanismo de aire acondicionado funciona con base en cinco componentes independientes, y la vida útil de cada uno sigue una distribución exponencial con media 5 (en años). Para que el mecanismo de aire acondicionado pueda funcionar se requiere que por lo menos dos de sus cinco componentes aún sirvan. Calcule la probabilidad de que el mecanismo de aire acondicionado continúe funcionando después de ocho años.

7) Los paquetes grandes de café marca X señalan en la etiqueta un contenido neto que debería ser de 4 Kg. En el departamento de empaque saben que el contenido neto en Kg. es ligeramente variable y han estimado que la desviación estándar es de $\sigma = 0{,}04$ Kg. Además, aseguran que sólo el 2% de los paquetes contienen menos de 4 Kg. Si se supone que el peso de los paquetes tiene una distribución normal, ¿cuál es el contenido neto promedio de los paquetes?

8) Se sabe que el 30% de los residentes de una ciudad al momento de comprar prefieren un teléfono celular color gris más que cualquier otro color disponible, ¿cuál es la probabilidad aproximada de que entre los siguientes 1.000 teléfonos que se vendan en esta ciudad:
a) entre 250 y 350, inclusive sean grises?
b) menos de 280 sean grises?
c) menos de 680 sean de otro color diferente al gris?
d) exactamente 300 sean grises?

9) El período de tiempo que transcurre hasta que un empleado falta a su trabajo en una empresa particular tiene una distribución exponencial con media de 20 días.
a) ¿Cuál es la probabilidad de que un empleado no falte en los próximos 10 días?
b) Si se toma una muestra aleatoria de 5 empleados ¿cuál es la probabilidad que sólo uno de ellos falte a su trabajo en un período de 10 días?

10) Una oficina municipal que procesa permisos para remodelación de edificios tiene una política que indica que el permiso se entregará sin costo si no está listo al final de 5 días hábiles, a partir de la fecha de solicitud. Suponga que los tiempos de procesamiento tienen distribución normal. Si el proceso tiene una media de 3 días y una desviación estándar de un día, ¿qué proporción de los permisos será gratis?
En el proceso descrito anteriormente, para disminuir la proporción de permisos sin costo, ¿qué le conviene a la oficina municipal, reducir el promedio a 2 días o la desviación estándar a 0,75 días? Explique.

11) Una fábrica usa un producto a granel. La cantidad que usa un día se puede describir mediante una distribución exponencial con $\beta = 4$ (medida en toneladas).

a) Calcule la probabilidad de que la planta use más de 4 toneladas determinado día.
b) ¿Qué cantidad del producto se debe almacenar para ser usado determinado día si se espera que la probabilidad de agotar la existencia sea de 0,05?
c) Determine la función de distribución acumulada y represéntela gráficamente.

12) Delta Airlines publica un tiempo de vuelo de 2 horas y 5 minutos entre dos ciudades. Suponga que el tiempo real de vuelo se distribuye uniformemente entre 2 horas y dos horas 20 minutos.
a) ¿Cuál es la probabilidad de que el vuelo llegue más de 10 minutos tarde?
b) En los próximos 10 vuelos ¿cuál es la probabilidad de que ninguno de ellos se retrase más de 10 minutos? Explicite los supuestos realizados.

13) La vida útil en horas de un dispositivo electrónico es una variable aleatoria normal con esperanza 50 y desvío estándar 13.
a) ¿Cuál es la probabilidad de que el dispositivo falle en las primeras 25 horas?
b) Si una empresa A compra un lote de 3 dispositivos, ¿cuál es la probabilidad de que ningún dispositivo falle en las primeras 25 horas?
c) El fabricante de los dispositivos gana $20 en un dispositivo que dura más de 25 horas y gana $5 si el dispositivo dura menos de 25 horas, puesto que tienen este período de garantía. Indique la ganancia esperada del fabricante en la venta realizada a la empresa A.

14) La cooperativa proveedora de electricidad en la ciudad de Yerba Buena ha realizado durante el último año una campaña para disminuir el consumo de energía domiciliaria. Si el usuario mantiene el consumo por debajo de estándares establecidos se beneficia con la aplicación de una tarifa promocional. Un reporte realizado por la gerencia establece que sólo el 30% de los residentes en el municipio se beneficiarán por la aplicación de la tarifa promocional.
a) Calcule la probabilidad de que, si se selecciona aleatoriamente cinco usuarios, todos sean beneficiados por el descuento.
b) Se ha emitido la facturación de doscientos consumidores, ¿cuál es la probabilidad de que más de 50 consumidores estén beneficiados por la tarifa promocional?
c) Indique los supuestos que debió realizar para responder el inciso anterior.
d) Si se sabe que el consumo en kw/h tiene una distribución normal con media 450 y desvío estándar 160, ¿cuál es el rango de consumo de los usuarios beneficiados?

15) Sea Z una variable aleatoria normal estándar. Determine los valores de k tales que:
a) $P(Z > k) = 0{,}3264$ b) $P(Z < 1{,}96) = 1 - k$
c) $P(Z \geq k) = 0{,}996$ d) $P(Z < k) = 0{,}025$
e) $P(|Z| > k) = 0{,}05$ f) $P(|Z| < k) = 0{,}95$
g) $P(-0{,}58 < Z < k) = 0{,}3138$

16) Durante un largo período se observó que los gastos semanales de una empresa, por concepto de mantenimiento y reparaciones, se aproximan a una distribución normal con una media de $400 y una desviación estándar de $20.
a) Si se destina un presupuesto de $450 para la siguiente semana, ¿cuál es la probabilidad de que los costos superen el presupuesto?
b) ¿Qué presupuesto debe destinarse a reparaciones y mantenimiento semanales con el fin de que la probabilidad de superar el presupuesto cierta semana sea de sólo 0,1?
c) Encuentre el primer y tercer cuartil para el gasto semanal.
d) Si se destina un presupuesto de $450 para cada una de las 5 semanas siguientes, ¿cuál es la probabilidad de que los costos no superen el presupuesto en 3 de las 5 semanas?

17) Una empresa de bebidas gaseosas considera que el 31% de los consumidores la prefieren. Si se tomara una muestra aleatoria de 500 consumidores, ¿sería raro encontrar que 232 o más de los entrevistados la prefieren?

18) En un mismo gráfico represente las densidades de los grupos de distribuciones normales X_1 y X_2 definidas en cada apartado:
   a) X_1 y X_2 tienen la misma esperanza μ=5, pero desviaciones típicas distintas: σ_1=1, y σ_2=2, respectivamente.
   b) X_1 y X_2 tienen el mismo desvío estándar σ= 4, pero sus esperanzas difieren: μ_1=8 y μ_2=10, respectivamente.

19) Sea X una variable aleatoria uniformemente distribuida en $[-\alpha,\alpha]$ con $\alpha>0$. Cuando sea posible, determine α de modo que se satisfaga lo siguiente:
   a) $P(X> 1) = 1/3$ b) $P(X< 1/8) = 0{,}7$ c) $P(|X| < 1) = P(|X| > 1)$

20) Una variable aleatoria X tiene la distribución uniforme continua con parámetros α y β. Para el caso particular en que $\alpha = 3$ y $\beta = 7$, determine:
   a) La función de densidad y represéntela gráficamente.
   b) La función de distribución acumulada.
   c) $E(X)$ y $V(X)$.

21) El contenido neto de un paquete de azúcar es una variable aleatoria uniforme entre 990 y 1010 gramos.
   a) Determine las funciones de densidad y de distribución acumulada del contenido neto y grafíquelas.
   b) Calcule $E(X)$ y $V(X)$.
   c) ¿Cuál es la probabilidad de que un paquete pese menos de un kilo?
   d) Los compradores tienen derecho a devolver un paquete de azúcar si detectan que el peso es menor a 995 gramos, ¿cuál es la probabilidad de que un cliente solicite el cambio de un paquete?

22) El Ministerio de Educación ha estimado que sólo el 50% de los alumnos que han finalizado el cursado de la escuela secundaria se presentan al año siguiente para rendir algún examen de las materias pendientes.
   a) Si se toma una muestra aleatoria de seiscientos estudiantes que finalizaron el cursado de la secundaria en 2022 con materias pendientes, ¿sería raro encontrar que menos de 280 estudiantes se presentaron a rendir en el año 2023?
   b) ¿Bajo qué condiciones se puede usar a la distribución normal como una aproximación de la binomial?

23) El tiempo (en minutos) durante el cual no hay clientes esperando en la caja de un supermercado sigue una distribución de probabilidad exponencial con 1,2 minutos de promedio.
   a) Indique la función de densidad de probabilidad para esta distribución.
   b) ¿Cuál es la probabilidad de que el siguiente cliente llegue entre 0,5 y 1,0 minutos después de salir el último cliente?
   c) ¿Cuál es la probabilidad de que la fila de una caja esté vacía durante más de un minuto?

24) El contenido neto (en gramos) de un frasco de café marca Sol, que en su etiqueta indica un contenido de 200 gramos, es en realidad una variable aleatoria uniforme en [195, 205].

a) Determine la función de densidad y la función de distribución acumulada del contenido neto. Grafíquelas.
b) Calcule el valor esperado y la varianza.
c) Si se elige un frasco al azar, ¿cuál es la probabilidad de que su contenido neto sea de por lo menos 200 gramos? Interprete gráficamente.
d) Un cliente puede solicitar el cambio del frasco si su contenido es inferior a 198 gramos. ¿Cuál es la probabilidad de que entre los próximos cinco compradores de este producto exactamente dos pidan un cambio de frasco? Indique los supuestos realizados para poder calcular la probabilidad.

25) La media de la deuda hipotecaria, en familias estadounidenses cuyo jefe tiene 35 años de edad o menos, es de 63.000 dólares. Suponga que esta deuda se distribuye normalmente y que su desviación estándar es 15.000 dólares.
a) ¿Cuál es la máxima deuda hipotecaria del 10% con menor deuda?
b) ¿Qué porcentaje de las familias tienen deuda hipotecaria superior a 68.000 dólares?
c) ¿Cuál es el valor de la deuda tal que sólo el 5% sea superior a ella?

EJERCICIOS SUGERIDOS DEL LIBRO “PROBABILIDAD Y ESTADÍSTICA PARA INGENIERÍA Y CIENCIAS” de WALPOLE, MYERS y MYERS. (9º Ed.)

Capítulo 6:

Secciones 6.1 a 6.4, página 185: ejercicios 6.8, 6.15 y 6.22.
Sección 6.5, página 193: ejercicios 6.32 y 6.37.

CAPÍTULO 7: DISTRIBUCIONES MUESTRALES

Muestreo aleatorio. Distribuciones muestrales. Teorema central del límite.

Objetivos:

El alumno debe ser capaz de:

- Definir población y muestra.
- Reconocer las condiciones para que una muestra particular sea aleatoria.
- Reconocer la diferencia entre estadístico, estimador y estimación.
- Identificar la distribución de algunos estadísticos.
- Analizar bajo qué condiciones es válido aplicar el Teorema Central del Límite.
- Conocer la importancia del Teorema Central del Límite en la Estadística.

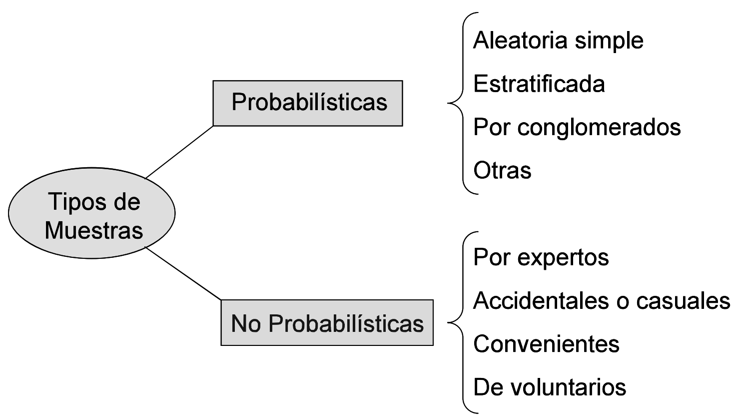
Resumen

D1. *Población*: Es el conjunto de todas las unidades/objetos en los cuales se quiere conocer el comportamiento de alguna característica de interés.

Observación 1: el conjunto de los valores asumidos por la característica de interés en la población se denomina *Población Estadística*, esta puede ser modelada por alguna distribución de probabilidad.

D2. *Muestra*: Es un subconjunto de la población.

Observación 2: Dependiendo el proceso por el cual se selecciona una muestra, las mismas se pueden clasificar según el siguiente esquema



D3. *Muestra Aleatoria*: Se dice que $\boldsymbol{X_1, X_2, \ldots, X_n}$ constituye una muestra aleatoria (m.a.) de tamaño *n* de una población *X*, si $\boldsymbol{X_1, X_2, \ldots, X_n}$ son variables aleatorias independientes e idénticamente distribuidas como *X*. Si $f(x)$ es la función de distribución de probabilidades de *X*, la distribución de probabilidades conjunta de la m.a. será:

$$\boldsymbol{f(x_1, x_2, \ldots, x_n) = f_X(x_1) f_X(x_2) \ldots f_X(x_n)}.$$

Observación 3: Cuando se habla de conocer alguna información sobre alguna característica de interés en una población, y si se supone que esa población (variable aleatoria) sigue alguna distribución, el problema se reduce a conocer los parámetros que caracterizan a esa distribución particular.

D4. *Estadístico*: Sea $\boldsymbol{X_1, X_2, \ldots, X_n}$ m.a., un estadístico (*T*) es cualquier función de la m.a., es decir $\boldsymbol{T = g(X_1, X_2, \ldots, X_n)}$.

Notación: Generalmente los parámetros se denotan con letras griegas μ, θ, π, σ y los estadísticos con letras de imprenta mayúsculas o cuando ya están direccionados a estimar algún parámetro se los denota por $\hat{\mu}, \hat{\theta}, \hat{\pi}, \hat{\sigma}$.

Algunos estadísticos importantes
Media muestral: $\boldsymbol{\bar{X} = n^{-1} \sum_{i=1}^{n} X_i}$

Mediana muestral: $\boldsymbol{Me} = \begin{cases} \boldsymbol{X_{[(n+1)/2]}} & \text{si } \boldsymbol{n} \text{ es impar} \\ \dfrac{\left(X_{[\frac{n}{2}]} + X_{[\frac{n}{2}+1]}\right)}{2} & \text{si } \boldsymbol{n} \text{ es par} \end{cases}$, siendo $\boldsymbol{X_{[1]}, X_{[2]}, \ldots, X_{[n]}}$ la m.a. de tamaño *n* acomodada en orden de magnitud creciente.

Moda: Es el valor de la muestra que ocurre con mayor frecuencia

Rango: $R = X_{[n]} - X_{[1]}$ siendo $X_{[1]}$ y $X_{[n]}$ las observaciones más pequeña y más grande de la muestra respectivamente.
Varianza Muestral: $S^2 = (n-1)^{-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$
Desviación estándar muestral: es la raíz cuadrada positiva de la varianza muestral

$$\boldsymbol{S = \sqrt{(n-1)^{-1} \sum_{i=1}^{n} (X_i - \bar{X})^2}}$$

T1. $\boldsymbol{S^2 = [n(n-1)]^{-1} \{ n \sum_{i=1}^{n} {X_i}^2 - (\sum_{i=1}^{n} X_i)^2 \}}$
D5. *Distribución muestral*: Es la distribución de probabilidad de un estadístico.
T2 *Reproducibilidad de la distribución Normal***:** Si $\boldsymbol{X_1, X_2, \ldots X_n}$ son variables aleatorias independientes que tienen distribuciones normales con medias $\boldsymbol{\mu_1, \mu_2, \ldots \mu_n}$ y varianzas $\boldsymbol{\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2}$ .Entonces: $\boldsymbol{Y = a_1 X_1 + a_2 X_2 + \cdots + a_n X_n}$
tiene una distribución Normal con media $\boldsymbol{\mu_Y = a_1 \mu_1 + a_2 \mu_2 + \cdots + a_n \mu_n}$ y varianza $\boldsymbol{\sigma_Y^2 = a_1^2 \sigma_1^2 + a_2^2 \sigma_{12}^2 + \cdots + a_n^2 \sigma_n^2}$.
T3. Teorema del límite central: Si$\bar{X}$ es la media de una muestra aleatoria de tamaño *n* que se toma de una población con media μ y varianza finita σ^2, entonces la forma límite de la distribución de

$$\boldsymbol{Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}}$$

conforme$\boldsymbol{n} \to \infty$, es la distribución normal estándar *n*(*z*;0,1).
Observación 4: Para conocer las distribuciones muestrales de otros estadísticos que se utilizarán en esta materia consulte el Anexo I.

Ejemplos

Ejemplo 1: Se extrae una muestra aleatoria de la población $\{1, 2, 3\}$.
a) Determine la distribución muestral de $\overline{X}$en los siguientes casos:
   i) $n = 2$, muestreo con reposición.
   ii) n= 2, muestreo sin reposición.
   iii)n= 3, muestreo con reposición.

b) Responde las siguientes preguntas:
   a) ¿La distribución muestral depende del tamaño de la muestra?
   b) ¿La distribución muestral depende del tipo de muestreo realizado?

Resolución:

La **distribución muestral** es la distribución de probabilidad de un estadístico. Para conocerla, en situaciones donde la población es finita y el tamaño muestral es pequeño, es suficiente con enumerar todas las posibles muestras con sus respectivas probabilidades y el valor que asume el estadístico para cada muestra.

a) i) La siguiente tabla presenta la enumeración de todas las posibles muestras de tamaño *n*=2, con reposición, conjuntamente con el valor asumido por el estadístico al ser evaluado en cada muestra.

Muestra $\{s_1, s_2\}$	{1, 1}	{1, 2}	{1, 3}	{2, 1}	{2, 2}	{2, 3}	{3, 1}	{3, 2}	{3, 3}
$\bar{x}$	1	1,5	2	1,5	2	2,5	2	2,5	3

Se debe observar que en este caso, cada muestra tiene la misma probabilidad de aparecer. $P\{s_1, s_2\} = 1/9$para cualquier $\{s_1, s_2\}$. Consecuentemente, la distribución de probabilidad del estadístico es la siguiente:

$\bar{x}$	1	1,5	2	2,5	3
$f(\bar{x})$	1/9	2/9	1/3	2/3	1/9

ii) La siguiente tabla presenta la enumeración de todas las posibles muestras de tamaño *n*=2, sin reposición, conjuntamente con el valor asumido por el estadístico al ser evaluado en cada muestra.

Muestra	{1, 2}	{1, 3}	{2, 1}	{2, 3}	{3, 1}	{3, 2}
$\bar{x}$	1,5	2	1,5	2,5	2	2,5

Considerando que cada muestra tiene probabilidad 1/6 de aparecer, la distribución de probabilidad del estadístico es la siguiente:

$\bar{x}$	1,5	2	2,5
$f(\bar{x})$	1/3	1/3	1/3

iii) La siguiente tabla presenta la enumeración de todas las posibles muestras de tamaño *n*=3, con reposición, conjuntamente con el valor asumido por el estadístico al ser evaluado en cada muestra.

Muestra	$\bar{x}$	Muestra	$\bar{x}$
{1, 1 ,1}	1	{2, 2, 3}	7/3
{1, 1, 2}	4/3	{2, 3, 1}	2
{1, 1, 3}	5/3	{2, 3, 2}	7/3
{1, 2, 1}	4/3	{2, 3, 3}	8/3
{1, 2, 2}	5/3	{3, 1, 1}	5/3
{1, 2, 3}	2	{3, 1, 2}	2
{1, 3, 1}	5/3	{3, 1, 3}	7/3
{1, 3, 2}	2	{3, 2, 1}	2
{1, 3, 3}	7/3	{3, 2, 2}	7/3
{2, 1, 1}	4/3	{3, 2, 3}	8/3
{2, 1, 2}	5/3	{3, 3, 1}	7/3
{2, 1, 3}	2	{3, 3, 2}	8/3
{2, 2, 1}	5/3	{3, 3, 3}	3
{2, 2, 2}	2		

Considerando que cada muestra tiene probabilidad 1/27 de aparecer, la distribución de probabilidad del estadístico es la siguiente:

$\bar{x}$	1	4/3	5/3	2	7/3	8/3	3
$f(\bar{x})$	1/27	1/9	2/9	7/27	2/9	1/9	1/27

b) De los incisos i) y iii) anteriores se puede observar que si cambia el tamaño de la muestra puede cambiar la distribución del estadístico. Entonces, la distribución muestral sí depende del tamaño de muestra seleccionado.

De los incisos i) y ii) es evidente que si cambia el tipo de muestreo puede cambiar la distribución del estadístico. Entonces, la distribución muestral también depende del tipo de muestreo realizado.

Ejemplo 2: Se extrae una muestra aleatoria de tamaño 54 de una población con distribución de probabilidad Binomial con parámetros $n = 3$, $\pi=1/3$.
a) ¿Cuál es la probabilidad de que la media muestral supere a 10/9? ¿Qué teorema permite asegurar este resultado?
b) ¿Se llegaría a un resultado diferente si la población fuera normal con $\mu = 1, \sigma^2 = 2/3$?
c) ¿Podría usar el mismo teorema utilizado en a) si la muestra fuera de tamaño 9?

Resolución:

a) Para poder calcular una probabilidad necesitamos conocer previamente la distribución muestral del estadístico. Cuando el tamaño muestral (n') es grande, el Teorema Central del Límite asegura que la distribución de la media se aproxima a una normal. El tamaño de muestra n'= 54 es un tamaño razonablemente grande para usar tal aproximación.

$$E(\overline{X}) = E(X) = n\pi = 3 \times \frac{1}{3} = 1, V(\overline{X}) = \frac{V(X)}{n'} = \frac{n\pi(1-\pi)}{n'} = \frac{3 \times 1/3 \times 2/3}{54} = \frac{1}{81}$$

Para dejar en claro la notación empleada, n' denota el tamaño muestral, es decir el tamaño de la muestra tomada que en este caso es 54. Por otra parte, n denota el parámetro de la distribución Binomial que está siendo muestreada, en este caso n=3.

$$P\left(\overline{X} > \frac{10}{9}\right) = P\left(\frac{\overline{X} - 1}{\sqrt{1/81}} > \frac{10/9 - 1}{1/9}\right) \cong P(Z > 1) = 0{,}1587$$

b) En el caso en que la población de origen fuera normal con $\mu = 1, \sigma^2 = \frac{2}{3}$, también se obtiene

$$E(\overline{X}) = E(X) = 1, Var(\overline{X}) = \frac{Var(X)}{n'} = \frac{\sigma^2}{n'} = \frac{2/3}{54} = \frac{1}{81},$$

y se procede a realizar la misma estandarización realizada en a), obteniéndose el mismo resultado 0,1587.
La diferencia en este caso es que, si la muestra proviene de una población normal, la media muestral tiene distribución normal, por lo tanto, la probabilidad antes calculada sería exactamente igual a 0,1587, no aproximadamente igual a ese valor.

c) No se podría haber usado el Teorema Central del Límite para un tamaño muestral 9, que es muy pequeño como para considerar válida la aproximación de las probabilidades con la distribución asintótica.

Ejemplo 3: La primera tarea de un curso introductorio de programación consiste en codificar, editar y correr un breve programa. La experiencia indica que el tiempo que les insume esta tarea a los estudiantes es una variable aleatoria uniforme en el intervalo (24 minutos, 48 minutos).
a) Genere con su computadora 1.000 muestras de tamaño 75 de esta población y construya la distribución de medias.

b) Calcule la media y varianza de la distribución muestral construida. Compare estos resultados con los valores de los parámetros poblacionales.
c) ¿Qué observa de la forma de la distribución muestral que construyó?

Resolución:

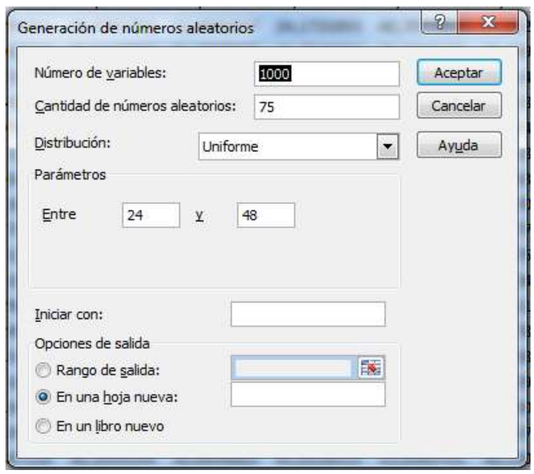
Para resolver el ejercicio usando Microsoft Excel se requiere de la herramienta Análisis de Datos ubicada en el Menú Datos; en caso de no contar con la misma a continuación se detallan los pasos para su instalación:

1° Abrir Excel

2° En el Menú Archivo, seleccionar Opciones y luego Complementos (en la parte inferior del cuadro de dialogo),

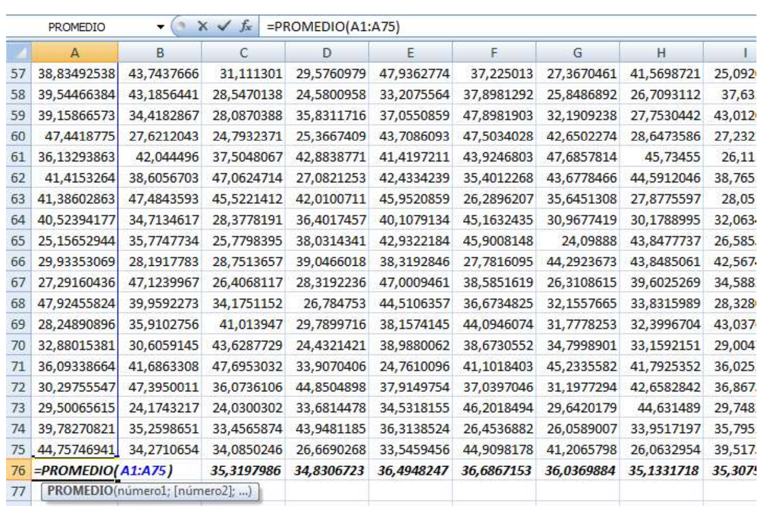
3° En complemento de aplicaciones inactivas seleccionar Herramientas para análisis y Aceptar.

a) Para generar las 1.000 muestras de tamaño 75 se selecciona en el Menú Datos, Análisis de Datos y luego Generación de Números Aleatorios. En la ventana de generación de números aleatorios se llena el cuadro con los datos del ejercicio:



El software genera las 1.000 muestras de tamaño 75 y las presenta en una hoja nueva, donde cada columna es una muestra.

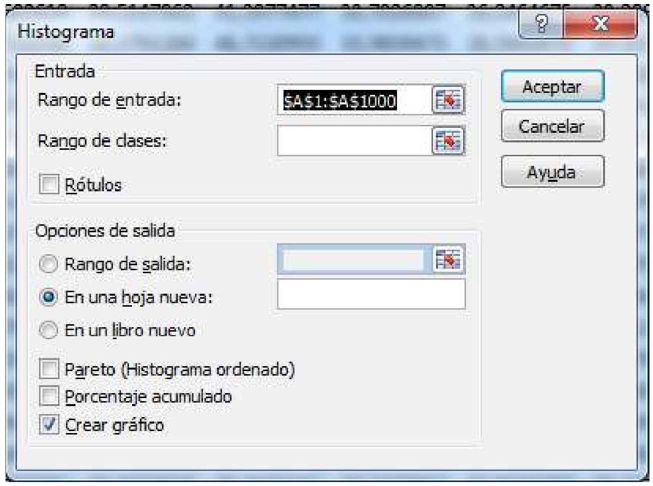
Para construir la distribución muestral de la media se calcula la media de cada muestra, utilizando la función PROMEDIO. A continuación, se muestra la pantalla de Excel con las muestras generadas y las medias calculadas resaltadas:



PROMEDIO =PROMEDIO(A1:A75)

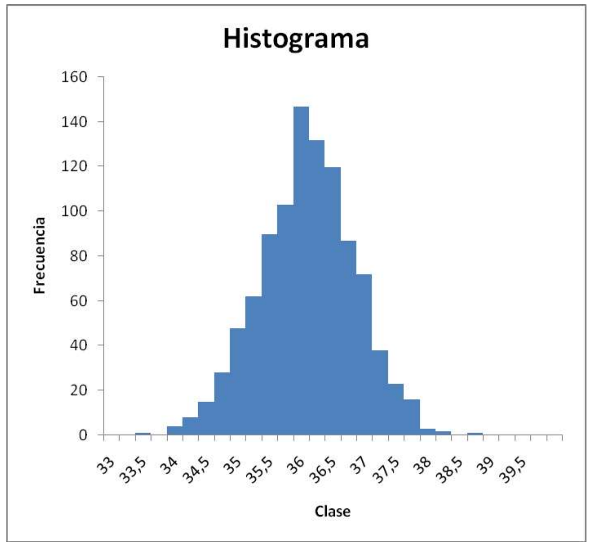
	A	B	C	D	E	F	G	H	I
57	38,83492538	43,7437666	31,111301	29,5760979	47,9362774	37,225013	27,3670461	41,5698721	25,092
58	39,54466384	43,1856441	28,5470138	24,5800958	33,2075564	37,8981292	25,8486892	26,7093112	37,63
59	39,15866573	34,4182867	28,0870388	35,8311716	37,0550859	47,8981903	32,1909238	27,7530442	43,012
60	47,4418775	27,6212043	24,7932371	25,3667409	43,7086093	47,5034028	42,6502274	28,6473586	27,232
61	36,13293863	42,044496	37,5048067	42,8838771	41,4197211	43,9246803	47,6857814	45,73455	26,11
62	41,4153264	38,6056703	47,0624714	27,0821253	42,4334239	35,4012268	43,6778466	44,5912046	38,765
63	41,38602863	47,4843593	45,5221412	42,0100711	45,9520859	26,2896207	35,6451308	27,8775597	28,05
64	40,52394177	34,7134617	28,3778191	36,4017457	40,1079134	45,1632435	30,9677419	30,1788995	32,063
65	25,15652944	35,7747734	25,7798395	38,0314341	42,9322184	45,9008148	24,09888	43,8477737	26,585
66	29,93353069	28,1917783	28,7513657	39,0466018	38,3192846	27,7816095	44,2923673	43,8485061	42,567
67	27,29160436	47,1239967	26,4068117	28,3192236	47,0009461	38,5851619	26,3108615	39,6025269	34,588
68	47,92455824	39,9592273	34,1751152	26,784753	44,5106357	36,6734825	32,1557665	33,8315989	28,328
69	28,24890896	35,9102756	41,013947	29,7899716	38,1574145	44,0946074	31,7778253	32,3996704	43,037
70	32,88015381	30,6059145	43,6287729	24,4321421	38,9880062	38,6730552	34,7998901	33,1592151	29,004
71	36,09338664	41,6863308	47,6953032	33,9070406	24,7610096	41,1018403	45,2335582	41,7925352	36,025
72	30,29755547	47,3950011	36,0736106	44,8504898	37,9149754	37,0397046	31,1977294	42,6582842	36,867
73	29,50065615	24,1743217	24,0300302	33,6814478	34,5318155	46,2018494	29,6420179	44,631489	29,748
74	39,78270821	35,2598651	33,4565874	43,9481185	36,3138524	26,4536882	26,0589007	33,9517197	35,795
75	44,75746941	34,2710654	34,0850246	26,6690268	33,5459456	44,9098178	41,2065798	26,0632954	39,517
76	=PROMEDIO(A1:A75)		35,3197986	34,8306723	36,4948247	36,6867153	36,0369884	35,1331718	35,307
77	PROMEDIO(número1; [número2]; ...)								

Para graficar la distribución muestral de la media se puede construir un Histograma. Excel permite realizarlo con la función Histograma en el menú análisis de datos.



En Rango de entrada se completa con las celdas de las 1.000 medias muestrales calculadas, ordenadas de menor a mayor, y en Rango de clases se colocan los límites superiores de los intervalos de clase que se desea, si no se completa esta ventana el programa los calculará automáticamente. De esta manera, en una hoja nueva se obtienen los intervalos y las frecuencias de los mismos junto con el Histograma como se muestra a continuación:

Intervalo de clases (minutos)	*Frecuencia*
más de 33 a 33,25	0
más de 33,25 a 33,5	1
más de 33,5 a 33,75	0
más de 33,75 a 34	4
más de 34 a 34,25	8
más de 34,25 a 34,5	15
más de 34,5 a 34,75	28
más de 34,75 a 35	48
más de 35 a 35,25	62
más de 35,25 a 35,5	90
más de 35,5 a 35,75	103
más de 35,75 a 36	147
más de 36 a 36,25	132
más de 36,25 a 36,5	120
más de 36,5 a 36,75	87
más de 36,75 a 37	72
más de 37 a 37,25	38
más de 37,25 a 37,5	23
más de 37,5 a 37,75	16
más de 37,75 a 38	3
más de 38 a 38,25	2
más de 38,25 a 38,5	0
más de 38,5 a 38,75	1
más de 38,75 a 39	0



Distribución muestral de la media de 1.000 muestras de una distribución *U(24, 48)*

b) La media de la distribución de $\bar{X}$, se estima con el promedio de las medias $\bar{X}_1, \bar{X}_2, \ldots,$ $\bar{X}_{1000}$, donde $\overline{X_j}$ denota la media muestral de la *j*-ésima muestra, para *j*=1,.., 1.000. Este promedio se denota con $\overline{\bar{X}}$, medias de medias, y en el caso de las muestras seleccionadas, su valor observado resulta $\bar{\bar{x}} = 35{,}98$, que es la estimación del parámetro solicitado.

La varianza de $\bar{X}$ se estima usando los valores $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_{1000}$ observados, y su estimador se denota con $S^2_{\bar{X}}$ y en el caso de las muestras seleccionadas la estimación resultó $s^2_{\bar{X}} = 0{,}576$.

Se observa que $\bar{\bar{x}}$ es muy próximo a 36, valor de la media de la distribución poblacional a partir de la cual se simularon las muestras y también coincide con la esperanza de la distribución muestral de $\bar{X}$. Por otro lado, se sabe que $V(\overline{X})$ es $\frac{V(X)}{n}$=0,64, valor no muy distante de la estimación realizada con $s^2_{\bar{X}}$.

c) La distribución muestral construida es claramente simétrica, con forma aproximadamente de campana, que podría parecerse a una campana de Gauss (distribución Normal), forma que era de esperar en virtud del Teorema Central del Límite.

TRABAJO PRÁCTICO Nº 7

1) Defina brevemente y con precisión:
   a) Población. b) Muestra.

c) Muestra aleatoria.
d) Mediana muestral.
e) Varianza y desvío estándar muestrales.
f) Distribución muestral.
g) Estadístico

2) Encuentre la media, la mediana y el desvío estándar de una muestra cuyas observaciones son 825, 3300, 2000, 550, 625, 450, 575, 425, 462 y 425, valores que representan el ingreso mensual, en dólares estadounidenses, de diez empleados elegidos al azar en una tienda de indumentaria deportiva. ¿Qué valor parece ser la mejor medida de posición? Justifique su respuesta.

3) Se está estudiando una población formada por los pesos de seis niños recién nacidos, medidos en kilogramos. Los mismos son los que se detallan a continuación:
x_1 = 2,5, x_2 = 3,3, x_3 = 3,5, x_4 = 3,8, x_5 =2,9, x_6 = 3,7.
   a) Calcule la media y la varianza de la población.
   b) Calcule $\bar{x}$ y s^2 para todas las muestras posibles de tamaño n = 2 con reposición de esa población de seis niños.
   c) Construya la distribución muestral de la media muestral y la distribución muestral del S^2. Luego calcule la media y varianza de dichas distribuciones. Compare con lo obtenido en el apartado (a).

4) El consumo de energía eléctrica en los hogares de Tucumán está distribuido normalmente con valor esperado 260 kWh y desviación estándar de 190 kWh.
   a) ¿Cuál es la probabilidad de que al medir el consumo en un hogar este sea mayor a 250 kWh en valor absoluto?
   b) Si se seleccionan aleatoriamente 15 hogares y se registran sus consumos mensuales ¿cuál es probabilidad de que el desvío estándar supere 190 kWh?

5) Para evitar dificultades con la Liga de Defensa al Consumidor, un embotellador debe asegurarse de que las botellas de 500 ml en realidad contengan esta cantidad de bebida. Para determinar si la embotelladora está funcionando bien, se selecciona una muestra al azar de 10 botellas por hora y se mide la cantidad de bebida que contiene cada una. Se reajusta la máquina si la cantidad promedio de bebida en las botellas seleccionadas es inferior a 500 ml. Si en los registros se observa que la cantidad de líquido por botella está normalmente distribuida, con una desviación estándar de 15 ml, y si la máquina embotelladora fue ajustada para introducir un llenado medio por botella de 510 ml, ¿cuál es la probabilidad de que se decida ajustar la máquina?

6) El ingreso familiar, registrado en dólares estadounidenses, de los socios de un club de tenis es una variable aleatoria Normal con μ =1500 y desvío estándar σ =150. Se selecciona al azar una muestra de n=25 socios. Sea $\bar{X}$el ingreso familiar medio de los socios en la muestra.
   a) ¿Cuál es la distribución de$\bar{X}$? Razone si se trata de una distribución exacta o aproximada.
   b) Calcular la probabilidad de que el ingreso familiar exceda \$1550 y la probabilidad de que la media de la muestra exceda \$1550.
   c) Si no se supiese que la distribución del ingreso de los socios es Normal y se seleccionase una muestra de veinte individuos, ¿se puede saber cuál es la distribución de la media muestral?
   d) Si no se supiese que la distribución de los ingresos familiares es Normal y se seleccionan 120 socios, ¿se puede saber cuál es la distribución de la media muestral? La distribución que mencionó ¿es exacta o aproximada?
   e) Genere con su computadora cien muestras de tamaño 20 de esta población y grafique la distribución de medias.

7) Se desea comparar la variabilidad en las notas de estudiantes del mismo grado escolar pero distintos salones. Se toman muestras independientes de tamaño $n_1 = 15$ y $n_2 = 17$ en los salones. Asumiendo la misma varianza poblacional, calcule el número b para el cual $P\left(\frac{S_1^2}{S_2^2} \leq b\right) = 0{,}99$.

8) La sección de carnicería de un hipermercado de la provincia recibe la media res de novillo de dos proveedores diferentes, uno de la provincia de Santa Fe y otro de Córdoba. El peso medio de la media res del proveedor de Santa Fe es 240 kg y el desvío estándar 20 kg, mientras que el peso medio de las medias res del proveedor de Córdoba es 230 kg y desvío estándar 20 kg.
   a) Si se tomaran dos muestras independientes de cincuenta reses de cada proveedor Calcule la probabilidad de que los pesos promedios difieran como máximo en 10 kg. Interprete el resultado alcanzado e indique con claridad los supuestos que debió realizar.
   b) Si el desvío estándar de la muestra seleccionada del proveedor de Santa Fe es 19,15 kg y del proveedor de Córdoba 22,30, calcule la probabilidad de que el cociente de las varianzas de las muestras muestre un valor más extremo al obtenido en la muestra. De acuerdo al resultado de este cálculo ¿usted sostendría que las varianzas de las dos poblaciones son iguales?

9) Suponga que el tiempo entre llegadas de vehículos procedentes de Marruecos para embarcar en un puerto del sur de España sigue una distribución uniforme en el intervalo (0,10).
   a) Utilizando el programa Excel, grafique esta distribución de probabilidad uniforme.
   b) Genere con su computadora 120 muestras de tamaño 50 de esta población y construya la distribución de medias y varianzas.
   c) Calcule la media y varianza de la distribución muestral construida. Compare estos resultados con el valor del parámetro poblacional.
   d) ¿Qué forma tiene la distribución muestral que construyó?

10) Dada una población uniforme discreta $f(x) = \begin{cases} 1/3 & x = 2{,}4{,}6 \\ 0 & otrocaso \end{cases}$. Encuentre la probabilidad aproximada de que la media de una muestra aleatoria de tamaño 54, seleccionada con reemplazo, sea mayor que 4,1 y menor que 4,4.

11) Uno de los dos candidatos que compiten por el ballotage para la presidencia, (Candidato A), pide a una consultora de opinión política que estime la intención de votos que él tiene en todo el territorio del país una semana antes del ballotage.
Los responsables de la consultora deciden realizar una encuesta telefónica, llamando a celulares seleccionados de las líneas activas según todas las compañias que prestan servicio de telefonía celular en el país. Se realizaron cinco mil llamadas seleccionando los números en forma aleatoria.
Si la persona que llamaban no contestaba luego de cuatro intentos se reemplazaba dicho número en la muestra, de acuerdo a instrucciones que recibieron los encuestadores. Un dato que informó la consultora es que debió reemplazarse el 30% de los números de la muestra por falta de respuesta en cuatro llamados.
Luego de realizar el trabajo de campo comunicándose con los cinco mil ciudadanos tres mil respondieron que votarían en el ballotage al candidato “A”.
   a) ¿Cuál es la población objetivo que desean estudiar?
   b) ¿Cuál es la población estadística que se desea estudiar?

c) ¿La muestra seleccionada es adecuada para conocer la población objetivo?
d) ¿Cuál es la población muestreada?
e) ¿En su opinión es correcto generalizar los resultados obtenidos en la muestra con la de todos los votantes?

Ejercicios Adicionales

1) Los pesos al nacer, en kilogramos, de seis niños tucumanos son: 3,2; 3,1; 2,7; 2,5; 3,3 y 3,2. Se seleccionan muestras de tamaño dos, con reposición.
   a) Determine todas las muestras posibles y encuentre la distribución muestral de la media muestral.
   b) Calcule la media de la población y compare con la media de $\overline{X}$.
   c) Calcule la varianza de la distribución de medias y compare con la varianza de la variable original.

2) ¿Bajo qué condiciones:
   a) $\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$ tiene una distribución normal estándar?
   b) $\frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$ se distribuye como normal, exacta o aproximadamente?
   c) $\frac{\overline{X}-\mu}{\frac{S}{\sqrt{n}}}$ tiene una distribución t con $(n$-1) grados de libertad?
   d) $\frac{(n-1)S^2}{\sigma^2}$ tiene una distribución χ^2 con $\nu = n$-1 grados de libertad?
   e) $\frac{\frac{S_1^2}{\sigma_1^2}}{\frac{S_2^2}{\sigma_2^2}}$ tiene una distribución F con $\upsilon_1 = n_1 - 1$ y $\upsilon_2 = n_2 - 1$ grados de libertad?

3) Enuncie el Teorema Central del Límite. Indique claramente las condiciones necesarias para su aplicación

4) Indique si las siguientes aseveraciones son verdaderas o falsas, justificando en cada caso su respuesta.
   a) La distribución de la media muestral depende del tamaño de la muestra.
   b) La distribución de la media muestral depende de la forma en que se realizó el muestreo (con o sin reposición).
   c) La distribución de la media muestral depende de la distribución de la población.
   d) Si X_1, X_2,...,X_n es una muestra aleatoria de una población, con σ^2conocido, entonces $\bar{X}$se distribuye como normal.
   e) La varianza de la distribución muestral de la media se reduce a la mitad si el tamaño de la muestra se duplica.
   f) $\bar{X}$se distribuye como normal, independientemente del tamaño de la muestra cuando la distribución de la que provienen los datos es normal.

5) Sea $\bar{X}_1$ la media de una muestra aleatoria de tamaño n_1, seleccionada con reemplazo de la población discreta

X	1	3	5
$f(x)$	1/3	1/3	1/3

y sea $\bar{X}_2$ la media de una muestra aleatoria de tamaño n_2, seleccionada con reemplazo de la población discreta

X	1	3
$f(x)$	1/3	2/3

Si $n_1 = 60$ y $n_2 = 50$, ¿cuál es la probabilidad de que $\bar{X}_1 - \bar{X}_2$ sea mayor que 0,75 y menor que 1?

6) El tiempo diario de inactividad de un equipo de computación tiene una media de 4 horas y una desviación estándar de 1,5 horas.
   a) Si se desea calcular probabilidades relacionadas con los tiempos diarios promedio de inactividad durante un período de 30 días:
      i) ¿Qué condiciones se deben cumplir para obtener una aproximación válida de las probabilidades relacionadas con el tiempo promedio de inactividad diario mediante la aplicación del TCL?
      ii) De acuerdo con las condiciones del apartado anterior, ¿cuál es la probabilidad aproximada de que el tiempo promedio de inactividad diario durante un período de 30 días esté entre 3,6 y 4,4 horas?
   b) De acuerdo con las condiciones del apartado a), ¿cuál es la probabilidad aproximada de que el tiempo total de inactividad en 30 días sea inferior a 115 horas?

7) El período de tiempo que un cajero de supermercado atiende a un cliente es una variable aleatoria con una media de $\mu = 5$ minutos y desviación estándar de $\sigma = 2$ min.
   a) Si se observa una muestra aleatoria de 56 clientes, encuentre la probabilidad de que en la muestra el tiempo promedio de atención del cajero por cliente sea al menos de 3 minutos, pero no mayor a 6 minutos.
   b) ¿Podría contestar la pregunta anterior si se observó una muestra de tan sólo 16 clientes? Justifique. ¿Agregando algún supuesto podría responderla? ¿Por qué?

8) Considere que un productor de tornillos está lanzando a la venta una clase alternativa de tornillos, con las mismas características, pero con diferente material. Él informa que el diámetro de los tornillos nuevos tiene la misma media y varianza que el diámetro de los tornillos antiguos. Para verificar que efectivamente tienen la misma variabilidad y considerando que el diámetro es una variable normal, se diseñó un muestreo aleatorio en ambas poblaciones de tornillos con tamaño 8 entre los viejos y 9 entre los nuevos. Suponiendo que los resultados obtenidos sean (a)$s_N^2 = 20mm^2$, $s_V^2 = 1{,}5mm^2$ o (b))$s_N^2 = 2mm^2$, $s_V^2 = 23mm^2$, calcule la probabilidad de que el cociente de las varianzas de las muestras presente un valor igual o más extremo al obtenido en las muestras ¿Qué opina Ud., en base al cálculo realizado, sobre la afirmación del productor de que la variabilidad del diámetro no ha cambiado en los nuevos tornillos?

9) Aerolíneas Argentinas afirma que el tiempo de vuelo entre Mendoza y Santiago de Chile es de 45 minutos. Suponga que el tiempo real de vuelo se distribuye uniformemente entre 40 y 60 minutos.
   a) ¿Cuál es el tiempo esperado de vuelo?
   b) ¿Cuál es la probabilidad de que el vuelo llegue más de 5 minutos atrasado con relación al tiempo estipulado por la empresa?
   c) En los próximos 10 vuelos ¿cuál es la probabilidad de que ninguno de ellos se retrase más de 5 minutos? Explicite los supuestos realizados.

d) Si se sacara una muestra aleatoria de 50 tiempos de vuelos ¿Podría decir cuál es la distribución de probabilidad del promedio de ellos? En caso que su respuesta sea afirmativa, justifique su respuesta y especifique la distribución y sus parámetros.

10) Suponga que la distribución de los ingresos de los habitantes de una localidad de la provincia de Tucumán, registrados en dólares estadounidenses, es una variable aleatoria normal con media igual a U$S 500 y desviación estándar U$S 50.
a) ¿Cuál es el ingreso máximo entre el 33% más pobre de esa localidad?
b) Si se seleccionara una muestra aleatoria de tamaño 25 de los habitantes de dicha localidad y se calculara la media muestral $\bar{X}$ y la varianza muestral S^2.
i. ¿Cuál es la probabilidad de que la media muestral no esté alejada más de 5 unidades de la media de la población?
ii. ¿Cuál es la probabilidad de que la varianza muestral sea mayor que 3078,4375?

11) Los pesos de los bebés recién nacidos en un hospital del interior de la provincia se distribuyen como una Normal con media 2.900 gr. y varianza 190 gr.
a) Genere con su computadora 1000 muestras de tamaño 20 de esta población y grafique la distribución muestral de la media.
b) Calcule la media y varianza de la distribución muestral construida. Compare estos resultados con los valores de los parámetros poblacionales.
c) ¿Qué observa de la forma de la distribución muestral que construyó?

12) Encuentre la media, la mediana y el desvío estándar de una muestra cuyas observaciones son 15, 7, 8, 95, 19, 12, 8, 22 y 14, valores que representan el número de días durante los que se reportaron ausencias por enfermedad en 9 legajos de empleados de una empresa. ¿Qué valor parece ser la mejor medida de posición? Justifique su respuesta.

13) Supongamos que se está estudiando una población formada por los pesos de seis perros, medidos en kilogramos. Los mismos son los que se detallan a continuación:

$$x_1 = 11, x_2 = 16, \quad x_3 = 12, \quad x_4 = 15, \quad x_5 = 16 \text{ y } \quad x_6 = 14$$

a) Calcule la media y la varianza de la población.
b) Calcule $\bar{x}$ y s^2 para todas las muestras posibles de tamaño $n = 2$ con reposición de esa población de 6 perros.
c) Construya la distribución muestral de la media muestral y la distribución muestral del S^2. Luego calcule la media y varianza de dichas distribuciones. Compare con lo obtenido en el apartado (a).

14) Los puntajes obtenidos por los aspirantes al ingreso a una universidad tienen una distribución Normal con $\mu = 86$ puntos y $\sigma = 14{,}262$.
Si se toma una muestra aleatoria de 20 aspirantes al ingreso de esta universidad ¿cuál es la probabilidad de obtener un desvío estándar muestral menor o igual a 12,82 puntos?

15) Si se toman muestras independientes de tamaño $n_1 = 6$ y $n_2 = 10$ de dos poblaciones normales con la misma varianza poblacional, calcule el número b para el cual $P\left(\frac{S_1^2}{S_2^2} \leq b\right) = 0{,}95$.

16) Un fabricante de papel que se utiliza para embalaje conoce que el papel que fabrica tiene una resistencia media de 20 libras por pulgada cuadrada y una desviación estándar de 2 libras por pulgada cuadrada, con distribución normal. Cada hora se toma una muestra aleatoria de 10 piezas de papel y se mide su resistencia. ¿Cuál es la probabilidad de obtener un desvío estándar mayor o igual a 2,18?

17) Las calificaciones en el examen de ingreso a una Facultad tienen una media de 540 puntos con una desviación estándar de 50. La media y la desviación estándar es la misma entre los ingresantes de las dos carreras que se brindan en esa facultad. Si se toman 2 muestras independientes de 31 y 61 estudiantes de cada carrera,
   a) Calcule la probabilidad de que las calificaciones promedio de las muestras difieran en 10 puntos o más y la probabilidad que difieran en 20 puntos o más. Interprete este resultado (Indique claramente los supuestos que utilizó).
   b) Si s_1=52,7 y s_2=32, calcule la probabilidad de que el cociente de las varianzas de las muestras presente un valor igual o más extremo al obtenido en la muestra ¿cree Ud., en base al cálculo realizado, que los grupos de donde sacó las muestras son homogéneos respecto a su variabilidad?

18) Un programa de televisión que se transmite en vivo el sábado de 20 a 22 horas realiza una encuesta sobre si la policía debe intervenir cuando las protestas callejeras interrumpen el tránsito y la circulación, para garantizar los derechos del resto de la ciudadanía. Si el televidente responde "sí" debe llamar a un teléfono que se muestra en la pantalla, y si responde "no" a otro. Unos minutos antes de finalizar el programa el conductor muestra los porcentajes de votantes a favor y en contra de la intervención policial como representativas de la "opinión de la calle".
   a) Describa la población de donde se extrajo la muestra.
   b) ¿Qué opinión tiene Ud. sobre la generalización que realiza el conductor

19) Se está estudiando una población formada por los pesos de ocho gatos, medidos en kilogramos. Los mismos son los que se detallan a continuación:
$x_1 = 10$, $x_2 = 14$, $x_3 = 8$, $x_4 = 12$, $x_5 = 6$, $x_6 = 11$, $x_7 = 13$ y $x_8 = 9$.
   a) Calcule la media y la varianza de la población.
   b) Calcule $\bar{x}$ y s^2 para todas las muestras posibles de tamaño $n = 2$ con reposición de esa población de 8 gatos.
   c) Construya la distribución muestral de la media muestral y la distribución muestral del S^2. Luego calcule la media y varianza de dichas distribuciones. Compare con lo obtenido en el apartado (a).

20) Se sabe que los errores en cierto instrumento para medir longitudes están distribuidos normalmente con valor esperado cero y desviación estándar de 0,1 pulgadas.
   a) ¿Cuál es la probabilidad de que al medir un error sea mayor de 0,12 pulgadas en valor abosoluto?
   b) Si se tomaran diez mediciones aleatoriamente y se registraran los errores cometidos en la medición ¿cuál es probabilidad de que el desvío estándar sea mayor o igual a 0,11 pulgadas?

EJERCICIOS SUGERIDOS DEL LIBRO "PROBABILIDAD Y ESTADÍSTICA PARA INGENIERÍA Y CIENCIAS" de WALPOLE, MYERS y MYERS. (9º Ed.)

Capítulo 8:

Secciones 8.1 y 8.2, página 230: ejercicios 8.1 y 8.14
Secciones 8.3 y 8.4, página 241: ejercicios 8.21 y 8.34
Secciones 8.5 a 8.8, página 259: ejercicios 8.41,8.47, 8.49 y 8.53

CAPÍTULO 8: ESTIMACIÓN PUNTUAL

Inferencia Estadística. Métodos clásicos de estimación. Estimación Puntual. Propiedades de los estimadores. Estimación de la media. Estimación de diferencias entre dos medias. Estimación de una proporción. Estimación de diferencia entre dos proporciones. Estimación de varianzas. Estimación de la razón entre dos varianzas. Estimadores de máxima verosimilitud.

Objetivos:

El alumno debe ser capaz de:

- Deducir conclusiones acerca de los parámetros de la población a partir de la información obtenida en una muestra aleatoria.
- Identificar las propiedades de los estimadores o estadísticos.
- Dada una muestra aleatoria de tamaño *n*, con distribución de probabilidad conocida con excepción de alguno de sus parámetros, calcular los estimadores máximo verosímiles de los mismos.

Resumen

D1. *Parámetro*: Las características de la población, en general desconocidas, sobre las cuales tenemos interés se denominan parámetros y usualmente se representan por letras griegas tales como μ, θ, π, σ, entre otras.

D2. *Estimador*: Un estimador es un estadístico cuya forma funcional no involucra ningún parámetro y que se construye con la intención de atribuir un valor que pudiera representar un parámetro de interés desconocido en la población.

D3. *Estimación puntual:* Es el valor asumido por un estimador cuando se evalúa en los valores obtenidos en una muestra particular.

D4. *Estimador insesgado*: Se dice que $\hat{\theta}$ es un estimador insesgado para el parámetro θ si $E(\hat{\theta}) = \theta$.

D5. *Sesgo*: $E(\hat{\theta}) - \theta$

Obs: Un estimador se dice *sesgado* cuando el sesgo es no nulo.

D6. *Error Cuadrático Medio*: $ECM(\hat{\theta}) = E(\hat{\theta} - \theta)^2$

Estimadores para media, proporción y varianza

Parámetro	Estimador	Propiedades
μ	$\bar{X} = \dfrac{X_1+\ldots+X_n}{n}$	Insesgado
π	$\hat{\pi} = \dfrac{\text{Nro de éxitos en la muestra}}{n}$	Insesgado
σ^2	$S^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$	Insesgado
σ^2	$\hat{\sigma}^2 = \dfrac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$	Sesgado

D7. *Eficiencia*: Dados dos estimadores $\hat{\theta}_1$y$\hat{\theta}_2$ insesgados para el parámetro θ, decimos que $\hat{\theta}_1$ es más eficiente que $\hat{\theta}_2$ si $Var(\hat{\theta}_1) < Var(\hat{\theta}_2)$.

D8. *Función de Verosimilitud*: Sean $X_1, X_2, \dots, X_n$ i.i.d. como X con función de distribución de probabilidad $f_X(x;\theta)$. Entonces la función de verosimilitud es

$$L(\theta) = f(\theta; x_1, x_2, \dots, x_n) = \prod_{i=1}^{n} f_X(x_i;\theta)$$

considerada como función de θ dados los valores de la muestra $(x_1, x_2, \dots, x_n)$.

Estimación por Máxima Verosimilitud: Dadas las observaciones $x_1, x_2, \dots, x_n$ de una muestra aleatoria $X_1, X_2, \dots, X_n$ de una población con f.d.p. $f_X(x;\theta)$, la estimación por máxima verosimilitud de θ, $\hat{\theta}$, es aquella función de los datos que maximiza la función verosimilitud

$$L(\theta) = f(\theta; x_1, x_2, \dots, x_n).$$

Obs: En la búsqueda de un estimador para un parámetro, la elección del estimador podrá basarse en sus propiedades o en el método de obtención del mismo.

Ejemplos

Ejemplo 1:

Se obtuvo una muestra al azar de una población exponencial con media β, resultando los valores 3, 5, 4, 3, 2, 4, 3 y 5. Obtenga la estimación de β por máxima verosimilitud.

Resolución:

Sean $X_1, X_2, \dots, X_n$ i.i.d $X \sim \text{Exp}(\beta)$

Considerando que la función de densidad de la distribución exponencial es:

$$\boldsymbol{f(x)} = \begin{cases} \boldsymbol{(1/\beta) e^{(-x/\beta)}} & \textbf{si } \boldsymbol{x > 0} \\ \boldsymbol{0} & \textbf{en otro caso,} \end{cases}$$

para obtener la estimación de β por máxima verosimilitud, se maximiza la función de verosimilitud

$$\boldsymbol{L(\beta; x_1, x_2, \dots, x_n) = (1/\beta) e^{(-x_1/\beta)} \dots (1/\beta) e^{(-x_n/\beta)} = \beta^{-n} \, exp \sum_{i=1}^{n} -(x_i/\beta),}$$

que es la distribución conjunta de las variables aleatorias de la muestra aleatoria vista como función del parámetro.

Puesto que el logaritmo es una función estrictamente creciente, el máximo de la función de verosimilitud coincide con el máximo de su logaritmo, y ya que en la mayoría de las funciones de probabilidad la búsqueda del máximo es más sencilla si se trabaja con el logaritmo, se recomienda aplicar el logaritmo natural cada vez que se busque maximizar la verosimilitud.

Aplicando logaritmo natural

$$\boldsymbol{ln\, L\,(\beta) = -n\, ln\, \beta - \sum_{i=1}^{n} x_i/\beta},$$

el máximo de la verosimilitud se busca derivando $ln\, L\,(\beta)$ con respecto a β e igualando a cero:

$$\frac{\partial\, ln[L(\beta)]}{\partial \beta} = -\frac{n}{\beta} - (-1)\beta^{-2} \sum_{i=1}^{n} x_i = 0$$

$$-\frac{n}{\beta} + \frac{\sum_{i=1}^{n} x_i}{\beta^2} = 0$$

$$\frac{-n\beta + \sum_{i=1}^{n} x_i}{\beta^2} = 0$$

$$-n\beta + \sum_{i=1}^{n} x_i = 0.$$

Luego el valor de $\boldsymbol{\beta}$ que maximiza la función de verosimilitud es $\frac{\sum_{i=1}^{n} x_i}{n}$.

Por lo tanto, el estimador máximo verosímil de β es $\hat{\boldsymbol{\beta}} = \frac{\sum_{i=1}^{n} X_i}{n} = \bar{\boldsymbol{X}}$.

En este ejemplo particular la estimación por máxima verosimilitud de $\boldsymbol{\beta}$ es 3,25.

Se puede verificar que la segunda derivada de la función de verosimilitud evaluada en el punto crítico encontrado es menor que cero lo que confirma que se trata de un punto de máximo de $L(\beta)$.

Ejemplo 2:

Evalúe si el estimador $S_*^2 = n^{-1} \sum_{i=1}^{n} (X_i - \overline{X})^2$ es un estimador insesgado del parámetro σ^2.

Resolución:

El numerador de $\boldsymbol{S_*^2}$ puede expresarse de la siguiente manera:

$$\sum_{i=1}^{n} (\boldsymbol{X_i} - \overline{\boldsymbol{X}})^2 = \sum_{i=1}^{n} \left[(\boldsymbol{X_i} - \boldsymbol{\mu}) - (\overline{\boldsymbol{X}} - \boldsymbol{\mu})\right]^2$$

$$= \sum_{i=1}^{n} (X_i - \mu)^2 - \sum_{i=1}^{n} 2(\overline{X} - \mu)(X_i - \mu) + \sum_{i=1}^{n} (\overline{X} - \mu)^2$$

$$= \boldsymbol{\sum_{i=1}^{n} (X_i - \mu)^2 - 2(\overline{X} - \mu) \sum_{i=1}^{n} (X_i - \mu) + n(\overline{X} - \mu)^2}$$

$$= \sum_{i=1}^{n} (X_i - \mu)^2 - 2n(\overline{X} - \mu)^2 + n(\overline{X} - \mu)^2 = \sum_{i=1}^{n} (X_i - \mu)^2 - n(\overline{X} - \mu)^2$$

Aplicando esperanza,

$$E(\boldsymbol{S_*^2}) = E\left[\frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{n}\right]$$

$$= \frac{1}{n}\left[\sum_{i=1}^{n} E(X_i - \mu)^2 - nE(\overline{X} - \mu)^2\right] = \frac{1}{n}\left(\sum_{i=1}^{n} \sigma_{X_i}^2 - n\sigma_{\overline{X}}^2\right)$$

Como, $\sigma_{X_i}^2 = \sigma^2$ para i = 1, 2,…, n y $\sigma_{\overline{X}}^2 = \frac{\sigma^2}{n}$, se tiene que,

$$E(S_*^2) = \frac{1}{n}\left(n\sigma^2 - n\frac{\sigma^2}{n}\right) = \left[\frac{n-1}{n}\right]\sigma^2 \neq \sigma^2.$$

Luego el estimador $\boldsymbol{S_*^2}$ no es insesgado.

Ejemplo 3:

Al finalizar un curso, los alumnos pueden promocionar la materia, quedar regulares o quedar libres. Estimar por máxima verosimilitud la probabilidad de promocionar, si se sabe que ésta es 1/3 de la probabilidad de regularizar la materia, y que en una muestra aleatoria de 40 alumnos se encontraron 7 que promocionaron, 17 de regularizaron y el resto quedaron libres.

Resolución:

Sea X la variable aleatoria que asume el valor 0 si el alumno queda libre, 1 si regulariza y 2 si promociona.

Debido a que la probabilidad de promocionares 1/3 de la de regularizar, la función de masa de X resulta ser:

x	0	1	2
$p(x)$	$\boldsymbol{1-4\theta}$	$\boldsymbol{3\theta}$	$\boldsymbol{\theta}$

donde $\boldsymbol{\theta}$ es la probabilidad de promocionar la materia.

La verosimilitud en este caso puede escribirse como

$$\boldsymbol{L(\theta; x_1, \dots, x_n) = \theta^{n_2}(3\theta)^{n_1}(1-4\theta)^{n_0},}$$

donde $\boldsymbol{n_k}$ es el número de elementos en la muestra que presentan el valor k, para k=0,1 y 2.

Para obtener el estimador por máxima verosimilitud, se escribe la log-verosimilitud y se deriva respecto del parámetro que se quiere estimar. Es decir:

$$\boldsymbol{\ln L\,(\theta; x_1, \dots, x_n) = n_2 \ln\theta + n_1(\ln 3 + \ln\theta) + n_0 \ln(1-4\theta)}$$

$$\boldsymbol{= (n_2 + n_1)\ln\theta + n_0 \ln(1-4\theta) + n_1 \ln 3}$$

$$\boldsymbol{\frac{d \ln L\,(\theta; x_1, \dots, x_n)}{d\theta} = \frac{n_2+n_1}{\theta} + \frac{n_0}{1-4\theta}(-4)}$$

$$\boldsymbol{= \frac{(1-4\theta)(n_2+n_1) - 4\theta n_0}{\theta(1-4\theta)}}$$

$$\boldsymbol{= \frac{(n_2+n_1) - 4\theta(n_2+n_1+n_0)}{\theta(1-4\theta)}}$$

Igualando a cero esta expresión se obtiene el estimador por MV de $\boldsymbol{\theta}$ es $\hat{\theta} = \frac{n_2+n_1}{4n}$, con $n = n_2 + n_1 + n_0$. La estimación por MV de $\boldsymbol{\theta}$ es 24/160=0,15.

Se puede verificar que la segunda derivada de la función de verosimilitud evaluada en el punto crítico encontrado es menor que cero.

Como en este caso se trata de una distribución discreta, la verosimilitud se puede interpretar como la probabilidad de obtener la muestra obtenida, y el valor estimado de θ es el valor que maximiza esa probabilidad.

TRABAJO PRÁCTICO Nº 8

1) El departamento de marketing de una empresa desea conocer la proporción de clientes que adquirieron un determinado producto a través de la página web en el último mes.
   Con el fin de estimar esta proporción, se seleccionaron aleatoriamente 350 compras del producto en el último mes y se registró si fueron realizadas en tienda online o física.
   De las 350 compras analizadas, 91 se realizaron en una tienda física.
   a) ¿Cuál es la población que desean estudiar?
   b) ¿La muestra seleccionada es adecuada para conocer la población objetivo?
   c) ¿Cuál es la población muestreada?
   d) ¿Cuál es el parámetro a estimar? Cuál es su distribución.
   e) ¿De qué forma puede informar a las autoridades de la empresa sobre el valor estimado de la proporción?

2) Sean X_1, X_2, X_3, X_4 independientes e idénticamente distribuidas (iid) como X, donde $E(X) = \mu$ con $\mu \neq 0$ y $V(X) = \sigma^2$. Considere los siguientes estimadores de μ:
   $i)$ ⤢ $\hat{\mu}_1 = X_1$ $\quad$ $ii)$ ⤢ $\hat{\mu}_2 = \frac{X_1+X_2+X_3+X_4}{3}$ $\quad$ $iii)$ ⤢ $\hat{\mu}_3 = X_1 + \frac{X_2-X_3}{2}$

Indique:

a) ¿Cuáles de estos estimadores son insesgados para μ?
b) ¿Cuál es el de menor varianza?
c) ¿Cuál es el más eficiente?
d) Teniendo en cuenta que tiene 4 elementos en la muestra, ¿Ud. podría proponer algún otro estimador más eficiente para estimar μ?

3) Usando la computadora genere trescientas muestras de tamaño diez y trescientas muestras de tamaño treinta de una distribución continuna N(50, 5). Grafique la distribución muestral de las medias para ambos tamaños de muestra. ¿Qué concluiones puede enuniciar luego de graficar ambas distribuciones?

4) Muestre que:
   a) $ECM(\hat{\theta}) = sesgo^2(\hat{\theta}) + V(\hat{\theta})$, donde $\hat{\theta}$ es un estimador de θ.
   b) Cualquier combinación lineal $\sum_{k=1}^{K} \lambda_k \hat{\theta}_k$ de estimadores insesgados de θ es insesgado para θ si $\sum_{k=1}^{K} \lambda_k = 1$.

5) Para estimar la media de una población se considera el estimador $a\bar{X}$. Encontrar el valor de *a* que minimiza el ECM.

6) a) Describa el método de estimación por máxima verosimilitud.
   b) Interprete la estimación por máxima verosimilitud de un parámetro de una distribución discreta. ¿Ocurre lo mismo para una población continua?
   c) Deduzca el estimador máximo verosímil para el parámetro μ de una distribución normal con σ^2 conocido.
   d) ¿Cuáles son las propiedades de los estimadores máximo-verosímiles?

7) Una empresa desea conocer la proporción de clientes que realiza consultas a través de WhatsApp. Para ello se seleccionó aleatoriamente una muestra de cuarenta consultas realizadas en el último año y se registró si fueron realizadas mediante Whatsapp o no. Con la información de esta muestra, estime por máxima verosimilitud la proporción de consultas realizadas mediante WhatsApp.

8) Un banco está interesado en analizar información sobre las consultas que se realizan en forma telefónica. En particular, se desea estimar el número de llamadas telefónicas que se reciben cada 10 minutos. Para ello se tomó una muestra aleatoria de *n*=15 periodos de 10 minutos, registrándose las llamadas recibidas. Las observaciones fueron: 13, 9, 7, 17, 5, 19, 16, 11, 8, 6, 12, 14, 21, 10 y 9.
   a) ¿Con qué distribución de probabilidad describiría Ud. la variable que cuenta el número de llamadas en un periodo de 10 minutos? ¿Por qué?
   b) Deduzca el estimador por máxima verosimilitud del parámetro de la distribución y calcule la estimación puntual del parámetro.
   c) Suponiendo que el verdadero parámetro de la distribución coincida con su estimación realizada en el apartado anterior, calcule la probabilidad que al elegir aleatoriamente 4 periodos de 10 minutos, en ninguno de ellos lleguen llamadas.
   d) Por cada veinte minutos, el mantenimiento del centro de atención tiene un costo fijo de 5 pesos, y un costo adicional de 2 pesos si se reciben más de 30 llamadas. Suponiendo que el verdadero parámetro de la distribución coincida con su estimación realizada en el apartado (b), calcule el costo esperado por hora de funcionamiento del centro de atención.

9) Deduzca el estimador máximo verosímil para el parámetro π de una distribución Binomial (π ; n) si si cuenta con una muestra tamaño 20.

Ejercicios Adicionales

1) Defina brevemente y con precisión:
   a) Estimador.
   b) Estimación puntual.
   c) Estimador insesgado.
   d) Estimador más eficiente.
2) Explique la diferencia entre estimador y estimación de un parámetro
3) Deduzca el estimador máximo verosímil para el parámetro β de una distribución Exponencial.
4) Sean $Y_1, Y_2, \ldots, Y_n$ variables aleatorias independientes con distribución normal, con media $\mu_i = \alpha + \beta x_i$ y varianza conocida, para$i = 1, \ldots, n$; donde$x_1, \ldots, x_n$ son constantes conocidas.
   a) ¿Se podría decir que las variables aquí definidas son i.i.d.?
   b) Obtenga los estimadores por máxima verosimilitud de α y β.
5) El control de recepción de una partida de rodillos se realiza clasificando los rodillos en defectuosos y no defectuosos. La proporción teórica de rodillos defectuosos es 10%, pero se sospecha que ha aumentado a $0{,}10 + \alpha$. Se analizan 5.000 rodillos y se observaron 572 rodillos defectuosos. Estime αpor máxima verosimilitud.
6) Una máquina puede averiarse por dos razones, A y B. Estime la probabilidad de avería diaria de cada tipo si se sabe que (i) la probabilidad de avería de tipo A es el doble que la de B, (ii) no existen otros tipos de averías posibles y (iii) al observar las averías en 30 días seleccionados aleatoriamente se encontraron dos averías tipo A, tres tipo B y 25 días sin avería.
7) Sean $X_1, X_2, \ldots, X_n$ variables aleatorias i.i.d. como *X*. Considere los siguientes estimadores:
   (i) $(n-1)^{-1}\sum_{i=1}^{n} X_i$ (ii) $n^{-1}\sum_{i=1}^{n} X_i$ (iii)$(n+1)^{-1}\sum_{i=1}^{n} X_i$
   a) Calcule la esperanza de cada uno de los estimadores e indique cuál es insesgado para la media de *X*.
   b) Obtenga las respectivas varianzas y los ECM.
   c) ¿Podría caracterizar las distribuciones de estos estimadores cuando (i) *n* es pequeño, (ii) *n* es pequeño y *X* tiene distribución normal y (iii) *n* es grande? Justifique convenientemente sus respuestas.
   d) Esboce las distribuciones muestrales de estos estadísticos en un mismo gráfico, para las situaciones (c.ii) y (c.iii).
   e) Suponiendo que $V(X)$=4, exprese y grafique los ECM como funciones de la media de *X*, e indique cuál de los tres estimadores es mejor para estimarla. ¿Cuál usaría Ud., en qué situaciones y por qué?
8) Sean $X_1, X_2, \ldots, X_n$ variables aleatorias independientes con distribución normal, con media conocida. Obtenga el estimador por MV de σ.
9) Se toman dos muestras independientes de una población $N(\mu, \sigma^2)$, de tamaños *n* y *m* respectivamente. Considere la familia de estimadores para μ:$\{a\bar{X} + b\bar{Z} / a \wedge b \in R\}$