INSTITUTO POLITECNICO NACIONAL
ESCUELA SUPERIOR DE CÓMPUTO
Departamento de Formación básica

Problemas resueltos de probabilidad y estadística.

Trabajo elaborado por: Cañedo Suárez Leticia (lcanedos@ipn.mx)
Tirado Lule Judith Margarita (jtirado@ipn.mx)

Justificación:

Con base en la experiencia, se ha observado que muchos de los estudiantes tienen dificultades a la hora de leer, interpretar y plantear problemas de probabilidad y estadística. Por esta razón se considero pertinente elaborar este trabajo, el cual incluye problemas ilustrativos de las cinco unidades temáticas consideradas en la unidad de aprendizaje de Probabilidad y Estadística, según el rediseño del año 2009.

I Elementos de probabilidad.
II Variables aleatorias discretas y continuas.
III Distribuciones de probabilidad para variables discretas y continuas.
IV Distribución de varias variables aleatorias.
V Estadística paramétrica usando estimación y prueba de hipótesis.

El grado de dificultad de los problemas va de lo simple a lo complejo, cuidando mucho el nivel que un Ingeniero en Sistemas Computacionales egresado de ESCOM debe tener en una unidad de aprendizaje como esta.

Objetivos:

- Servir de guía para los estudiantes que cursan probabilidad y estadística.
- Servir como material de apoyo para los profesores que imparten esta unidad de aprendizaje.

Notación.

1._ AB : Intersección entre los eventos A y B, también conocida como $A \cap B$.
2._ $(X > 2, X > 5)$: Intersección entre los eventos $(X > 2)$ y $(X > 5)$.
3._ A' : Complemento del evento A.
4._ v.a: variable aleatoria.
5._ f.d.p: función de probabilidad.
6._ f.g.m: función generadora de momentos.
7._ IC: Intervalo de confianza.
8._ $f_X(x)$: función de probabilidad de la v.a X.
9._ $A(B_1 \cup B_2)$: Intersección entre A y $(B_1 \cup B_2)$
10._ $P(A|B)$: Probabilidad condicional de A dado B.
11._ TCL : Teorema Central del límite.

UNIDAD I Elementos de probabilidad.

Problema 1._ De una encuesta aplicada a 60 estudiantes que asisten a la universidad, 9 habitan fuera del recinto universitario, 36 son estudiantes de licenciatura y 3 son estudiantes de licenciatura que habitan fuera del recinto.
a) Encuentra el número de estudiantes que están estudiando su licenciatura, que habitan fuera del recinto o que satisfacen ambas características.
b) ¿Cuántos estudiantes de licenciatura habitan en el recinto?
c) ¿Cuántos estudiantes ya tienen su licenciatura y habitan en el recinto?

Solución

Es muy conveniente colocar la información del problema en el cuadro siguiente:

Distribución de los estudiantes dentro y fuera del recinto.

	Lic.	Lic. Terminada	Total
Fuera	3	6	9
Dentro	33	18	51
Total	36	24	60

a) Los que estudian su licenciatura son 36 más los que viven fuera del recinto que son 9 hacen un total de 45 pero hay que restar los 3 que cumplen ambas condiciones quedando únicamente 42 alumnos.
b) Del cuadro se observa a simple vista que 33 alumnos estudian su licenciatura y viven en el recinto.
c) También se observa directamente del cuadro que son 18 los estudiantes que ya tienen su licenciatura terminada y habitan dentro del recinto.

Problema 2._ Una instalación consta de dos calderas y un motor. Sea A el evento de que el motor está en buenas condiciones, mientras que los eventos B_k , $k = 1{,}2$ son los eventos de que la k-ésima caldera esté en buenas condiciones. El evento C es que la instalación pueda funcionar. Si la instalación funciona cada vez que el motor y al menos una caldera funciona, expresa C y C' en términos de A y de los eventos B_k.

Solución

Sean los eventos

A: motor en buenas condiciones.
B_k: caldera k en buenas condiciones, con $k = 1{,}2$.
C: instalación funciona.

La instalación funciona cada vez que el motor y (intersección) al menos una (unión) caldera funciona es decir, el evento C es la intersección del evento A con la unión de los eventos B_k

$$C = A(B_1 \cup B_2)$$

Usando las Leyes de D'Morgan el complemento del evento es:

$$C' = [A(B_1 \cup B_2)]' = A' \cup (B_1 \cup B_2)' = A' \cup (B_1' \cup B_2') = (A' \cup B_1')(A' \cup B_2')$$

Problema 3._ Un mecanismo puede ponerse en cuatro posiciones, digamos *a*, *b*, *c* y *d*. Hay 8 de tales mecanismos en un sistema.
a) ¿De cuántas maneras puede instalarse este sistema?
b) Supón que dichos mecanismos están instalados en algún orden (lineal) pre asignado. ¿De cuántas maneras posibles se instalan los mecanismos, si dos mecanismos adyacentes no están en la misma posición?
c) ¿Cuántas maneras son posibles si sólo se usan las posiciones *a* y *b* con la misma frecuencia?
d) ¿Cuántas maneras son posibles si sólo se usan dos posiciones diferentes y una de ellas aparece tres veces más a menudo que la otra?

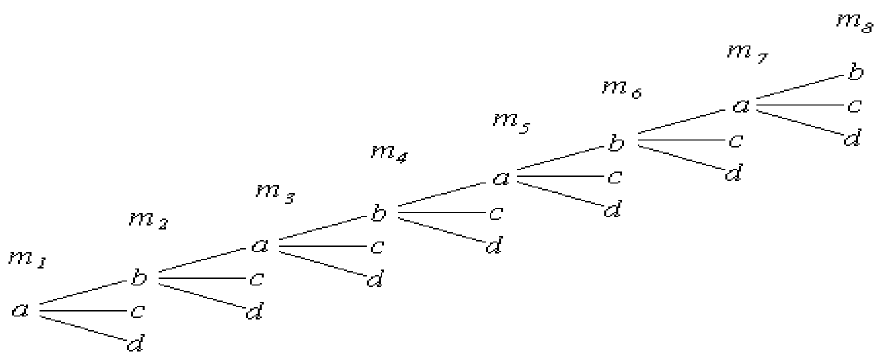
<u>Solución</u>

a) Sea m_i el mecanismo *i*-ésimo, *i* =1, 2, 3, 4, 5, 6, 7, 8. Cada uno de los mecanismos puede colocarse en cuatro posiciones así el sistema puede instalarse como sigue:

4	4	···	4
m_1	m_2	···	m_8

$= 4^8$ maneras de instalar el sistema.

b) Si el mecanismo 1 se coloca en posición *a*, el mecanismo 2 puede colocarse en posición *b*, *c* o *d*, pero no en *a*, si el mecanismo 2 se instala en posición *b* el mecanismo 3 puede instalarse en posición *a*, *c* o *d* pero no en *b*, etc. De tal manera que el primer mecanismo puede instalarse en cualquiera de las cuatro posiciones, el segundo sólo en tres posiciones ya que se descarta la posición del primer mecanismo, el tercer mecanismo se puede instalar en cualquiera de tres posiciones dado que se descarta la posición del mecanismo 2, etc.
Es muy útil hacer un diagrama de árbol como el siguiente:



De manera similar se construye el árbol para *b*, *c* y *d.*

O bien, con el método de las casitas en la primer casita se coloca el número 4 que representa el número de maneras en que se puede instalar el mecanismo 1, en la segunda casita se coloca un 3 ya que el segundo mecanismo se puede instalar sólo de tres maneras, en la tercer casita se coloca un 3 por la misma razón y así hasta llegar a la casita número 8 correspondiente a las posiciones disponibles del mecanismo ocho.

4	3	3	···	3

$= (4)3^7$

Cualquiera de las 4 — se quita la que queda en la casilla anterior

Por lo tanto el mecanismo se puede instalar de $4(3)^7 = 8748$ maneras.

c) Cuatro mecanismos están en la posición *a* y 4 mecanismos en la posición *b*.
Es como si tomara 4 mecanismos de 8 que tengo y a esos les asigno la posición *a*, si por ejemplo me salen los números 2, 3, 7, 8 significa que a los mecanismos 2, 3, 7 y 8 les asigno la posición *a* y si salen los números 8, 2, 7, 3 da lo mismo, entonces no hay orden por lo tanto son combinaciones.

$$\binom{8}{4} = \frac{8!}{4!4!} = \frac{4!(5)(6)(7)(8)}{4!(4)(3)(2)} = 70 \text{ maneras con 4 mecanismos en la posición } a \text{ y 4 en la posición } b.$$

d) Para que una posición aparezca tres veces más que la otra, se tiene que una aparece 2 veces y la otra aparece 6 veces ($2+6=8$ mecanismos).

¿De cuántas maneras puedo tomar 6 de 8 que tengo? $\binom{8}{6} = \frac{8!}{6!2!} = \frac{7(8)}{2} = 28$

Por ejemplo, si salen los números 1, 2, 3, 4, 5, 6 entonces a los mecanismos correspondientes los coloco en la posición *x* (*a, b, c* ó *d*) y los mecanismos que faltan los coloco en la posición *y*.

Pero por cada una de estas maneras de poner 6 mecanismos en la posición *x* y 2 en la posición *y*, hay una más, poner los 6 mecanismos en la posición *y* y dos en la posición *x*, entonces, necesitamos multiplicar por (2).
Además ¿de cuántas maneras puedo tomar 2 posiciones de 4 que tengo?

$$\binom{4}{2} = \frac{4!}{2!2!} = \frac{3(4)}{2} = 6$$

$\therefore$ El número total de maneras es 28(2)(6) = 336

También puede hacerse con permutaciones, quiero tomar 2 posiciones de 4 que tengo y el orden importa porque la primera posición será puesta en 6 mecanismos y la segunda en sólo 2 mecanismos y no es lo mismo poner 6 mecanismos en la posición *a* y 2 en *b* que poner 6 mecanismos en la posición *b* y 2 en la posición *a*.

$$P_2^4 = \frac{4!}{2!} = 3(4) = 12 \quad \text{Número total de maneras} = 28(12) = 336$$

Problema 4._ En una habitación 10 personas tienen insignias numeradas del 1 al 10. Se eligen tres personas al azar y se les pide que dejen la habitación simultáneamente y se anotan los números de las insignias. ¿Cuál es la probabilidad de que el número
a) menor de las insignias sea 5?
b) mayor de las insignias sea 5?

<u>Solución</u>

El espacio muestral consta de todas las triadas que se pueden formar de los 10 números, no importa el orden por lo tanto la cardinalidad del espacio muestral es:

$$\#(S) = \binom{10}{3} = 120$$

a) El 5 será el número más pequeño de las insignias sólo si él está en la triada y los otros dos son números tomados del 6 al 10, es decir necesitamos saber de cuántas manera se pueden sacar dos de los cinco números 6, 7, 8, 9, 10.
Sea el evento A: Conjunto de triadas donde 5 es el número más pequeño

$$\text{entonces} \quad P(A) = \frac{\#(A)}{\#(S)} = \frac{\binom{5}{2}}{120} = \frac{10}{120} = \frac{1}{12}$$

b) Para que 5 sea el mayor número, debo incluirlo en la triada y tomar los otros 2 de entre los números 1, 2, 3, 4.

Si A: Conjunto de triadas donde 5 es el número más grande.

Se tiene que $P(A) = \dfrac{\#(A)}{\#(S)} = \dfrac{\binom{4}{2}}{120} = \dfrac{1}{20}$

Problema 5._ Supón que A, B y C son eventos tales que $P(A) = P(B) = P(C) = 1/4$, $P(AB) = P(BC) = 0$ y $P(AC) = 1/8$. Calcula la probabilidad de que al menos uno de los eventos A, B o C ocurra.

__Solución__

Como A y B no se intersectan y B y C tampoco entonces $P(ABC) = 0$, así la probabilidad de la unión de los eventos está dada por:

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC) = \frac{3}{4} - \frac{1}{8} = \frac{5}{8}$$

Problema 6._ De 6 números positivos y 8 números negativos se eligen 4 números al azar sin sustitución y se multiplican. ¿Cuál es la probabilidad de que el producto sea un número positivo?

__Solución__

En el cuadro siguiente se presentan todas las opciones al elegir 4 números de los 14 que tenemos. Me interesa sólo el signo entonces no importa el orden.

Posibilidades de tomar 4 números de los 14 disponibles y signo del producto.

Números con signo positivo	Números con signo negativo	Signo del producto de los cuatro números elegidos.
4	0	+
3	1	-
1	3	-
2	2	+
0	4	+

Si A es el evento producto positivo tenemos que:

A = (4 números positivos) $\cup$(2 números positivos y 2 números negativos) $\cup$ (4 números negativos)

$$\#(A)=\binom{6}{4}+\binom{6}{2}\binom{8}{2}+\binom{8}{4}=15+420+70=505$$

y el espacio muestral S es el conjunto de todas las cuartetas de números de entre los disponibles.

Entonces la cardinalidad del espacio muestral es $\#(S)=\binom{14}{4}=1001$

Finalmente la probabilidad de que el producto de los cuatro números elegidos sea positivo es:

$$P(A)=\frac{\#(A)}{\#(S)}=\frac{505}{1001}=0.504$$

Problema 7._ Se eligen sin reemplazo dos dígitos al azar del 1 al 9. Si la suma de los dígitos es par, encuentra la probabilidad de que ambos dígitos sean impares.

Solución

Se trata de una probabilidad condicional donde el espacio muestral reducido es el conjunto de parejas que al ser sumadas resultan en un número par y el evento de interés es que ambos números sean impares.

1 3 5 7 9 $\rightarrow$ 5 números impares

2 4 6 8 $\rightarrow$ 4 números pares

para que la suma sea par se necesita que ambos sean pares o ambos impares

$$\left.\begin{array}{l}\binom{4}{2}=6 \quad \text{ambos pares}\\ \\ \binom{5}{2}=10 \quad \text{ambos impares}\end{array}\right\}16 \text{ maneras de que la suma sea par}$$

$$P(\text{ambos impares}|\text{suma par})=\frac{10}{16}=\frac{5}{8}$$

Problema 8._ Se ha observado que los hombres y las mujeres reaccionan de una manera diferente en ciertas circunstancias; el 70% de las mujeres reacciona positivamente en dichas circunstancias, mientras que el porcentaje en los hombres es solamente del 40%. Se sometió a prueba un grupo de 20 personas, 15 mujeres y 5 hombres, y se les pidió llenar un cuestionario para descubrir sus reacciones. Una respuesta escogida al azar de las 20 resultó negativa. ¿Cuál es la probabilidad de que haya sido contestada por un hombre?

Solución

Sean los eventos m: masculino.
f: femenino.
p: reacción positiva.

Los datos que tenemos en términos de probabilidades son:

$$P(p|f) = 0.7$$

$$P(p|m) = 0.4 \therefore P(p'|m) = 0.6$$

$$P(f) = \frac{3}{4}$$

$$P(m) = \frac{1}{4}$$

$$P(m|p') = ?$$

Usando la Regla de la multiplicación podemos encontrar las probabilidades siguientes:

$$P(pf) = P(p|f)P(f) = \frac{3}{4}\left(\frac{7}{10}\right) = \frac{21}{40}$$

$$P(Pm) = P(p|m)p(m) = \frac{1}{10}$$

$$P(p) = P(pf) + P(pm) = \frac{21}{40} + \frac{1}{10} = \frac{25}{40} = \frac{5}{8}$$

$$\therefore P(p') = \frac{3}{8}$$

Usando la definición de probabilidad condicional se tiene que la probabilidad de que haya contestado un hombre dado que la respuesta fue negativa es:

$$P(m|p') = \frac{P(mp')}{P(p')} = \frac{P(p'|m)P(m)}{P(p')} = \frac{0.6\left(\frac{1}{4}\right)}{\left(\frac{3}{8}\right)} = 0.4$$

Problema 9._ La probabilidad de que un doctor diagnostique de manera correcta una enfermedad particular es 0.7. Dado que el Dr. Hace un diagnóstico incorrecto, la probabilidad de que el paciente presente una demanda es 0.9. ¿Cuál es la probabilidad de que el Dr. haga un diagnóstico incorrecto y el paciente lo demande?

<u>Solución</u>

Sean los eventos I: Diagnóstico incorrecto.
I' : Diagnóstico correcto.
D: El paciente demanda al doctor.

Los datos que tenemos en términos de probabilidades son:

$$P(I') = 0.7 \quad \therefore P(I) = 0.3$$
$$P(D|I) = 0.9$$

Usando la Regla de la multiplicación se encuentra la probabilidad deseada.

$$P(ID) = P(I)P(D|I) = 0.3(0.9) = 0.27$$

Problema 10._ Se dio a una nueva secretaria n contraseñas para la computadora, pero solamente una de ellas dará acceso a un archivo. La secretaria no sabe cuál es la contraseña correcta y por tanto, escoge una al azar y la prueba. Si la contraseña es incorrecta, la quita y selecciona aleatoriamente otra de las que quedan, continuando de esta manera hasta encontrar la contraseña correcta.
a) ¿Cuál es la probabilidad de que obtenga la contraseña correcta en el primer intento?
b) ¿Cuál es la probabilidad de que obtenga la contraseña correcta en el segundo intento? y ¿en el tercero?
c) Se estableció un sistema de seguridad de tal manera que si se intentan 3 contraseñas incorrectas antes de encontrar la buena se cierra el archivo y se niega el acceso. Si $n = 7$ ¿Cuál es la probabilidad de que la secretaria tenga acceso al archivo?

<u>Solución</u>

Sea el evento C_i: contraseña correcta al intento i.

a) Sólo una contraseña es correcta de un total de n contraseñas disponibles, es decir $P(C_1) = \frac{1}{n}$

b) El evento *contraseña correcta al intento dos*, es en realidad la intersección de los eventos *contraseña incorrecta al intento uno* y *contraseña correcta al intento dos*. Usando la Regla de la multiplicación y el hecho de que sólo una contraseña es la correcta la probabilidad de que la primera contraseña sea incorrecta es el cociente entre el número de contraseñas incorrectas $(n-1)$ y el número de contraseñas en total (n) y la probabilidad de que la segunda contraseña sea correcta dado que la primera fue incorrecta es el cociente del número de contraseñas correctas (1) entre el número total de contraseñas que se ha reducido en una ya que la primera se deja fuera, es decir:

$$P(C_2) = P\left(C_1' C_2\right) = P\left(C_1'\right)P\left(C_2|C_1'\right) = \frac{n-1}{n}\left(\frac{1}{n-1}\right) = \frac{1}{n}$$

De manera semejante el evento *contraseña correcta al intento tres* es la intersección de los eventos *contraseña incorrecta al intento uno*, *contraseña incorrecta al intento dos* y *contraseña correcta al intento tres*. Usando la Regla de la multiplicación y el hecho de que una vez que se ha probado una contraseña incorrecta esta es desechada se tiene que:

$$P(C_3) = P\left(C_1' C_2' C_3\right) = P\left(C_1'\right) P\left(C_2' \middle| C_1'\right) P\left(C_3 \middle| C_1' C_2'\right) = \frac{n-1}{n}\left(\frac{n-2}{n-1}\right)\left(\frac{1}{n-2}\right) = \frac{1}{n}$$

c) La secretaria tendrá acceso al archivo si logra encontrar la contraseña correcta en cualquiera de los tres primeros intentos y como son eventos disjuntos la probabilidad de la unión de los eventos $(C_1 \cup C_2 \cup C_3)$ es la suma de las probabilidades individuales. Usando $n = 7$ se tiene que la probabilidad de que la secretaria tenga acceso al archivo es:

$$P(\text{acceso}) = P(C_1 \cup C_2 \cup C_3) = P(C_1) + P(C_2) + P(C_3) = \frac{3}{7}$$

Problema 11._ La urna 1 contiene x esferas blancas y y rojas. La urna 2 contiene z blancas y v rojas. Se escoge una esfera al azar de la urna 1 y se pone en la urna 2. Entonces se escoge una esfera al azar de la urna 2. ¿Cuál es la probabilidad de que esta esfera sea blanca?

Solución

Sean los eventos b_i : bola blanca de la urna i con i = 1, 2.
r_i : bola roja de la urna i con i = 1, 2.

La bola que se extrae de la urna 2 puede ser blanca bajo dos situaciones diferentes y excluyentes; la bola que se extrae de la urna 1 es blanca y la bola que se extrae de la urna 2 también es blanca o la bola que se extrae de la urna 1 es roja y la bola que se extrae de la urna 2 es blanca. Usando el hecho de que la probabilidad de unión de eventos disjuntos es la suma de las probabilidades individuales y la Regla de la multiplicación se tienen que la probabilidad de interés está dada por:

$$\begin{aligned}
P(b_2) &= P(b_1 b_2 \cup r_1 b_2) \\
&= P(b_1 b_2) + P(r_1 b_2) \\
&= P(b_1) P(b_2 | b_1) + P(r_1) P(b_2 | r_1) \\
&= \frac{x}{x+y}\left(\frac{z+1}{z+v+1}\right) + \frac{y}{x+y}\left(\frac{z}{z+v+1}\right)
\end{aligned}$$

Problema 12._ Un prisionero político será enviado a Siberia o a los Urales. Las probabilidades de que lo envíen a estos dos lugares son 0.6 y 0.4 respectivamente. Se sabe además que si un residente de Siberia se elige al azar hay una probabilidad de 0.5 de que lleve un abrigo de piel, en tanto que la probabilidad para lo mismo es de 0.7 en los Urales. Al llegar al exilio, la primera persona que ve el prisionero no lleva un abrigo de piel ¿Cuál es la probabilidad de que esté en Siberia?

<u>Solución</u>

Sean los eventos S: El prisionero está en Siberia.
U: El prisionero está en los Urales.
A: Una persona cualquiera usa abrigo.

Los datos que tenemos son:

$$P(S) = 0.6 \qquad P(A|S) = 0.5 \qquad \therefore P(A'|S) = 0.5$$
$$P(U) = 0.4 \qquad P(A|U) = 0.7 \qquad \therefore P(A'|U) = 0.3$$

Usando el Teorema de Bayes para encontrar la probabilidad de interés se tiene que:

$$\begin{aligned} P(S|A') &= \frac{P(SA')}{P(A')} \\ &= \frac{P(S)P(A'|S)}{P(S)P(A'S) + P(U)P(A'|U)} \\ &= \frac{0.6(0.5)}{0.6(0.5)+0.4(0.3)} = 0.714 \end{aligned}$$

Problema 13._ Una ciudad tiene dos carros de bomberos que operan de forma independiente. La probabilidad de que un carro específico esté disponible cuando se le necesite es 0.96. ¿Cuál es la probabilidad de que:
a) ninguno esté disponible cuando se le necesite?
b) un carro de bomberos esté disponible cuando se le necesite?

<u>Solución</u>

Sea el evento d_i : carro i disponible con $i = 1, 2$.
Se sabe que $P(d_i) = 0.96$ y que hay independencia entre la operación de los carros de bomberos. Recordemos que si dos eventos son independientes también lo son sus complementos.

a) $P\left(d_1' d_2'\right) = P\left(d_1'\right)P\left(d_2'\right) = (0.04)^2 = 0.0016.$

b)
$$\begin{aligned} P(d_1 \cup d_2) &= P(d_1) + P(d_2) - P(d_1)P(d_2) \\ &= 2(0.96) - (0.96)^2 \\ &= 0.9984. \end{aligned}$$

Problema 14._ Supón que A y B son eventos independientes, tales que la probabilidad de que no ocurra ninguno de los dos es a y la probabilidad de que ocurra B es b. Demuestra que $P(A) = \dfrac{1-b-a}{1-b}$.

Solución

Los datos en términos de probabilidades son:

$$P\left[(A \cup b)'\right] = a$$

$$P(B) = b$$

Con un poco de álgebra tenemos que:

$$\begin{aligned} P\left[(A \cup b)'\right] &= 1 - P(A \cup b) \\ &= 1 - P(A) - P(B) + P(A)P(B) \\ &= 1 - P(B) - P(A)[1 - P(B)] = a \end{aligned}$$

Despejando $P(A)$

$$P(A) = \frac{1 - P(B) - a}{1 - P(B)} = \frac{1-b-a}{1-b}$$

Problema 15._ Un detector de mentiras muestra una lectura positiva (es decir, indica una mentira) en 10% de los casos cuando la persona dice la verdad y en 95% de los casos cuando la persona miente. Supón que se sospecha de dos personas de haber cometido un delito, que fue ejecutado por una sola persona, y de hecho sólo una de ellas es la culpable. ¿Cuál es la probabilidad de que el detector:

a) muestre una lectura positiva para los dos sospechosos?
b) muestre una lectura positiva para el sospechoso culpable y una lectura negativa para el inocente?
c) esté completamente equivocado, es decir, que indique una lectura positiva para el inocente y una negativa para el culpable?
d) dé una lectura positiva para cualquiera de los dos o para ambos sospechosos?

Solución

Sean los eventos
M: detector indica una mentira.
V: la persona dice la verdad.

Los datos en términos de probabilidades son:

$$P(M|V) = 0.1 \quad \therefore P(M'|V) = 0.9$$

$$P(M|V') = 0.95 \quad \therefore P(M'|V') = 0.05$$

a) Una lectura positiva para los dos sospechosos significa *lectura positiva para el que miente* y *lectura positiva para el que dice la verdad*, es decir se trata de una intersección de eventos y dado que hay independencia entre los sospechosos la probabilidad de la intersección es el producto de las probabilidades:

$$P(M|V')P(M|V) = 0.95(0.1) = 0.095$$

b) Probabilidad de la intersección de los eventos *lectura positiva para el que miente* y *lectura negativa para el que dice la verdad*, dado que hay independencia entre los sospechosos nuevamente es el producto de las probabilidades:

$$P(M|V')P(M'|V) = 0.95(0.9) = 0.855$$

c) Lectura positiva para el que dice la verdad y lectura negativa para el que miente.

$$P(M|V)P(M'|V') = 0.1(0.05) = 0.005$$

d) Lectura positiva para el que dice la verdad, lectura positiva para el que miente o lectura positiva para el que miente y para el que dice la verdad.

$$P(M|V)P(M'|V') + P(M'|V)P(M|V') + P(M|V')P(M|V) = 0.1(0.05) + 0.9(0.95) + 0.95(0.1) = 0.955.$$

Problema 16._ La víctima de un accidente morirá a menos de que reciba en los próximos 10 minutos una cantidad de sangre tipo A, Rh positivo, que sea suministrada por un solo donante. Se tarda 2 minutos en definir el tipo de sangre de un posible donante y 2 minutos en realizar la transfusión. Hay una gran cantidad de donantes diferentes cuyo tipo de sangre se desconoce y 40% de ellos tienen el tipo de sangre A, Rh positivo. ¿Cuál es la probabilidad de que sobreviva la víctima si solamente se dispone de un equipo para determinar el tipo de sangre?

<u>Solución</u>

La víctima sobrevive si entre los primeros cuatro donantes hay uno con el tipo de sangre A, Rh positivo. Usando el complemento del evento de interés la víctima muere si ninguno de los cuatro primeros donantes tiene el tipo de sangre A, Rh positivo. Usando el hecho de que hay independencia entre los donantes y que la probabilidad de que no tenga el tipo de sangre de interés es 0.6 se tiene que:

$$P(sobreviva) = 1 - P(muera) = 1 - (0.6)^4 = 0.8704.$$

Problema 17._ Un número binario está compuesto sólo de los dígitos 0 y 1 (Por ejemplo 1011, 1100, etc.). Estos números tienen un papel importante en el uso de las computadoras. Supón que un número binario está formado por *n* dígitos. Supón que la probabilidad de que aparezca un dígito incorrecto es *p* y que los errores en dígitos diferentes son independientes unos de otros. ¿Cuál es la probabilidad de formar un número incorrecto?

Solución

Sean los eventos d_i :dígito incorrecto.

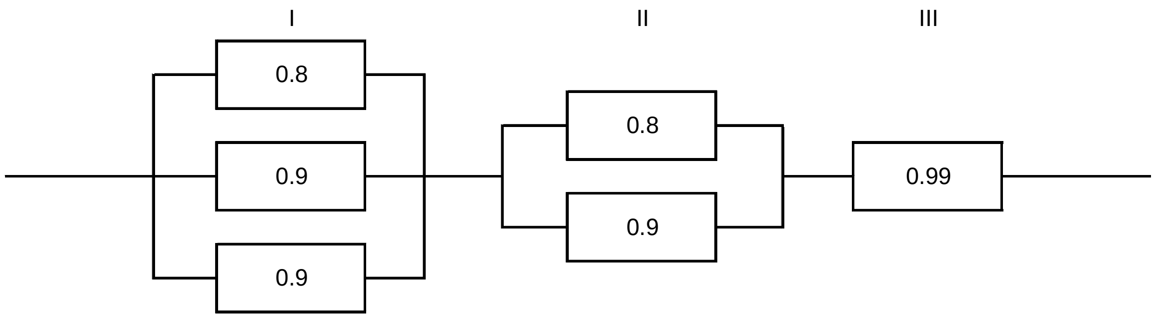
N_c : número binario (formado por n dígitos) correcto.

$P(d_i) = p \quad \therefore P(d_i') = 1 - p$

Usando complemento de un evento y la independencia entre números binarios se tiene que:

$P(N_c') = 1 - P(N_c) = 1 - (1-p)^n$

Problema 18._ Considera el diagrama de un sistema electrónico que muestra las probabilidades de que los componentes del sistema operen de modo apropiado. ¿Cuál es la probabilidad de que el sistema opere si el ensamble III y al menos uno de los componentes en los ensambles I y II deben operar para que funcione el ensamble? Supón que los componentes de cada ensamble operan independientemente y que la operación de cada ensamble también es independiente.



Solución

Sean los eventos: *A*: Funciona al menos un componente del ensamble I.
B: Funciona al menos un componente del ensamble II.
C: Funciona el ensamble III.
S: Funciona el sistema.

El ensamble I no funciona si los tres componentes no funcionan, es decir:

$P(A) = 1 - P(A') = 1 - 0.2(0.1)^2 = 0.998$

El ensamble II no funciona si los dos componentes no funcionan, es decir:

$P(B) = 1 - P(B') = 1 - 0.2(0.1) = 0.98$

El sistema funciona si los tres ensambles operan correctamente, es decir:

$P(ABC) = P(A)P(B)P(C) = 0.998(0.98)(0.99) = 0.968$

UNIDAD II Variables aleatorias discretas y continuas.

I) CASO DISCRETO.

Problema 1._ Considera un sistema de agua que fluye a través de unas válvulas de A a B. Las válvulas 1, 2 y 3 funcionan independientemente y cada una se abre correctamente mediante una señal con una probabilidad de 0.8. Encuentra la distribución de probabilidad para *Y*, el número de vías abiertas de A a B después de haber enviado la señal.

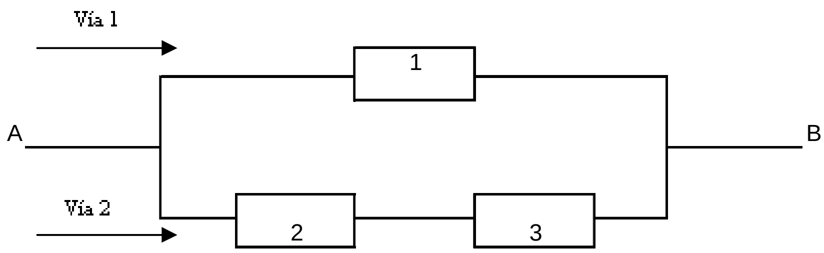


Diagrama: Sistema de válvulas para flujo de agua.

<u>Solución</u>

Sea la variable aleatoria *Y*: Número de vías abiertas, con valores posibles $y = 0, 1, 2$.
V_i : Vía i abierta, con $i = 1, 2$.
v_i : Válvula i abierta correctamente con $i = 1, 2$.

La vía 1 está abierta si la válvula 1 abre correctamente, lo cual ocurre con una probabilidad de 0.8.

$$P(V_1) = P(v_1) = 0.8 \qquad P(V_1^{'}) = 0.2$$

La vía 2 está abierta sólo si las válvulas 2 y 3 abren correctamente, es decir:

$$P(V_2) = P(v_2 v_3) = P(v_2)P(v_3) = (0.8)^2 = 0.64 \;\therefore P(V_2^{'}) = 0.36$$

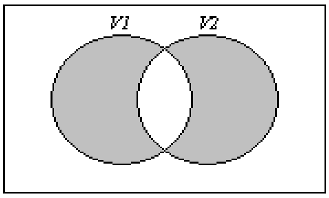
Las dos vías están cerradas con una probabilidad de: $P(V_1^{'}V_2^{'}) = P(V_1^{'})P(V_2^{'}) = 0.2(0.36) = 0.072$

Las probabilidades de que la v.a tome los valores 0, 1 y 2 son:

$$P(Y = 0) = 0.072$$

$$\begin{aligned}
P(Y = 1) &= P(V_1 \cup V_2) - P(V_1V_2) \\
&= P(V_1) + P(V_2) - 2P(V_1V_2) \\
&= 0.8 + 0.64 - 2(0.8)(0.64) \\
&= 0.416
\end{aligned}$$

El diagrama siguiente ilustra la probabilidad de la unión de los eventos de interés.



$$P(V_1 \cup V_2) = [P(V_1) + P(V_2) - P(V_1V_2)] - P(V_1V_2) = P(V_1) + P(V_2) - 2P(V_1V_2)$$

Y la probabilidad de que ambas vías estén abiertas es:

$$P(Y = 2) = P(V_1V_2) = 0.8(0.64) = 0.512$$

La función de probabilidad puede representarse en una tabla que incluya los valores posibles de la v.a y las probabilidades asociadas.

Función de probabilidad del número de vías abiertas.

$Y = y$	$P(Y = y)$
0	0.072
1	0.416
2	0.512
Suma	1

Problema 2._ De las personas que llegan a un banco de sangre, 1 de 3 tiene tipo sanguíneo O^+, y 1 de 15 tipo O^-. Considera 3 donantes, seleccionados aleatoriamente del banco de sangre. Sea *X* el número de donantes con sangre tipo O^+ y *Y* el número de donantes con sangre tipo O^-. Obtén las distribuciones de probabilidad para *X* y *Y*. Determina también la distribución de probabilidad para $X + Y$ el número de donantes con sangre tipo *O*.

<u>Solución</u>

La probabilidad de que un donante tenga sangre tipo O^+ es $P(O^+) = \frac{1}{3}$ entonces la probabilidad de que no tenga tipo O^+ es $P\left((O^+)'\right) = \frac{2}{3}$

Sea la variable aleatoria X : Número de donantes con sangre tipo O^+.

La probabilidad de que ninguno de los tres donantes tenga el tipo de sangre deseado es igual a la probabilidad de que los tres donantes tengan otro tipo de sangre, suponiendo independencia entre los donantes se tiene que la probabilidad de que los tres donantes no tengan sangre *O* positiva es igual al producto de las probabilidades de que cada uno de ellos no tenga el tipo de sangre *O* positiva. Es decir:

$$P(X = 0) = P\left((O^+)'\right)P\left((O^+)'\right)P\left((O^+)'\right) = \left(\frac{2}{3}\right)^3 = \frac{8}{27}$$

La probabilidad de que sólo un donante tenga el tipo de sangre de interés es igual a la probabilidad de que uno de ellos tenga el tipo de sangre deseado y los otros dos tengan otro tipo de sangre. Se deben contar todas las posibilidades, es decir ¿De cuántas maneras dos de los tres donantes tienen otro tipo de sangre?

$$\binom{3}{2} = \frac{3!}{2!\,1!} = 3$$

Por lo que se debe multiplicar por 3 la probabilidad de que un donante tenga el tipo de sangre de interés y los otros dos no lo tengan, asumiendo independencia entre los donantes se tiene que la probabilidad de que un donante tenga tipo de sangre *O* positiva está dada por:

$$P(X=1) = 3P(O^+)P\left((O^+)'\right)P\left((O^+)'\right) = 3\left(\frac{1}{3}\right)\left(\frac{2}{3}\right)^2 = \frac{12}{27}$$

La probabilidad de que dos donantes tengan tipo de sangre *O* positiva está dada por:

$$P(X=2) = 3P(O^+)^2 P\left((O^+)'\right) = 3\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right) = \frac{6}{27}$$

Razonando de esta manera se llena la tabla siguiente:

Función de probabilidad del número de donantes con sangre tipo O^+.

$X = x$	$P(X = x)$
0	$\left(\frac{2}{3}\right)^3 = \frac{8}{27}$
1	$3\left(\frac{2}{3}\right)^2\left(\frac{1}{3}\right) = \frac{12}{27}$
2	$3\left(\frac{2}{3}\right)\left(\frac{1}{3}\right)^2 = \frac{6}{27}$
3	$\left(\frac{1}{3}\right)^3 = \frac{1}{27}$

Para la variable aleatoria Y : Número de donantes con sangre tipo O^- se tiene que:

$$P(O^-) = \frac{1}{15} \text{ y } P(O^-)' = \frac{14}{15}$$

Función de probabilidad del número
de donantes con sangre tipo O^-.

$Y = y$	$P(Y = y)$
0	$\left(\frac{14}{15}\right)^3 = \frac{2744}{3375} = 0.813$
1	$3\left(\frac{1}{15}\right)\left(\frac{14}{15}\right)^2 = \frac{588}{3375} = 0.174$
2	$3\left(\frac{1}{15}\right)^2\left(\frac{14}{15}\right) = \frac{42}{3375} = 0.0124$
3	$\left(\frac{1}{15}\right)^3 = \frac{1}{3375} = 0.00029$

Los de tipo *O* positivo u *O* negativo tienen sangre tipo *O*, es decir:
$O = O^+ \cup O^-$. Dado que son eventos disjuntos ya que una persona no puede tener dos tipos de sangre diferentes se tiene que $P(O) = P(O^+) + P(O^-) = \frac{1}{3} + \frac{1}{15} = \frac{6}{15} = \frac{2}{5}$ y por lo tanto $P(O') = \frac{3}{5}$
$X + Y$: Número de donantes con sangre tipo *O*.

Función de probabilidad del número
de donantes con sangre tipo *O*.

$X + Y$	$P(X + Y)$
0	$\left(\frac{3}{5}\right)^3 = \frac{27}{125} = 0.216$
1	$3\left(\frac{3}{5}\right)^2\left(\frac{2}{5}\right) = \frac{54}{125} = 0.432$
2	$3\left(\frac{3}{5}\right)\left(\frac{2}{5}\right)^2 = \frac{36}{125} = 0.288$
3	$\left(\frac{2}{5}\right)^3 = \frac{8}{125} = 0.064$

Problema 3._ Sea X una v.a con posibles valores $x = 0,\ 1,\ 2$ tal que $P(X=0)=\frac{1}{3}$, $P(X=1)=\frac{1}{6}$ y $P(X=2)=\frac{1}{2}$. Encuentra y grafica la función de probabilidad acumulada para la variable aleatoria X.

Solución

La función de probabilidad acumulada está dada por $F_X(x) = P(X \le x)$.

$$F_X(0) = P(X=0) = \frac{1}{3}$$

$$F_X(1) = P(X=0) + P(X=1) = \frac{1}{3} + \frac{1}{6} = \frac{1}{2}$$

$$F_X(2) = P(X=0) + P(X=1) + P(X=2) = \frac{1}{3} + \frac{1}{6} + \frac{1}{2} = 1$$

Observa que:

$$F_X(0.5) = P(X \le 0.5) = P(X \le 0) = P(X=0) = \frac{1}{3}$$

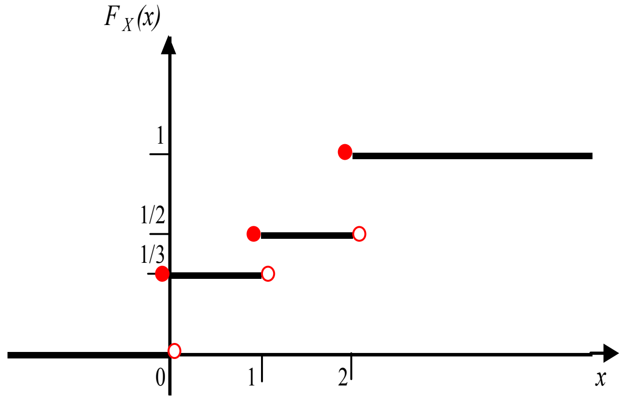
$$F_X(1.9) = P(X \le 1.9) = P(X=0) + P(X=1) = \frac{1}{2}$$

$$F_X(-1) = P(X \le -1) = 0$$

Por lo tanto la función de probabilidad acumulada es:

$$F_X(x) = \begin{cases} 0 & \text{si} \quad x < 0 \\ \frac{1}{3} & \text{si} \quad 0 \le x < 1 \\ \frac{1}{2} & \text{si} \quad 1 \le x < 2 \\ 1 & \text{si} \quad x \ge 2 \end{cases}$$

Y su gráfica es:



Función de probabilidad acumulada para la v.a X

Problema 4._ La demanda de un producto es –1, 0, 1, 2 por día con las probabilidades respectivas de $\frac{1}{5}, \frac{1}{10}, \frac{2}{5}, \frac{3}{10}$. Una demanda de –1 implica que se regresa una unidad. Encuentra la demanda esperada y la varianza.

<u>Solución</u>

Sea la v. a X la demanda por día del producto, con valores posibles $x = -1,\ 0,\ 1,\ 2$.

El valor esperado o esperanza de la variable está dado por:

$$\begin{aligned} E(X) &= \sum_x xP(X = x) \\ &= (-1)P(X = -1) + 0P(X = 0) + 1p(X = 1) + 2P(X = 2) \\ &= -\frac{1}{5} + \frac{2}{5} + 3\left(\frac{3}{10}\right) = -\frac{1}{5} + \frac{2}{5} + \frac{6}{10} = \frac{-2+4+6}{10} \qquad \therefore\ E(X) = 0.8 \end{aligned}$$

La varianza de la variable aleatoria está dada por el segundo momento menos el cuadrado del primero, es decir: $V(X) = E\left(X^2\right) - \left(E(X)\right)^2$.

Con $E\left(X^2\right) = \sum_x x^2 P(X = x) = (-1)^2 \frac{1}{5} + (1)^2 \frac{2}{5} + (2)^2 \frac{3}{10} = \frac{1}{5} + \frac{2}{5} + \frac{12}{10} = \frac{2+4+12}{10} = \frac{18}{10} = \frac{9}{5} = 1.8$

Entonces la varianza es: $V(X) = 1.8 - (0.8)^2 = 1.16$

Problema 5._ Una v.a discreta X tiene la función de probabilidad $f_X(x) = k\left(\frac{1}{2}\right)^x \quad x = 1, 2, 3$.

a) Determina el valor de *k*.
b) Encuentra la media y la varianza de *X*.
c) Encuentra la función de distribución acumulada $F_X(x)$.

<u>Solución</u>

a) Para que la función $f_X(x) = k\left(\frac{1}{2}\right)^x$ con $x = 1, 2, 3$ sea una f.d.p tiene que cumplir que $f_X(x) \in [0, 1]$ y además $\sum_x f_X(x) = 1$.

La primera propiedad dice que $k > 0$ y con la segunda propiedad se tiene que:

$$1 = k\left[\frac{1}{2} + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^3\right] = k\left[\frac{1}{2} + \frac{1}{4} + \frac{1}{8}\right] = k\frac{7}{8} \qquad \therefore\ k = \frac{8}{7}$$

b) Encuentra la media y la varianza de *X*.

$$E(X) = \sum_x x f_X(x) = 1\left(\frac{4}{7}\right) + 2\left(\frac{2}{7}\right) + 3\left(\frac{1}{7}\right) = \frac{11}{7}$$

El cálculo de la varianza es como sigue:

$$V(X)=E\left(X^2\right)-E^2(X)=\left[1\left(\frac{4}{7}\right)+2^2\left(\frac{2}{7}\right)+3^2\left(\frac{1}{7}\right)\right]-\left(\frac{11}{7}\right)^2=\frac{26}{49}=0.53$$

c) Encuentra la función de distribución acumulada $F_X(x)$.

$$F_X(x)=\begin{cases}0 & x<1\\ \frac{4}{7} & 1\le x<2\\ \frac{6}{7} & 2\le x<3\\ 1 & x\ge 3\end{cases}$$

Problema 6._ La v.a discreta N ($n=0,1,\ldots$) tiene probabilidades de ocurrencia de kr^n ($0<r<1$). Encuentra el valor apropiado de k.

__Solución__

La f.d.p de la variable aleatoria es $f_N(n)=P(N=n)=kr^n$ con valores posibles $n=0,\ 1,\ 2\ldots$ y $0<r<1$. Por propiedades de la función de probabilidad se tiene que:

$$1=\sum_n f_N(n)=\sum_{n=0}^{\infty}kr^n=k\sum_{n=0}^{\infty}r^n=k\left(\frac{1}{1-r}\right)\quad\therefore\quad k=1-r$$

Problema 7._ Supón que la v.a J tiene valores posibles $1,2,3,\ldots$ y $P(J=j)=\frac{1}{2^j},\quad j=1,2,\ldots$

a) Verifica que es una función de probabilidad.
b) Calcula P(J sea par).
c) Calcula $P(J\ge 5)$.
d) Calcular P(J sea divisible entre 3).

__Solución__

a) Para que la función sea una f.d.p debe cumplir las dos propiedades:
Propiedad_1 La $P(J=j)\in[0,1]$

$$P(J=j)=\frac{1}{2^j}\text{ Siempre es positiva para } j=1,2,\ldots$$

Propiedad_2 La suma de las probabilidades para todos los valores posibles de la variable aleatoria debe ser uno.

$$\sum_j f_J(j)=\sum_{j=1}^{\infty}\frac{1}{2^j}=\sum_{j=1}^{\infty}\left(\frac{1}{2}\right)^j=\frac{\frac{1}{2}}{1-\frac{1}{2}}=1$$

b) $P(X \text{ sea par}) = P(X=2)+P(X=4)+P(X=6)+\cdots$

$$= \frac{1}{2^2}+\frac{1}{2^4}+\frac{1}{2^6}+\cdots = \left[\left(\frac{1}{2}\right)^2\right]^1 + \left[\left(\frac{1}{2}\right)^2\right]^2 + \left[\left(\frac{1}{2}\right)^2\right]^3 + \cdots$$

$$= \sum_{i=1}^{\infty}\left[\left(\frac{1}{2}\right)^2\right]^i = \frac{\left(\frac{1}{2}\right)^2}{1-\left(\frac{1}{2}\right)^2} = \frac{\frac{1}{4}}{1-\frac{1}{4}} = \frac{\frac{1}{4}}{\frac{3}{4}} = \frac{4}{3(4)} = \frac{1}{3}$$

c) $P(X \geq 5) = 1 - P(X \leq 4)$

$$= 1 - \left[\left(\frac{1}{2}\right)^1 + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^3 + \left(\frac{1}{2}\right)^4\right] = 1 - \left[\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\frac{1}{16}\right]$$

$$= 1 - \frac{8+4+2+1}{16} = 1 - \frac{15}{16} = \frac{1}{16}$$

d) $P(\text{divisible entre } 3) = P(X=3)+P(X=6)+P(X=9)+P(X=12)+\cdots$

$$= \left(\frac{1}{2}\right)^3 + \left(\frac{1}{2}\right)^6 + \left(\frac{1}{2}\right)^9 + \left(\frac{1}{2}\right)^{12} + \cdots$$

$$= \left[\left(\frac{1}{2}\right)^3\right]^1 + \left[\left(\frac{1}{2}\right)^3\right]^2 + \left[\left(\frac{1}{2}\right)^3\right]^3 + \left[\left(\frac{1}{2}\right)^3\right]^4 + \cdots$$

$$= \sum_{i=1}^{\infty}\left(\frac{1}{2}\right)^3 = \frac{\left(\frac{1}{2}\right)^3}{1-\left(\frac{1}{2}\right)^3} = \frac{\frac{1}{8}}{\frac{7}{8}} = \frac{8}{7(8)} = \frac{1}{7}$$

Problema 8._ Sea X una v.a con resultados posibles 0, 1, 2,... Supón que $P(X=x)=(1-a)a^x$.
a) ¿Para qué valores de a es significativo el modelo anterior?
b) Demuestra que para dos enteros positivos cualesquiera s y t, $P(X > s+t \mid X > t) = P(X \geq s)$

Solución

a) La función de probabilidad debe cumplir que la suma de las probabilidades sobre todos los valores posibles de la v.a debe ser uno, es decir: $\sum_{x=0}^{\infty} P(X=x)=1.$

Sustituyendo el valor de la probabilidad se tiene que: $\sum_{x=0}^{\infty}(1-a)a^x = (1-a)\sum_{x=0}^{\infty}a^x = (1-a)\frac{1}{1-a} = 1.$

Esta igualdad tiene sentido sólo si $|a| < 1$, entonces el modelo es significativo para valores de a entre (0, 1).

b) Usando la definición de probabilidad condicional:

$$P(X > s+t | X > t) = \frac{P(X > t+s, X > t)}{P(X > t)} = \frac{P(X > t+s)}{P(X > t)} = \frac{1 - P(X \le t+s)}{1 - P(X \le t)}$$

$$= \frac{1 - \sum_{j=0}^{t+s} (1-a)a^j}{1 - \sum_{j=0}^{t} (1-a)a^j} = \frac{1 - (1-a)\sum_{j=0}^{t+s} a^j}{1 - (1-a)\sum_{j=0}^{t} a^j}$$

$$= \frac{1 - (1-a)\left[\frac{1 - a^{t+s+1}}{1-a}\right]}{1 - (1-a)\left[\frac{1 - a^{t+1}}{1-a}\right]} = \frac{a^{t+s+1}}{a^{t+1}} = a^s$$

Por otro lado se tiene que:

$$P(X \ge s) = 1 - P(X \le s-1) = 1 - (1-a)\sum_{j=0}^{s-1} a^j = 1 - (1-a)\left[\frac{1 - a^{s-1+1}}{1-a}\right] = a^s$$

$$\therefore P(X > s+t | X > t) = P(X \ge s)$$

Problema 9._ Determina la media y la varianza de la v.a X si se sabe que su función generadora de momentos (f.g.m) está dada por $\psi_X(t) = \frac{1}{4}\left(3e^t + e^{-t}\right)$ para $-\infty < t < \infty$.

Solución

La primera derivada de la f.g.m evaluada en 0 es la media de la v.a y la segunda derivada de la f.g.m evaluada en cero es el segundo momento al origen, entonces:

$$\psi'_X(t=0) = \frac{1}{4}\left(3e^t - e^{-t}\right)\Big|_{t=0} = \frac{1}{4}(3-1) = \frac{1}{2}$$

$$\therefore \mu_X = \frac{1}{2}$$

$$\psi''_X(t=0) = \frac{1}{4}\left(3e^t + e^{-t}\right)\Big|_{t=0} = \frac{1}{4}(3+1) = 1 = E\left(X^2\right)$$

$$V(X) = E\left(X^2\right) - \mu_X^2 = 1 - \left(\frac{1}{2}\right)^2$$

$$\therefore V(X) = \frac{3}{4}$$

Problema 10._ Sea *X* una v.a con media μ y varianza σ^2 y sea $\psi_X(t)$ su f.g.m para $-\infty < t < \infty$. Si *c* es una constante positiva y *Y* una v.a con f.g.m $\psi_Y(t) = e^{c[\psi_X(t)-1]}$ para $-\infty < t < \infty$. Determina las expresiones de la media y la varianza de *Y* en función de la media y la varianza de *X*.

Solución

La media de la v.a Y es la primera derivada de su f.g.m evaluada en cero.

$$\mu_Y = \psi'_Y(t=0)$$
$$= e^{c[\psi_X(t)-1]} c\psi'_X(t)\Big|_{t=0}$$
$$= e^{c(1-1)} c\mu$$
$$\therefore \mu_Y = c\mu$$

Ya que $\psi_X(t) = E\left(e^{tX}\right) \therefore \psi_X(0) = E\left(e^0\right) = E(1) = 1$

Para el cálculo de la varianza, se necesita la segunda derivada de la f.g.m de *Y* evaluada en cero.

$$\psi'_Y(t) = e^{c[\psi_X(t)-1]} c\psi'_X(t)$$
$$\psi'_Y(t) = ce^{-c} e^{c\psi_X(t)} \psi'_X(t)$$

$$\psi''_Y(t) = ce^{-c}\left[e^{c\psi_X(t)}\psi''_X(t) + \psi'_X(t) e^{c\psi_X(t)} c\psi'_X(t)\right]\Big|_{t=0}$$
$$\psi''_Y(0) = ce^{-c}\left[e^{c\psi_X(0)}\psi''_X(0) + \mu^2 e^{c\psi_X(0)} c\right]$$
$$= ce^{-c}\left(e^c\psi''_X(0) + c\mu^2 e^c\right)$$
$$= c\psi''_X(0) + c^2\mu^2$$

Pero $\psi''_X(0) = E\left(X^2\right) = V(X) + E^2(X) = \sigma^2 + \mu^2$

Entonces $\psi''_Y(0) = c\left(\sigma^2 + \mu^2\right) + c^2\mu^2$

Por lo tanto $V(Y) = E\left(Y^2\right) - E^2(Y)$

$$= c\sigma^2 + c\mu^2 + c^2\mu^2 - c^2\mu^2$$
$$\therefore \quad \sigma_Y^2 = c\left(\sigma^2 + \mu^2\right)$$

Problema 11._ Cinco pelotas numeradas del 1 al 5 se encuentran en una urna. Se sacan 2 pelotas al azar y se anotan sus números. Para el mayor número seleccionado encuentra:
a) La función de probabilidad.
b) Valor esperado y varianza.
c) Función generadora de momentos.
d) Función de probabilidad acumulada.

Solución

a) Sea la v.a de interés X, el mayor de los dos números seleccionados. En el cuadro siguiente se presentan todos los resultados posibles del experimento y el valor que en cada caso toma la v.a.

Resultados posibles y valores de la v.a.

Pareja Posible	$X = x$
(1,2)	2
(1,3)	3
(1,4)	4
(1,5)	5
(2,3)	3
(2,4)	4
(2,5)	5
(3,4)	4
(3,5)	5
(4,5)	5

Función de probabilidad del mayor número seleccionado.

$X = x$	$f_X(x) = P(X = x)$
2	0.1
3	0.2
4	0.3
5	0.4

La f.d.p también se puede expresar con la fórmula $f_X(x) = \dfrac{x-1}{10}$ para $x = 2,\ 3, 4.$

b) Para el cálculo de la media y la varianza es bueno crear un cuadro auxiliar que será de gran ayuda.

$X = x$	$f_X(x) = P(X = x)$	$xf_X(x)$	$x^2 f_X(x)$
2	0.1	0.2	0.4
3	0.2	0.6	1.8
4	0.3	1.2	4.8
5	0.4	2	10

La esperanza de la v.a es entonces $E(X) = \sum_{x=2}^{5} xf_X(x) = 0.2 + 0.6 + 1.2 + 2 = 4$

Para calcular la varianza se necesita el segundo momento:

$$E(X^2) = \sum_{x=2}^{5} x^2 f_X(x) = 0.4 + 1.8 + 4.8 + 10 = 17$$

Por la tanto la varianza es $V(X) = E(X^2) - E^2(X) = 17 - (4)^2 = 1$

c) La f.g.m según la definición es:

$$\psi_X(t) = E(e^{tx}) = \sum_{x=2}^{5} e^{tx} P(X = x) = 0.1e^{2t} + 0.2e^{3t} + 0.3e^{4t} + 0.4e^{5t}$$

O equivalentemente $\psi_X(t) = \frac{1}{10}\sum_{x=2}^{5}(x-1)e^{tx}$

d) Para la función de probabilidad acumulada $F_X(x) = P(X \le x)$, primero se encuentra la acumulada en cada uno de los valores posibles de la v.a.

$$F_X(2) = P(X \le 2) = P(X = 2) = 0.1$$

$$F_X(3) = P(X \le 3) = P(X = 2) + P(X = 3) = 0.1 + 0.2 = 0.3$$

$$F_X(4) = P(X \le 4) = P(X = 2) + P(X = 3) + P(X = 4) = 0.1 + 0.2 + 0.3 = 0.6$$

$$F_X(5) = P(X \le 5) = P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5) = 1$$

Finalmente podemos escribir la función de distribución acumulada como sigue:

$$F_X(x) = \begin{cases} 0 & x < 2 \\ 0.1 & 2 \le x < 3 \\ 0.3 & 3 \le x < 4 \\ 0.6 & 4 \le x < 5 \\ 1 & x \ge 5 \end{cases}$$

II) CASO CONTINUO.

Problema 1._ La variación en la profundidad de un río de un día al otro, medida (en pies) en un sitio específico, es una variable aleatoria *Y* con la siguiente función de densidad

$$f_Y(y) = k \quad -2 \le y \le 2$$

a) Determina el valor de *k*.
b) Encuentra la función de probabilidad acumulada para la v. a *Y*.

Solución

Sea la v.a *Y*: Variación en la profundidad de un río de un día al otro.

a) Se debe cumplir la propiedad de que la integral de la f.d.p en todo el rango de la v.a debe ser uno.

$$\int_{-2}^{2} k\,dy = 1 = ky\Big|_{-2}^{2} = k(2+2) = 4k \qquad \therefore\ k = \frac{1}{4}$$

b) $F_Y(y) = \int_{-2}^{y} \frac{1}{4}\, dt = \frac{1}{4} y \Big|_{-2}^{y} = \frac{y+2}{4}$

Por lo tanto la función de probabilidad acumulada es la siguiente función por partes:

$$F_Y(y) = \begin{cases} 0 & \text{si} \quad y < -2 \\ \dfrac{y+2}{4} & y \in (-2, 2) \\ 1 & \text{si} \quad y > 2 \end{cases}$$

Problema 2._ Un maestro de ESCOM nunca termina su clase antes que suene la campana, y siempre termina su clase a menos de 1 min después que suena la campana. Sea X el tiempo que transcurre entre la campana y el término de la clase y supón que la f.d.p de X es $f_X(x) = kx^2$ con $0 < x < 1$.

a) Encuentra el valor de k.

b) ¿Cuál es la probabilidad de que la clase termine a menos de medio minuto de que suene la campana?

c) ¿Cuál es la probabilidad de que la clase continúe entre 15 y 30 segundos después que suene la campana?

d) ¿Cuál es la probabilidad de que la clase continúe por lo menos 40 segundos después que suene la campana?

Solución

a) Sea la variable aleatoria X el tiempo que transcurre entre el sonido de la campana y el término de la clase.
Se sabe que la integral de la f.d.p en todo el rango de la v.a debe ser uno así que:

$$1 = \int_0^1 kx^2 dx = k\frac{x^3}{3}\Big|_0^1 = k\left(\frac{1}{3}\right) \quad \therefore \quad k = 3$$

b) $P(X \le 0.5) = 3\int_0^{0.5} x^2 dx = 3\frac{x^3}{3}\Big|_0^{0.5} = x^3\Big|_0^{0.5} = 0.5^3 = 0.125$

c) 15 segundos es la cuarta parte de un minuto así que la probabilidad pedida está dada por:

$$P(0.25 \le X \le 0.5) = 3\int_{0.25}^{0.5} x^2 dx = x^3\Big|_{0.25}^{0.5} = 0.5^3 - 0.25^3 = 0.125 - 0.015625 = 0.1094$$

d) 40 segundos son dos tercios de un minuto por lo que la probabilidad de interés es:

$$P\left(X \ge \frac{2}{3}\right) = 3\int_{2/3}^{1} x^2 dx = x^3\Big|_{2/3}^{1} = 1 - \left(\frac{2}{3}\right)^3 = 1 - \frac{8}{27} = \frac{19}{27} = 0.7037$$

Problema 3._ Para la siguiente función:

$$f_X(x) = \begin{cases} \dfrac{4(1-x^3)}{3} & 0 < x < 1 \\ 0 & \text{o.c} \end{cases}$$

a) Demuestra que es una función de densidad de probabilidad.
b) Encuentra la media y la varianza.
c) Encuentra la $F_X(x)$.

<u>Solución</u>

a) Hay que verificar que se cumplen las dos condiciones de una f.d.p.

i) Observa que para *x* entre 0 y 1 el valor de la función es siempre no negativo.

ii) $\displaystyle\int_{-\infty}^{\infty} f_X(x)dx = \int_0^1 \frac{4(1-x^3)}{3}dx = \frac{4}{3}\left[\int_0^1 dx - \int_0^1 x^3 dx\right] = \frac{4}{3}\left[x - \frac{x^4}{4}\right]\Bigg|_0^1 = \frac{4}{3}\left(\frac{3}{4}\right) = 1$

Por estas dos propiedades se concluye que esta función es una f.d.p.

b) $\displaystyle E(X) = \int_0^1 x f_X(x)dx = \int_0^1 x\frac{4(1-x^3)}{3}dx = \frac{4}{3}\left[\int_0^1 xdx - \int_0^1 x^4 dx\right]$

$$= \frac{4}{3}\left[\frac{x^2}{2} - \frac{x^5}{5}\right]_0^1 = \frac{4}{3}\left[\frac{1}{2} - \frac{1}{5}\right] = \frac{2}{5}$$

Para la varianza se necesita el segundo momento:

$$E(X^2) = \int_0^1 x^2 f_X(x)dx = \int_0^1 x^2 \frac{4(1-x^3)}{3}dx = \frac{4}{3}\left[\int_0^1 x^2 dx - \int_0^1 x^5 dx\right]$$

$$= \frac{4}{3}\left[\frac{x^3}{3} - \frac{x^6}{6}\right]_0^1 = \frac{4}{3}\left[\frac{1}{3} - \frac{1}{6}\right] = \frac{4}{3}\left(\frac{1}{6}\right) = \frac{2}{9}$$

$$\therefore\ V(X) = E(X^2) - \mu_X^2 = \frac{2}{9} - \left(\frac{2}{5}\right)^2 = \frac{50-36}{9(25)} = \frac{14}{225}$$

c) $\displaystyle F_X(x) = \int_0^x f_X(t)dt = \int_0^x \frac{4(1-t^3)}{3}dt = \frac{4}{3}\left[t - \frac{t^4}{4}\right]_0^x = \frac{4}{3}\left[x - \frac{x^4}{4}\right] = \frac{4x}{3} - \frac{x^4}{3}$

Entonces la función de probabilidad acumulada queda de la siguiente manera:

$$F_X(x) = \begin{cases} 0 & x < 0 \\ \dfrac{1}{3}(4x - x^4) & 0 < x < 1 \\ 1 & x > 1 \end{cases}$$

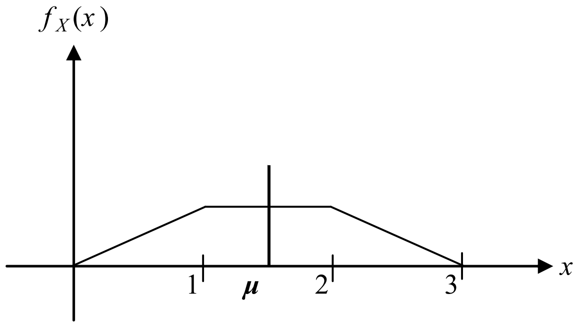
Problema 4._ Si la función de densidad de la variable aleatoria X está dada por

$$f_X(x) = \begin{cases} \frac{x}{2} & \text{para} \quad 0 < x \leq 1 \\ \frac{1}{2} & \text{para} \quad 1 < x \leq 2 \\ \frac{3-x}{2} & \text{para} \quad 2 < x < 3 \\ 0 & \text{en otra parte} \end{cases}$$

Determina el valor esperado de $g(X) = X^2 - 5X + 3$.

Solución

La gráfica de la f.d.p de la variable aleatoria es la siguiente:



El valor esperado de la función $g(X)$ está dado por:

$$E(g(X)) = E(X^2 - 5X + 3) = E(X^2) - 5E(X) + 3 \quad \text{--------------} \ (1)$$

Se necesita la esperanza y el segundo momento de la v. a X para sustituir en la ecuación (1).

$$E(X) = \int_{-\infty}^{\infty} x f_X(x)dx \quad \text{y} \quad E(X^2) = \int_{-\infty}^{\infty} x^2 f_X(x)dx$$

$$E(X) = \int_0^1 x\left(\frac{x}{2}\right)dx + \int_1^2 x\left(\frac{1}{2}\right)dx + \int_2^3 x\left(\frac{3-x}{2}\right)dx$$

$$= \int_0^1 \frac{x^2}{2}dx + \frac{1}{2}\int_1^2 xdx + \frac{1}{2}\int_2^3 (3x - x^2)dx$$

$$= \frac{1}{2}\left[\left.\frac{x^3}{3}\right|_0^1 + \left.\frac{x^2}{2}\right|_1^2 + \left.\frac{3x^2}{2}\right|_2^3 - \left.\frac{x^3}{3}\right|_2^3\right]$$

$$= \frac{1}{2}\left[\frac{1}{3} + \left(2 - \frac{1}{2}\right) + \frac{3}{2}(9-4) - \left(9 - \frac{8}{3}\right)\right]$$

$$= \frac{1}{2}\left[\frac{1}{3} + \frac{3}{2} + \frac{15}{2} - \frac{19}{3}\right] = \frac{1}{2}\left(\frac{2+9+45-38}{6}\right) = \frac{1}{2}\left(\frac{18}{6}\right) \quad \therefore \ E(X) = \frac{3}{2}$$

El cálculo del segundo momento es como sigue:

$$E\left(X^2\right)=\int_0^1 x^2\left(\frac{x}{2}\right)dx+\int_1^2 x^2\left(\frac{1}{2}\right)dx+\int_2^3 x^2\left(\frac{3-x}{2}\right)dx$$

$$=\int_0^1\frac{x^3}{2}dx+\frac{1}{2}\int_1^2 x^2dx+\frac{1}{2}\int_2^3\left(3x^2-x^3\right)dx=\frac{1}{2}\left[\left.\frac{x^4}{4}\right|_0^1+\left.\frac{x^3}{3}\right|_1^2+\left.x^3\right|_2^3-\left.\frac{x^4}{4}\right|_2^3\right]$$

$$=\frac{1}{2}\left[\frac{1}{4}+\left(\frac{8}{3}-\frac{1}{3}\right)+(27-8)-\left(\frac{81}{4}-\frac{16}{4}\right)\right]$$

$$=\frac{1}{2}\left[\frac{1}{4}+\frac{7}{3}+19-\frac{65}{4}\right]=\frac{1}{2}\left(\frac{3+28+228-195}{12}\right)$$

$$=\frac{1}{2}\left(\frac{64}{12}\right)\quad\therefore\quad E\left(X^2\right)=\frac{8}{3}$$

Finalmente se tiene que: $E(g(X))=E\left(X^2\right)-5E(X)+3$

$$E(g(X))=\frac{8}{3}-5\left(\frac{3}{2}\right)+3=\frac{16-45+18}{6}=-\frac{11}{6}$$

Problema 5._ Determina el primer y segundo momento al origen así como la desviación estándar para una variable aleatoria que tenga la densidad de probabilidad siguiente:

$$f_X(x)=\begin{cases}\dfrac{x}{2} & \text{para}\quad 0<x<2\\ 0 & \text{en otra parte}\end{cases}$$

Solución

El $k-$ésimo momento al origen está dado por $\mu_k^o=E\left(X^k\right)$, entonces:

$$\mu_1^o=E(X)=\int_0^2 x\left(\frac{x}{2}\right)dx=\frac{1}{2}\left.\left(\frac{x^3}{3}\right)\right|_0^2=\frac{4}{3}$$

$$\mu_2^o=E\left(X^2\right)=\int_0^2 x^2\left(\frac{x}{2}\right)dx=\frac{1}{2}\left.\left(\frac{x^4}{4}\right)\right|_0^2=2$$

La varianza de la v.a está dada por: $V(X)=E\left(X^2\right)-\mu^2=2-\left(\frac{4}{3}\right)^2=2-\frac{16}{9}=\frac{2}{9}$

Finalmente la desviación estándar de la v.a es $\sigma=\sqrt{V(X)}=\sqrt{\frac{2}{9}}=\frac{\sqrt{2}}{3}$

Problema 6._ Supón que el tiempo que tardan en atender a un individuo en una cafetería es una variable aleatoria con la densidad de probabilidad dada por:

$$f_X(x) = \begin{cases} \frac{1}{4}e^{-x/4} & \text{para} \quad x > 0 \\ 0 & \text{en otra parte} \end{cases}$$

Encuentra la media y la varianza de esta distribución.

<u>Solución</u>

$$E(X) = \int_0^\infty x\frac{1}{4}e^{-x/4}dx$$

Observa que es una integral por partes y recuerda que $\int e^u du = e^u$

Tomando $u = x$, $du = dx$, $dv = e^{-x/4}\left(\frac{1}{4}dx\right)$, $v = -e^{-x/4}$ y usando la fórmula $\int udv = uv - \int vdu$

$$\int_0^\infty x\frac{1}{4}e^{-x/4}dx = -xe^{-x/4}\Big|_0^\infty - \int_0^\infty -e^{-x/4}dx$$

$$= -xe^{-x/4}\Big|_0^\infty - 4\int_0^\infty e^{-x/4}dx\left(-\frac{1}{4}\right) = -xe^{-x/4}\Big|_0^\infty - 4e^{-x/4}\Big|_0^\infty \qquad \therefore\ \mu = 4$$

Segundo momento: $E\left(X^2\right) = \int_0^\infty x^2\frac{1}{4}e^{-x/4}dx$

Tomando $u = x^2$, $du = 2xdx$, $dv = e^{-x/4}\left(\frac{1}{4}dx\right)$, $v = -e^{-x/4}$ y usando la fórmula $\int udv = uv - \int vdu$

$$\int_0^\infty x^2\frac{1}{4}e^{-x/4}dx = -x^2e^{-x/4}\Big|_0^\infty - \int_0^\infty -e^{-x/4}2xdx$$

$$= 2\int_0^\infty xe^{-x/4}dx \quad \text{--------------- (2)}$$

Tomando $u = x$, $du = dx$, $dv = e^{-x/4}dx$, $v = -4e^{-x/4}$ y usando la fórmula $\int udv = uv - \int vdu$

$$\int_0^\infty xe^{-x/4}dx = -4xe^{-x/4}\Big|_0^\infty - \int_0^\infty -4e^{-x/4}dx$$

$$= 4\int_0^\infty e^{-x/4}dx = 4(-4)e^{-x/4}\Big|_0^\infty = 16$$

Sustituyendo en la ecuación (2) el segundo momento de la variable aleatoria es $E\left(X^2\right) = 2(16) = 32$.
La varianza de la v.a está dada por $V(X) = E\left(X^2\right) - \mu^2 = 32 - 4^2 = 16$.

Problema 7._ El tiempo requerido por los estudiantes para presentar un examen de una hora es una variable aleatoria con una función de densidad dada por:

$$f_X(x) = \begin{cases} cx^2 + x & 0 \le x \le 1 \\ 0 & \text{en otra parte} \end{cases}$$

a) Determina el valor de c.

b) Obtén la función de distribución acumulada $F_X(x)$.

c) Utiliza el inciso anterior para encontrar $F_X(-1)$, $F_X(0)$ y $F_X(1)$.

d) Calcula la probabilidad de que un estudiante termine en menos de media hora.

e) Dado que un estudiante necesita al menos 15 minutos para presentar el examen, encuentra la probabilidad de que necesite al menos 30 minutos para terminarlo.

<u>Solución</u>

a) La integral de la f.d.p en todo el rango de la v.a debe ser la unidad así que:

$$1 = \int_0^1 f_X(x)dx = \int_0^1 (cx^2 + x)dx = \left(c\frac{x^3}{3} + \frac{x^2}{2} \right)\Bigg|_0^1 = c\frac{1}{3} + \frac{1}{2} \quad \therefore \quad c = \frac{3}{2}$$

b) $$F_X(x) = \int_0^x f_X(t)dt = \int_0^t \left(\frac{3}{2}t^2 + t \right)dt = \left(\frac{3}{2}\frac{t^3}{3} + \frac{t^2}{2} \right)\Bigg|_0^x = \frac{1}{2}\left(x^3 + x^2\right)$$

La función de densidad acumulada se representa mediante la función por partes siguiente:

$$F_X(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{2}\left(x^3 + x^2\right) & 0 \le x \le 1 \\ 1 & x > 1 \end{cases}$$

c) Para evaluar la función de distribución acumulada en (-1) se debe tomar la primera parte de la función ya que (-1) es menor que cero. Para los valores cero y uno se tomara la segunda parte de $F_X(x)$.

$F_X(-1) = 0$, $F_X(0) = \frac{1}{2}\left(0^3 + 0^2\right) = 0 \quad F_X(1) = \frac{1}{2}\left(1^3 + 1^2\right) = 1$

d) $$\begin{aligned} P(X \le 0.5) &= F_X(0.5) \\ &= \frac{1}{2}\left[\left(\frac{1}{2}\right)^3 + \left(\frac{1}{2}\right)^2 \right] \\ &= \frac{1}{2}\left(\frac{1}{8} + \frac{1}{4} \right) = \frac{3}{16} = 0.1875 \end{aligned}$$

e) Se trata de una probabilidad condicional, además 15 minutos son la cuarta parte de una hora y 30 minutos la mitad de una hora.

$$
\begin{aligned}
P(X \geq 0.5 | X \geq 0.25) &= \frac{P(X \geq 0.25, X \geq 0.5)}{P(X \geq 0.25)} \\
&= \frac{P(X \geq 0.5)}{P(X \geq 0.25)} \\
&= \frac{1 - F_X(0.5)}{1 - F_X(0.25)} \\
&= \frac{1 - 0.1875}{1 - 0.0390625} \\
&= \frac{0.8125}{0.9609375} = 0.8455
\end{aligned}
$$

Utilizamos el hecho de que:

$$
\begin{aligned}
P(X \leq 0.25) &= F_X(0.25) \\
&= \frac{1}{2}\left[\left(\frac{1}{4}\right)^3 + \left(\frac{1}{4}\right)^2\right] \\
&= \frac{1}{2}\left(\frac{1}{64} + \frac{1}{16}\right) = \frac{5}{128} = 0.0390625
\end{aligned}
$$

UNIDAD III Distribuciones de probabilidad para variables discretas y continuas.

I) CASO DISCRETO.

Problema 1._ Un fabricante de válvulas admite que su control de calidad ha decaído, de modo que actualmente la probabilidad de producir una válvula defectuosa es 0.5. Si se fabrican un millón de válvulas al mes y eliges al azar entre estas válvulas 10,000 muestras cada una formada por 15 válvulas. ¿En cuántas muestras esperas encontrar

a) Exactamente 13 válvulas buenas?

b) Menos de 13 válvulas buenas?

Solución

Sea la variable aleatoria Y: Número de válvulas buenas en una muestra de tamaño 15.

La cual tiene una distribución binomial con parámetros 15 y 0.5, lo cual se escribe como:

$Y \sim \text{bin}(y;\ n_Y = 15,\ p_Y = 0.5)$.

Con ayuda de las tablas de distribución binomial se encuentran las probabilidades que se usarán en los incisos *a)* y *b)* respectivamente.

La probabilidad de exactamente 13 válvulas buenas en una muestra es:

$P(Y = 13) = F_Y(13) - F_Y(12) = 0.9995 - 0.9963 = 0.0032$

La probabilidad de menos de 13 válvulas buenas en una muestra es:

$P(Y < 13) = P(Y \le 12) = F_Y(12) = 0.9963$.

a) Sea X: Número de muestras en que hay exactamente 13 válvulas buenas.

$X \sim \text{bin}(x;\ n_X = 10{,}000,\ p_X = 0.0032)$

$E(X) = n_X p_X = 10{,}000(0.0032) = 32$ Muestras con exactamente 13 válvulas buenas.

b) Sea Z: Número de muestras con menos de 13 válvulas buenas.

$Z \sim \text{bin}(z;\ n_Z = 10{,}000,\ p_Z = 0.9963)$

$E(Z) = n_Z p_Z = 10{,}000(0.9963) = 9963$ Muestras con menos de 13 válvulas buenas.

Problema 2._ Un puente de cuota cobra \$1.00 por cada autobús de pasajeros y \$2.50 por otros vehículos. Supón que durante las horas diurnas, el 60% de todos los vehículos son autobuses de pasajeros. Si 25 vehículos cruzan el puente durante un periodo particular diurno, ¿Cuál es el ingreso resultante de cuotas esperado?

<u>Solución</u>

X: Número de autobuses de pasajeros.

P(autobuses de pasajeros) = 0.6

$X \sim \text{bin}(x;\ n = 25,\ p = 0.6)$

Primero se construye la función de costo en función de la v.a y luego se saca la esperanza:

$$C(X) = 1X + 2.5(25 - X) = X + 62.5 - 2.5X = 62.5 - 1.5X$$

$$E[C(X)] = 62.5 - 1.5E(X) = 62.5 - 1.5np = 62.5 - 1.5(25)(0.6) = 40.00$$

Problema 3._ Dada una distribución binomial con un valor fijo de n, ¿cuál es el valor de p en el que el valor de σ^2 es mayor?

<u>Solución</u>

Si $X \sim \text{bin}(x;\ n,\ p)$

y $V(X) = npq = np(1 - p) = n\left(p - p^2\right)$

Para maximizar el valor de la varianza hay que derivar respecto a p e igualar a cero:

$$\frac{\partial}{\partial p} V(X) = n(1 - 2p) = 0$$

1-2p = 0

Entonces la varianza alcanza su valor máximo cuando $p = \dfrac{1}{2}$.

Problema 4._ Un fabricante sabe que en promedio 20% de los tostadores eléctricos que fabrica requerirán de reparaciones dentro de un año después de su venta. Cuando se seleccionan al azar 20 tostadores, encuentra los números *x* y *y* apropiados tales que:
a) La probabilidad de que al menos *x* de ellos requieran reparaciones sea menor que 0.5.
b) La probabilidad de que al menos *y* de ellos no requieran reparaciones sea mayor de 0.8.

Solución

X : Número de tostadores que requieren reparación.
$X \sim \text{bin}(x;\ n = 20,\ p = 0.2)$

a) La probabilidad de que al menos *x* de ellos requieran reparaciones sea menor que 0.5 se escribe como:

$P(X \geq x) < 0.5$
$1 - F_X(x-1) < 0.5$
$1 - 0.5 < F_X(x-1)$

$0.5 < F_X(x-1)$ Se busca en tablas acumuladas para la binomial con $n = 20$ y $p = 0.2$ y se observa que esta condición se cumple cuando $x - 1 = 4 \quad \therefore \quad x = 5$.

b) La probabilidad de que al menos *y* de ellos no requieran reparaciones sea mayor de 0.8.

Y: Número de tostadores que no requieren reparaciones.

Observa que el éxito del experimento ha cambiado, ahora son de interés los tostadores que no requieren reparaciones.

$Y \sim \text{bin}(y;\ n = 20,\ p = 0.8)$
$P(Y \geq y) > 0.8$
$1 - F_Y(y-1) > 0.8$

$0.2 > F_Y(y-1)$ de las tablas binomiales se observa que $y - 1 = 14 \quad \therefore \quad y = 15$.

Problema 5._ Un agente de bienes raíces estima que la probabilidad de vender una casa es 0.1. El día de hoy tiene que ver 4 clientes. Si tiene éxito en las primeras tres visitas ¿Cuál es la probabilidad de que su cuarta visita no sea exitosa?

Solución

X: Número de visitas hasta un fracaso.
$q = P(\text{vender}) = 0.1$ $\qquad p = P(\text{no vender}) = 0.9$

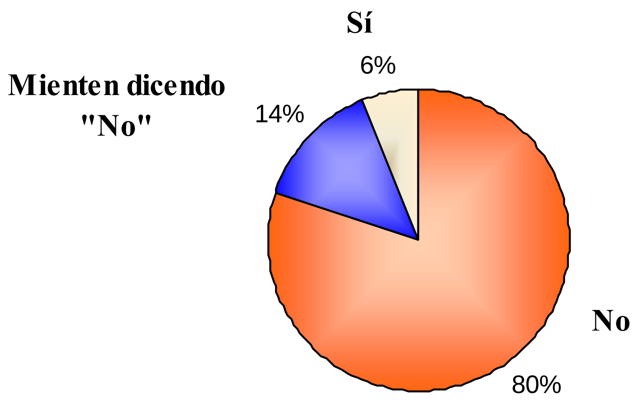
$X \sim G(x; p = 0.9)$

$P(X = 4 | X > 3) = P(X = 4) = (0.1^3)(0.9) = 0.0009$

Problema 6._ Al responder una pregunta con respecto a un tema controversial (como "¿Alguna vez ha fumado mariguana?"), muchas veces la gente no quiere contestar afirmativamente. Obtén la distribución de probabilidad para *Y*, el número de personas que se necesitaría entrevistar hasta obtener una sola respuesta afirmativa, sabiendo que el 80% de la población contestaría verídicamente "no" a la pregunta y que del 20% que deberían contestar verídicamente "sí", un 70% miente.

Solución

¿Alguna vez ha fumado mariguana? (Porcentaje)



P(obtener un Sí) = 0.3(0.2) = 0.06

Y: Número de personas que se necesita entrevistar hasta obtener el primer Sí.

La v.a Y tiene una distribución geométrica con probabilidad de éxito igual a 0.06, lo cual se escribe como $Y \sim G(y; p = 0.06)$.

La distribución de la v.a es de la forma $f_Y(y) = P(Y = y) = \left(0.94^{y-1}\right)(0.06)$ con $y = 1, 2, 3, \ldots$

Problema 7._ Supón que el costo de efectuar un experimento es $1,000. Si el experimento falla, se incurre en un costo adicional de $300 debido a ciertos cambios que deben efectuarse antes de que se intente un nuevo experimento. Si la probabilidad de éxito en cualquiera de los ensayos es 0.2, si los ensayos aislados son independientes y si los experimentos continúan hasta que se obtiene el primer resultado exitoso, ¿Cuál es el costo esperado del procedimiento completo?

Solución

X: Número de ensayos hasta el primer experimento exitoso.
P(éxito en un ensayo) = 0.2

$X \sim G(x; p = 0.2)$

La función de costo para este experimento está dada por:
$C = 1000X + 300(X-1) = 1300X - 300$

El costo esperado es la esperanza de la función de costo en función de la v.a geométrica.

$$\therefore \quad E(C) = 1300E(X) - 300 = 1300\frac{1}{p} - 300 = 1300\left(\frac{1}{0.2}\right) - 300.$$

El costo esperado del procedimiento completo es \$6,200.00.

Problema 8._ De acuerdo con un estudio publicado por un grupo de sociólogos de la Universidad de Massachusetts, cerca de dos tercios de los 20 millones de personas que en este país consumen Valium son mujeres. Supón que esta cifra es una estimación válida, encuentra la probabilidad de que en un día dado la quinta prescripción de Valium que escribe un doctor es la primera que prescribe Valium para una mujer.

Solución

X: Número de prescripciones de Valium hasta que se prescribe a una mujer.

$$X \sim G\left(x,\ p = \frac{2}{3}\right)$$

$$P(X = 5) = q^4 p = \left(\frac{1}{3}\right)^4 \frac{2}{3} = \frac{1}{54}$$

Problema 9._ Estudios de biología y el ambiente a menudo etiquetan y sueltan a sujetos a fin de estimar el tamaño y el grado de ciertas características en la población. Se capturan 10 animales de cierta población que se piensa extinta o cerca de la extinción, se etiquetan y se liberan en cierta región. Después de un período se selecciona en la región una muestra aleatoria (m.a) de 15 animales del tipo. ¿Cuál es la probabilidad de que cinco de estos seleccionados sean animales etiquetados si hay 25 animales de este tipo en la región?

Solución

X: Número de animales etiquetados.

$$X \sim H(x;\ r_1 = 10,\ r_2 = 15,\ n = 15)$$

$$P(X = 5) = \frac{\binom{10}{5}\binom{15}{10}}{\binom{25}{15}} = 0.2315$$

Problema 10._ Una fuerza de tarea gubernamental sospecha que algunas fábricas violan los reglamentos contra la contaminación ambiental con respecto a la descarga de cierto tipo de producto, 20 empresas están bajo sospecha pero no todas se pueden inspeccionar. Supón que tres de las empresas violan los reglamentos. ¿Cuál es la probabilidad de que
a) en la inspección de 5 empresas no se encuentre ninguna violación?
b) el plan anterior encuentre 2 que violan el reglamento?

<u>Solución</u>

X: Número de fábricas que violan el reglamento.

$X \sim H(x;\ r_1 = 3,\ r_2 = 17,\ n = 5)$

a) $$P(X=0) = \frac{\binom{3}{0}\binom{17}{5}}{\binom{20}{5}} = \frac{\frac{17!}{5!12!}}{\frac{20!}{5!15!}} = \frac{17!15!5!}{20!12!5!} = 0.3991$$

b) $$P(X=2) = \frac{\binom{3}{2}\binom{17}{3}}{\binom{20}{5}} = \frac{\frac{3!}{2!1!}\,\frac{17!}{3!14!}}{\frac{20!}{5!15!}} = \frac{17!5!15!}{2!14!20!} = 0.1315$$

Problema 11._ Para evitar la detección en la aduana un viajero coloca 6 tabletas de narcótico en una botella que contiene 9 píldoras de vitamina similares en apariencia. Si el oficial de la aduana selecciona 3 de las tabletas al azar para su análisis, ¿Cuál es la probabilidad de que el viajero sea arrestado por posesión ilegal de narcóticos?

<u>Solución</u>

X: Número de pastillas de narcótico.
$X \sim H(x;\ r_1 = 6,\ r_2 = 9,\ n = 3)$, $x = 0, \ldots, 3$.

6 narcóticos
9 vitaminas
____
15 total

El viajero será arrestado si el oficial encuentra al menos una pastilla de narcótico en la muestra seleccionada.

$$P(X=0) = \frac{\binom{9}{3}}{\binom{15}{0}} = 0.1846$$

Entonces la probabilidad de interés se encuentra por complemento: $P(X \geq 1) = 1 - P(X = 0) = 0.815$

Problema 12._ Supón que *X* tiene una distribución de Poisson. Si $P(X=2)=\frac{2}{3}P(X=1)$. Calcula $P(X=0)$ y $P(X=3)$.

Solución

$$P(X=2)=\frac{2}{3}P(X=1)$$
$$e^{-\lambda}\frac{\lambda^2}{2}=\frac{2}{3}e^{-\lambda}\lambda$$
$$\lambda=\frac{4}{3}=1.33$$
$$P(X=0)=e^{-1.33}\frac{1.33^0}{0!}=0.264$$
$$P(X=3)=e^{-1.33}\frac{1.33^3}{6}=0.104$$

Problema 13._ Una fuente radiactiva se observa durante 7 intervalos cada uno de 10 segundos de duración y se cuenta el número de partículas emitidas durante cada periodo. Supón que el número de partículas emitidas, digamos *X*, durante cada periodo observado tiene una distribución de Poisson con parámetro $\lambda=5$. ¿Cuál es la probabilidad de que:
a) en cada uno de los 7 intervalos de tiempo, se emitan 4 o más partículas?
b) Al menos en uno de los 7 intervalos de tiempo se emitan 4 o más partículas?

Solución

X: Número de partículas emitidas durante un período.
$X \sim P(x;\lambda=5)$

a) La probabilidad de que en un período se emitan 4 o más partículas es:

$$P(X\geq 4)=1-F_X(3)=1-0.265=0.735$$

Como hay independencia en los 7 períodos se tiene que:

$$P(X\geq 4 \text{ en I1 y } X\geq 4 \text{ en I2 y}\cdots\text{y } X\geq 4 \text{ en I7})=(0.735)^7=0.1158$$

b) Y: Número de intervalos en que se emiten 4 o más partículas.

Y~bin(y; $n=7$, $p=0.735$)

$$P(Y\geq 1)=1-P(Y=0)=1-\binom{7}{0}(0.735^0)(0.265)^7=1\text{-}0.265^7=0.999.$$

Problema 14._ El número de partículas emitidas por una fuente radioactiva durante un período específico es una v.a con distribución de Poisson. Si la probabilidad de ninguna emisión es igual a 1/3. ¿Cuál es la probabilidad de que ocurran 2 o más emisiones?

<u>Solución</u>

X : Número de partículas emitidas por una fuente radioactiva en un período específico.
$X \sim P(x;\lambda)$

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

La probabilidad de ninguna emisión es $P(X = 0) = \dfrac{e^{-\lambda}\lambda^{0}}{0!} = e^{-\lambda} = \dfrac{1}{3}$

Sacando logaritmo natural en ambos lados se tiene:

$$\ln\left(e^{-\lambda}\right) = \ln\left(\frac{1}{3}\right)$$

$$-\lambda = \ln(1) - \ln(3) \quad \therefore \lambda = \ln(3)$$

Una vez encontrado el valor del parámetro de la distribución se saca la probabilidad de 2 o más emisiones usando el complemento del evento de interés. Como ya se conoce el valor de la probabilidad de cero emisiones, ahora se calcula la probabilidad de una emisión.

$$P(X = 1) = \frac{e^{-\ln(3)}(\ln(3))^{1}}{1} = 3^{-1}\ln(3) = \frac{1}{3}\ln(3)$$

Finalmente la probabilidad de 2 o más emisiones es:

$$\begin{aligned} P(X \geq 2) &= 1 - P(X \leq 1) \\ &= 1 - \left[P(X = 0) + P(X = 1)\right] = 1 - \left[\frac{1}{3} - \frac{1}{3}\ln(3)\right] = \frac{2}{3} - \frac{1}{3}\ln(3) = \frac{2 - \ln(3)}{3} \end{aligned}$$

Problema 15._ El dueño de una tienda tiene existencias de cierto artículo y decide utilizar la siguiente promoción para disminuir la existencia. El artículo tiene un precio de $100. El dueño reducirá el precio a la mitad por cada cliente que compre el artículo durante un día en particular. Así el primer cliente pagará $50 por el artículo, el segundo pagará $25, y así sucesivamente. Supón que el número de clientes que compra el artículo durante el día tiene una distribución de Poisson con media 2. Encuentra el costo esperado del artículo al final de día.

<u>Solución</u>

Y : Número de clientes que compra el artículo durante el día.
$Y \sim P(y;\lambda = 2)$

La función de costo del artículo según la promoción del dueño está dada por:

$$C(Y) = 100\left(\frac{1}{2}\right)^{y}$$

El valor esperado de la función costo se saca usando el hecho de que:

$$E(g(Y) = \sum_{y=0}^{\infty} g(y)P(Y = y)$$

$$E[C(Y)] = \sum_{y=0}^{\infty} 100\left(\frac{1}{2}\right)^{y} e^{-\lambda} \frac{\lambda^{y}}{y!}$$

$$= \sum_{y=0}^{\infty} 100\left(\frac{1}{2}\right)^{y} e^{-2} \frac{2^{y}}{y!}$$

$$= 100\,e^{-2} \sum_{y=0}^{\infty} \frac{\left(\frac{1}{2}2\right)^{y}}{y!}$$

$$= 100\,e^{-2} \sum_{y=0}^{\infty} \frac{1^{y}}{y!}$$

$$= 100\,e^{-2} e = 100\,e^{-1}$$

Recuerda que : $e^{x} = \sum_{x=0}^{\infty} \frac{x^{i}}{i!}$

Problema 16._ Un estacionamiento tiene dos entradas. Los coches llegan a la entrada 1 de acuerdo con una distribución de Poisson con una media de tres por hora, y a la entrada 2 de acuerdo con una distribución de Poisson con una media de 4 por hora. ¿Cuál es la probabilidad de que tres coches lleguen al estacionamiento durante una hora dada?

Solución

X_1 : Número de coches que llagan al estacionamiento por la entrada 1 en una hora.
$X_1 \sim P(x_1; \lambda_1 = 3)$

X_2 : Número de coches que llagan al estacionamiento por la entrada 2 en una hora.
$X_2 \sim P(x_2; \lambda_2 = 4)$

X : Número de coches que llagan al estacionamiento.
$X = X_1 + X_2$

La suma de variables aleatorias Poisson es una Poisson con parámetro igual a la suma de los parámetros individuales. Es decir:

$$X \sim G(x; \lambda = \lambda_1 + \lambda_2 = 7)$$

$$P(X = 3) = \frac{e^{-\lambda}\lambda^{x}}{x!} = \frac{e^{-7}\left(7^{3}\right)}{3!} = 0.0521$$

Problema 17._ Se sabe que el proceso de producción de luces de un tablero de automóvil de indicador giratorio produce uno por ciento de luces defectuosas. Si este valor permanece invariable, y se selecciona al azar una muestra de 100 luces, encuentre $P(\hat{p} \leq 0.03)$, donde $\hat{p}$ es la fracción de defectos de la muestra.

__Solución__

Sea la v.a *X*: Número de luces defectuosas.

P (luz defectuosa) = 0.01

La v.a tiene una distribución binomial de la forma $X \sim bin(x; n = 100, p = 0.01)$.

Dado que n es grande se hace una aproximación usando la distribución de Poisson con parámetro $\lambda = np = 100\ (0.01) = 1$.

Observa que la fracción de luces defectuosas en la muestra en términos de la v.a es:

$$\hat{p} = \frac{X}{n}$$

Etonces:

$$P(\hat{p} \leq 0.03) = P\left(\frac{X}{n} \leq 0.03\right) = P(X \leq n(0.03))$$

$$= P(X \leq 100(0.03)) = P(X \leq 3) = F_X(3)$$

Buscando en tablas de la distribución de Poisson con parámetro $\lambda = 1$ y acumulando en tres se tiene que $P(\hat{p} \leq 0.03) = 0.981$.

Si no se cuenta con las tablas de la distribución de Poisson se puede calcular directamente de la fórmula de la f.d.p Poisson:

$$P(X \leq 3) = \sum_{i=0}^{3} \frac{e^{-\lambda}\lambda^x}{x!}$$

$$= \sum_{i=0}^{3} \frac{e^{-1}1^x}{x!}$$

$$= e^{-1}\sum_{i=0}^{3} \frac{1}{x!}$$

$$= e^{-1}\left(1 + 1 + \frac{1}{2!} + \frac{1}{3!}\right)$$

$$= e^{-1}\left(\frac{6+6+3+1}{6}\right)$$

$$= e^{-1}\left(\frac{16}{6}\right) = e^{-1}\left(\frac{8}{3}\right) = 0.981$$

Problema 18._ La probabilidad de que un ratón inoculado con un suero contraiga cierta enfermedad es 0.2. Encuentra la probabilidad de que a lo más tres de 30 ratones inoculados contraigan la enfermedad, utilizando una aproximación de Poisson.

Solución

La variable aleatoria de interés es X: Número de ratones inoculados que contraen la enfermedad.

Su distribución es binomial con parámetros 30 y 0.2, denotada como $X \sim bin(x; n = 30, p = 0.2)$.

Si se aproxima con una Poisson $\lambda = np = 30(0.2) = 0.6$.

$$P(X \le 3) = \sum_{x=0}^{3} e^{-\lambda} \frac{\lambda^x}{x!}$$

$$= e^{-\lambda}\left[\frac{\lambda^0}{0!} + \frac{\lambda^1}{1!} + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!}\right] = e^{-6}\left[\frac{6^0}{0!} + \frac{6^1}{1!} + \frac{6^2}{2!} + \frac{6^3}{3!}\right]$$

$$= e^{-6}[1 + 6 + 18 + 36] = 61e^{-6} = 0.151$$

Problema 19._ En promedio una persona en 1000 comete un error numérico al preparar su declaración de impuestos. Si se seleccionan 10,000 formas al azar y se examinan, ¿cuál es la probabilidad de que 6, 7 u 8 de las formas contengan un error?

Solución

X: Número de personas que cometen un error en su declaración de impuestos.
La probabilidad de que una persona cometa un error en su declaración es $\frac{1}{1000} = 0.001$, y la v.a tiene una distribución binomial con los siguientes parámetros: $X \sim bin(x; n = 10{,}000, p = 0.001)$.

Aproximando con una Poisson se debe hacer $\lambda = np = 10{,}000(0.001) = 10$.
La probabilidad de interés se puede encontrar usando las tablas Poisson de la siguiente manera:

$$P(6 \le X \le 8) = F_X(8) - F_X(5)$$
$$= 0.3328 - 0.0671 = 0.2657$$

O usando la fórmula de la f.d.p Poisson como se muestra a continuación:

$$P(6 \le X \le 8) = \sum_{x=6}^{8} \frac{e^{-\lambda}\lambda^x}{x!}$$

$$= e^{-10}\left(\frac{10^6}{6!} + \frac{10^7}{7!} + \frac{10^8}{8!}\right) = 0.2657$$

Problema 20._ El chef de un restaurante prepara una ensalada revuelta que contiene, en promedio, cinco vegetales. Encuentra la probabilidad de que la ensalada contenga más de 5 vegetales:
a) en un día dado.
b) en tres de los siguientes 4 días.
c) por primera vez en abril el día 5.

Solución

La variable aleatoria X: Número de vegetales en 1 ensalada, tiene una distribución de Poisson con parámetro cinco. $X \sim P(x; \lambda = 5)$

a) $P(X > 5) = 1 - F_X(5) = 1 - 0.615 = 0.385$.

b) Sea la v.a Y: Número de días en que la ensalada tuvo más de 5 vegetales. Esta v.a tiene una distribución binomial con 4 ensayos independientes y probabilidad de éxito igual a la probabilidad de que la ensalada tenga más de cinco vegetales un día cualquiera. $Y \sim bin(y; n = 4, p = 0.385)$

$$P(Y = 3) = \binom{4}{3} 0.385^3 (0.615) = 4(0.385)^3 (0.615) = 0.1404.$$

c) Dado que se cuentan los días hasta una ensalada con la característica de interés se necesita una nueva v.a con distribución geométrica.

Z: Número de días hasta una ensalada con más de 5 vegetales. $Z \sim G(z; p = 0.385)$

$$P(Z = 5) = q^4 p = 0.615^4 (0.385) = 0.055.$$

Problema 21._ Supón que un libro con n páginas contiene, en promedio λ erratas por página. ¿Cuál es la probabilidad de que al menos haya m páginas que contengan más de k erratas?

Solución

X: Número de erratas por página. $X \sim P(x; \lambda)$

La probabilidad de que una página contenga más de k erratas es:

$$P(X > k) = \sum_{x=k+1}^{\infty} \frac{e^{-\lambda} \lambda^x}{x!}$$

Sea ahora la v.a Y: Número de páginas de un total de n con más de k erratas.

$$P(\textit{éxito}) = P(\text{la página con más de } k \text{ erratas}) = \sum_{x=k+1}^{\infty} \frac{e^{-\lambda} \lambda^x}{x!} = p$$

Entonces $Y \sim b(y; n, p)$ y la probabilidad de que al menos haya m páginas con el éxito está dada por:

$$P(Y \geq m) = \sum_{y=m}^{n} \binom{n}{y} \left[\sum_{x=k+1}^{\infty} \frac{e^{-\lambda} \lambda^x}{x!} \right]^y \left[\sum_{x=0}^{k} \frac{e^{-\lambda} \lambda^x}{x!} \right]^{n-y}$$

Problema 22._ El número promedio de ratas de campo por acre en un campo de cinco acres de trigo se estima en 12. Encuentre la probabilidad de que se encuentren menos de siete ratas de campo
a) En un acre dado.
b) En dos de los siguientes tres acres que se inspeccionan.

Solución

a) X : Número de ratas por acre. $X \sim P(x; \lambda = 12)$

$$P(X < 7) = P(X \leq 6) = F_X(6) = 0.0458$$

b) Sea la variable aleatoria Y **:** Número de acres con menos de 7 ratas.

La probabilidad de éxito es la probabilidad que se acaba de encontrar, p = 0.0458.

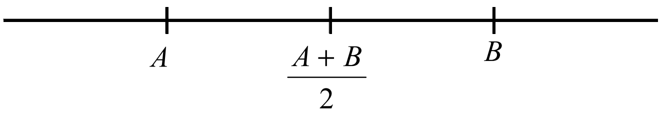
$$Y \sim bin(y;\, p = 0.0458,\ n = 3)$$

$$P(Y = 2) = \binom{3}{2} 0.0458^2 (1 - 0.0458) = \frac{3!}{2!}(0.0458)^2(0.9542) = 6.00 \times 10^{-3} = 0.006$$

II) CASO CONTINUO.

Problema 1._ Si un paracaidista cae en un sitio aleatorio de la línea recta entre los marcadores A y B
a) Encuentra la probabilidad de que esté más cerca de A que de B.
b) Calcula la probabilidad de que la distancia con respecto a A sea más de tres veces la distancia con respecto a B.
c) Si tres paracaidistas actúan independientemente, ¿cuál es la probabilidad de que exactamente uno de los tres caiga después del punto medio entre A y B?

Solución



Región de aterrizaje del paracaidista.

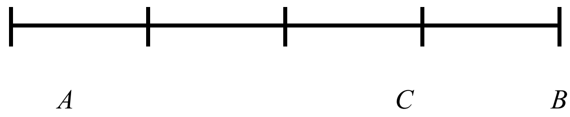
X: Punto en que cae el paracaidista.
La variable aleatoria se distribuye de manera uniforme de A a B. lo cual se escribe cmo:
$X \sim U(A, B)$

La f.d.p de esta v.a es $f_X(x) = \dfrac{1}{B - A}$

a) El paracaidista estará más cerca de *A* que de *B* si la distancia en la que cae es menor que el punto medio entre los dos extremos, es decir si:

$$P\left(X < \frac{A+B}{2}\right) = \int_A^{\frac{A+B}{2}} \frac{1}{B-A}dx = \frac{1}{B-A} x\Big|_A^{\frac{A+B}{2}} = \frac{1}{B-A}\left(\frac{A+B}{2} - A\right) = \frac{1}{B-A}\left(\frac{B-A}{2}\right) = \frac{1}{2} = 0.5$$

b) Si el paracaidista cae exactamente en el punto *C* la distancia de *A* a *C* es exactamente tres veces la distancia de *C* a *B*, así que para que cumpla la condición pedida tendría que caer en cualquier punto entre *C* y *B*.



Región de aterrizaje del paracaidista.

$$P\left(X > A + \frac{3}{4}(B-A)\right) = \int_{\frac{1}{4}A+\frac{3}{4}B}^{B} \frac{1}{B-A}dx = \frac{x}{B-A}\Big|_{\frac{1}{4}A+\frac{3}{4}B}^{B} = \frac{B - \frac{1}{4}A - \frac{3}{4}B}{B-A} = \frac{\frac{1}{4}(B-A)}{B-A} = \frac{1}{4}$$

c) Se supone independencia entre los tres paracaidistas y se construye una nueva v.a con distribución binomial.

Y: Número de paracaidistas que caen después del punto medio.

$$Y \sim bin\left(y; n = 3, p = \frac{1}{2}\right)$$

$$P(Y=1) = \binom{3}{1}\left(\frac{1}{2}\right)^1\left(\frac{1}{2}\right)^2 = 3\left(\frac{1}{2}\right)^3 = \frac{3}{8}$$

Problema 2._ El tiempo de un viaje (ida y vuelta) de los camiones que transportan concreto hacia una obra de construcción en una carretera, está distribuido uniformemente en un intervalo de 50 a 70 minutos. ¿Cuál es la probabilidad de que la duración del viaje sea mayor a 65 minutos si se sabe que la duración del viaje es mayor a 55 minutos?

<u>Solución</u>

X: Tiempo de un viaje (ida y vuelta). $X \sim U(50,70)$

Recuerda que si $X \sim U(a,b)$ su función de probabilidad acumulada está dada por $F_X(x) = \frac{x-a}{b-a}$, así que $P(X > x) = 1 - F_X(x) = 1 - \frac{x-a}{b-a} = \frac{b-a-x+a}{b-a} = \frac{b-x}{b-a}$.

Entonces: $P(X > 65 \mid X > 55) = \frac{P(X > 55, X > 65)}{P(X > 55)} = \frac{P(X > 65)}{P(X > 55)} = \frac{70-65}{70-55} = \frac{5}{15} = \frac{1}{3}$

Problema 3._ Un corredor de bienes raíces carga comisión fija de \$50 más el 6% a las ganancias de los propietarios. Si la ganancia se distribuye de modo uniforme entre \$0 y \$2000, obtén la distribución de probabilidad de las remuneraciones totales del corredor.

Solución

R: Remuneraciones totales del corredor.
X: Ganancia de los propietarios.

$X \sim U(0, 2000)$
La remuneración total del corredor puede expresarse mediante la fórmula: $R = 50 + 0.6X$

Para encontrar la función de probabilidad de esta variable aleatoria primero se encuentra la función de distribución acumulada y después se deriva ya que $F_X^{'}(x) = f_X(x)$.

$$F_R(r) = P(R \le r) = P(50 + 0.06X \le r) = P\left(X \le \frac{r-50}{0.06} \right) = \int_0^{\frac{r-50}{0.06}} \frac{1}{2000}dx = \frac{r-50}{2000(0.06)}$$

$$F_R(r) = \frac{r-50}{120} \quad \therefore \quad f_R(r) = F_R{'}(r) = \frac{1}{120}$$

Cuando $x = 0$ $R = 50$ y cuando $x = 2000$ $R = 170$.
Así que $R \sim U(50,170)$, las remuneraciones del corredor también se distribuyen uniformemente en el intervalo 50 a 170.

Problema 4._ Supón que cinco estudiantes van a realizar un examen independientemente unos de otros y que el número de minutos que cualquier estudiante necesita para terminar el examen tiene una distribución exponencial con media 80. Supón que el examen empieza a las nueve de la mañana, determina la probabilidad de que al menos uno de los estudiantes termine el examen antes de las diez menos veinte de la mañana.

Solución

Sea Z: Tiempo que tarda un estudiante en resolver el examen.

$$Z \sim \exp\left(z; \beta = \frac{1}{80} \right)$$

Utilizando la función de probabilidad acumulada para la distribución exponencial se tiene que

$$P(Z \le 40) = 1 - e^{-\beta\, x} = 1 - e^{-\frac{1}{2}} = 0.3935$$

Se necesita otra variable aleatoria que sea Y: Número de estudiantes que terminan el examen antes de los 40 minutos.

$$Y \sim bin(y; n = 5, p = 0.3935)$$

$$P(Y \ge 1) = 1 - P(Y = 0) = 1 - (0.6065)^{.5} = 0.9179$$

Problema 5._ El tiempo *X*, en segundos, que tarda un bibliotecario para localizar una ficha de un archivo de registros sobre libros prestados tiene una distribución exponencial con tiempo esperado de 20 segundos.
a) Calcula las probabilidades $P(X<30)$, $P(20<X)$, $P(20<X<30)$.
b) ¿Para qué valor de t es $P(X<t)=0.5$?

Solución

La v.a es X: Tiempo para que el bibliotecario localice una ficha.

$X \sim \exp\left(x; \beta = \frac{1}{20}\right)$

a) $P(X<30) = 1-e^{-\beta x} = 1-e^{-\frac{30}{20}} = 1-e^{-1.5} = 1-0.223 = 0.777$

$P(X>20) = e^{-\frac{20}{20}} = e^{-1} = 0.367$

$P(20<X<30) = F_X(30) - F_X(20) = 1-e^{-1.5} - \left(1-e^{-1}\right) = e^{-1} - e^{-1.5} = 0.367 - 0.223 = 0.144$

b) $P(X<t) = 0.5 = 1-e^{-\frac{t}{20}}$

Despejando se tiene que:

$$e^{-\frac{t}{20}} = 1-0.5 = 0.5$$

Aplicando logaritmo natural en la ecuación anterior se tiene que

$-\dfrac{t}{20} = \ln(0.5)$ $\quad\quad \therefore\ t = -20\ln(0.5) = 13.863$

Problema 6._ Un transistor tiene una distribución de tiempo de falla exponencial con tiempo medio de falla de 20,000 hrs. El transistor ha durado 20,000 hrs. En una aplicación particular. ¿Cuál es la probabilidad de que el transistor falle a las 30,000hrs?

Solución

La v.a X: Tiempo de falla del transistor. $X \sim \exp\left(x; \frac{1}{20{,}000}\right)$.

La probabilidad de interés es una probabilidad condicional

$$P(X<30{,}000 \mid X>20{,}000) = \frac{P(20{,}000<X<30{,}000)}{P(X>20{,}000)} = \frac{F_X(30{,}000) - F_X(20{,}000)}{1-F_X(20{,}000)}$$

$$= \frac{1-e^{-\frac{30000}{20000}} - \left(1-e^{-\frac{20000}{20000}}\right)}{e^{-1}} = \frac{e^{-1}-e^{-1.5}}{e^{-1}} = \frac{0.144}{0.367} = 0.3935$$

Recuerda que la acumulada para la distribución exponencial está dada por $F_X(x) = 1-e^{-\beta x}$.

Problema 7._ Se consideran dos procesos de manufactura. El costo por unidad para el proceso *I* es *C*, en tanto que para el proceso *II* es 3*C*. Los productores de ambos procesos tienen densidades de tiempo de falla exponenciales con tasas de 1/25 fallas por hora y 1/35 fallas por hora, respectivamente, para los procesos *I* y *II*. Si un producto falla antes de 15 hrs. debe reemplazarse a un costo de *Z* dólares. ¿Qué proceso recomiendas?

Solución

Primero se debe construir la función de costo para cada uno de los procesos, en función del tiempo de falla. Si C_I es el costo para el proceso I y C_{II} el costo para el proceso II, tenemos que:

Proceso *I*

$$C_I = \begin{cases} C & X > 15 \\ C + Z & X < 15 \end{cases}$$

$$\beta_I = \frac{1}{25}$$

Proceso *II*

$$C_{II} = \begin{cases} 3C & X > 15 \\ 3C + Z & X < 15 \end{cases}$$

$$\beta_{II} = \frac{1}{35}$$

El proceso a recomendar es aquel que tenga un menor costo esperado de producción, así que necesitamos encontrar el costo esperado para cada uno de los procesos. Observa que el costo es una variable aleatoria discreta con dos valores posibles así que la esperanza está dada por:

$$\begin{aligned} E(C_I) &= CP(X > 15) + (C + Z)P(X < 15) \\ &= Ce^{-\frac{15}{25}} + (C + Z)\left(1 - e^{-\frac{15}{25}}\right) \\ &= C + Z\left(1 - e^{-\frac{15}{25}}\right) = C + 0.451Z \end{aligned}$$

$$\begin{aligned} E(C_{II}) &= 3CP(X > 15) + (3C + Z)P(X < 15) \\ &= 3Ce^{-\frac{15}{35}} + (3C + Z)\left(1 - e^{-\frac{15}{35}}\right) = 3Ce^{-\frac{3}{7}} + 3C + Z - 3Ce^{-\frac{3}{7}} - Ze^{-\frac{3}{7}} = 3C + Z\left(1 - e^{-\frac{3}{7}}\right) = 3C + 0.348Z \end{aligned}$$

Si $E(C_I) < E(C_{II})$ se recomienda el proceso *I*. Es decir si:

$$C + 0.451Z < 3C + 0.348Z$$

$$(0.451 - 0.348)Z < 2C$$

$$0.103Z < 2C$$

Si $Z < 19.42$ se recomienda el proceso *I*.

Problema 8._ El tiempo Y que tarda en realizarse cierta tarea clave en la construcción de una casa es una v.a que tiene una distribución exponencial con una media de 10 hrs. El costo C para completar esa tarea está relacionado con el cuadrado del tiempo que tarda en completarse mediante la fórmula $C = 100 + 4Y + 3Y^2$ encuentra el valor esperado C.

<u>Solución</u>

Y: Tiempo que tarda la tarea. $Y \sim \exp\left(y; \frac{1}{10}\right)$

$E(Y) = 10hrs$

$C = 100 + 40Y + 3Y^2$

$E(C) = 100 + 40E(Y) + 3E(Y^2)$

$$= 100 + 40\frac{1}{\beta} + 3\frac{2}{\beta^2}$$

$$= 100 + 40(10) + 3(200) = 1100$$

Recuerda que $V(Y) = E(X^2) - \mu^2 \ \therefore \ E(X^2) = V(Y) + \mu^2$

Y que la varianza para la exponencial es $\frac{1}{\beta^2}$

Problema 9._ Si la v. a X tiene una distribución exponencial con parámetro β, encuentra la expresión para el k-ésimo momento al origen.

<u>Solución</u>

$X \sim \exp(x; \beta)$

La f.d.p de la variable aleatoria exponencial es $f_X(x) = \beta e^{-\beta x}$ y su f.g.m: $\Psi_X(t) = \frac{\beta}{\beta - t} = \beta(\beta - t)^{-1}$

El k-ésimo momento al origen es la k-ésima derivada de la f.g.m evaluada en cero.

$$\Psi_X{}'(t=0) = \beta(\beta - t)^{-2}\Big|_{t=0} = \frac{1}{\beta}$$

$$\Psi_X{}''(t=0) = 2\beta(\beta - t)^{-3}\Big|_{t=0} = \frac{2}{\beta^2}$$

$$\Psi_X{}'''(t=0) = 2(3)\beta(\beta - t)^{-4}\Big|_{t=0} = \frac{2(3)}{\beta^3}$$

$$\Psi_X{}^{(IV)}(t=0) = 2(3)(4)\beta(\beta - t)^{-5}\Big|_{t=0} = \frac{2(3)(4)}{\beta^4}$$

Procediendo de la misma manera se llega a que el k-ésimo momento al origen está dado por la expresión $\Psi_X{}^{(k)}(t=0) = \frac{k!}{\beta^k}$

Problema 10._ En cierta Ciudad el consumo diario de agua (en millones de litros) sigue aproximadamente una distribución gamma con parámetros $\alpha = 2$ y $\beta = \frac{1}{3}$. Si la capacidad diaria de dicha ciudad es 9 millones de litros de agua

a) ¿Cuál es la probabilidad de que en cualquier día dado el suministro de agua sea inadecuado?

b) Encuentra la media y la varianza del consumo diario de agua.

<u>Solución</u>

Sea la variable aleatoria X: Consumo diario de agua. Con función de probabilidad

$$X \sim Gamma\left(x;\ \alpha = 2, \beta = \frac{1}{3}\right)$$

a) El suministro de agua será inadecuado si el consumo de agua es mayor a los 9 millones de lts.

Recuerda que para la distribución Gamma se tiene que $P(X > x) = \sum_{k=0}^{\alpha-1} \frac{e^{-\beta x}(\beta x)^k}{k!}$, es decir la probabilidad de que la v.a sea mayor que el valor x es la distribución acumulada para una Poisson en $(\alpha - 1)$ con parámetro $\lambda = \beta x$. Entonces:

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

$$P(X > 9) = F_X^P(1) = 0.1991 \qquad \text{con} \qquad \lambda = \beta x = \frac{1}{3}(9) = 3$$

El valor se puede obtener de tablas o directamente de la fórmula de la f.d.p de la Poisson.

$$P(X > 9) = \sum_{x=0}^{1} P(X = x) = \sum_{x=0}^{1} e^{-\lambda} \frac{\lambda^x}{x!} = e^{-3}\left(\frac{3^0}{0!} + \frac{3^1}{1!}\right) = 4e^{-3} = 0.1991$$

b) $E(X) = \dfrac{\alpha}{\beta} = \dfrac{2}{\frac{1}{3}} = 6 \qquad V(X) = \dfrac{\alpha}{\beta^2} = \dfrac{2}{\left(\frac{1}{3}\right)^2} = 9(2) = 18$

Problema 11._ Se sabe que la duración de cierto tipo de transistor sigue una distribución gamma con media de 10 semanas y desviación estándar de $\sqrt{50}$ semanas. ¿Cuál es la probabilidad de que el transistor:

a) dure a lo más 50 semanas?

b) No sobreviva las primeras 10 semanas?

<u>Solución</u>

La variable aleatoria X duración del transistor tiene una distribución $X \sim Gamma\left(x;\ \alpha = 2, \beta = \frac{1}{5}\right)$.

Los valores de los parámetros se encuentran utilizando la información adicional del problema sobre la media y la desviación estándar.

$$E(X) = 10\,\text{semanas} = \frac{\alpha}{\beta}$$

$\sigma = \sqrt{50}$ semanas $\quad \therefore\ V(X) = 50\,\text{semanas}^2 = \dfrac{\alpha}{\beta^2}$

$$50 = \frac{\alpha}{\beta}\frac{1}{\beta} = 10\left(\frac{1}{\beta}\right) \qquad \therefore\ \beta = \frac{10}{50} = \frac{1}{5} \quad y \quad \alpha = \beta(10) = \frac{1}{5}(10) = 2$$

a) $P(X \le 50) = 1 - F_X^P(1)$ con $\lambda = \beta\, x = \frac{1}{5}(50) = 10$

$$= 1 - 0.0005 = 0.9995$$

b) $P(X \le 10) = 1 - F_X^P(1)$ con $\lambda = \frac{1}{5}(10) = 2$

$$= 1 - \sum_{x=0}^{1} e^{-\lambda}\frac{\lambda^x}{x!}$$

$$= 1 - \sum_{x=0}^{1} e^{-2}\frac{2^x}{x!} = 1 - e^{-2}(1+2)$$

$$= 1 - 3e^{-2} = 0.594$$

Problema 12._ Supón que cuando un transistor de cierto tipo se somete a una prueba acelerada de vida útil, la duración X (en semanas) tiene una distribución gamma con media de 24 semanas y desviación estándar de 12 semanas.

a) ¿Cuál es la probabilidad de que un transistor dure entre 12 y 24 semanas?

b) ¿Cuál es la probabilidad de que un transistor dure a lo sumo 24 semanas?

c) Si la prueba en realidad termina después de t semanas. ¿Qué valor de t es tal que sólo la mitad de 1% de todos los transistores estén todavía funcionando al terminar la prueba?

__Solución__

X: Duración en semanas de un transistor.

La función de probabilidad de la variable aleatoria es:

$$X \sim Gamma\left(x;\ \alpha = 4, \beta = \frac{1}{6}\right)$$

Para encontrar los valores de los parámetros utilizamos la información sobre la media y la desviación estándar de la variable aleatoria.

$$E(X) = \frac{\alpha}{\beta} = 24$$

La varianza de la v.a es el cuadrado de la desviación estándar entonces la varianza es 144, es decir;

$$V(X) = 144 = \frac{\alpha}{\beta^2} = \left(\frac{\alpha}{\beta}\right)\frac{1}{\beta} = E(X)\frac{1}{\beta} = \frac{24}{\beta} \quad \therefore\ \beta = \frac{24}{144} = \frac{1}{6}$$

Sustituyendo en la fórmula de la esperanza se encuentra el valor del parámetro α:

$$E(X) = 24 = \frac{\alpha}{\beta} = \alpha\left(\frac{1}{\beta}\right) = \alpha(6) \quad \therefore\ \alpha = 4$$

a) $P(12 < X < 24) = F_X(24) - F_X(12) = \left[1 - F_X^{P1}(3)\right] - \left[1 - F_X^{P2}(3)\right]$

El parámetro de la Poisson del primer corchete denotada como $P1$ es $\lambda = \beta x = \frac{1}{6}(24) = 4$ y el parámetro para la distribución Poisson del segundo corchete denotada como $P2$ es $\lambda = \frac{1}{6}(12) = 2$, entonces:

$$P(12 < X < 24) = F_X^{P2}(3) - F_X^{P1}(3) = 0.8571 - 0.4335 = 0.4236$$

b) $P(X < 24) = 1 - F_X^{P}(3) = 1 - 0.4335 = 0.5665 \qquad \lambda = \beta x = \frac{24}{6} = 4$

c) $P(X > t) = 0.005 = F_X^{P}(3) \quad$ con $\quad \lambda = t\beta = \frac{1}{6}t$

Se busca en las tablas de probabilidad acumulada para la Poisson una que se aproxime al valor 0.005 observando que:
$F_X(3) = 0.0049 \qquad$ con $\quad \lambda = 11 \quad \therefore t = 6\lambda = 66$

Problema 13._ Supón que la resistencia a la ruptura de una cuerda, en libras, se distribuye normalmente con media de 100 y varianza 16. Cada 100 pies de alambre para cuerda produce una utilidad de $25 si la resistencia a la ruptura es mayor de 95. Si la resistencia es menor o igual a 95, la cuerda puede utilizarse con un propósito diferente y se obtiene una utilidad de $10 por alambre. Encuentra la utilidad esperada por alambre.

<u>Solución</u>

X: Resistencia a la ruptura de la cuerda. Con $X \sim N(x;\ \mu = 100,\ \sigma^2 = 16)$

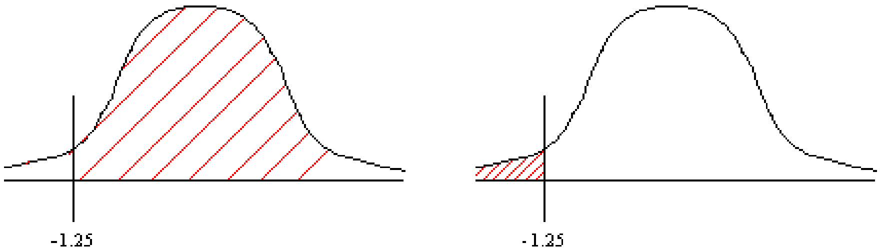
Primero hay que construir la función de utilidad:

$$U(X) = \begin{cases} 25 & X > 95 \\ \\ 10 & X \le 95 \end{cases}$$

Utilizando la definición del valor esperado de una función de una variable aleatoria tenemos que:

$$\begin{aligned} E(U(X)) &= \sum_u uP(U = u) = 25P(U = 25) + 10P(U = 10) \\ &= 25P(X > 95) + 10P(X \le 95) \\ &= 25P\left(Z > \frac{95-100}{4}\right) + 10P\left(Z \le \frac{95-100}{4}\right) \\ &= 25P(Z > -1.25) + 10P(Z \le -1.25) \end{aligned}$$

La parte sombreada de las siguentes graficas muestran las probabilidades de que $Z > -1.25$ y de que $Z < -1.25$ respectivamente:



Ubicación física de la región de interés.

Finalmente la utilidad esperada por alambre es:
$E(U(X)) = 25(0.8944)+10(0.1056)22.36+1.056 = 23.416$

Problema 14._ Se especifica que el diámetro exterior de un árbol de transmisión (flecha), llamémoslo D , debe ser de 4 pulgadas. Supón que D es una variable aleatoria distribuida normalmente con media de 4 pulgadas y varianza 0.01 pulgadas cuadradas. Si el diámetro real se diferencia del valor especificado por más de 0.05 pulgadas, pero en menos de 0.08 pulgadas, la pérdida del fabricante es de \$0.5, si el diámetro real se diferencia del diámetro especificado en más de 0.08 pulgadas, la pérdida es de \$1.00. La pérdida L, es una variable aleatoria, encuentra la distribución de probabilidad de L y calcula el valor esperado de la pérdida.

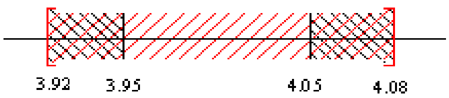
Solución

D: diámetro exterior del árbol de transmisión. $D \sim N(d;\ \mu = 4,\ \sigma^2 = 0.01)$

Primero se debe encontrar la función de pérdida en función del diámetro exterior del árbol de transmisión.

$$\text{Pérdida } L(D) = \begin{cases} \$0.5 & \text{si} \quad 0.05 < |D-4| < 0.08 \\ \\ \$1.00 & \text{si} \quad |D-4| > 0.08 \end{cases}$$

Para encontrar la función de probabilidad de la pérdida es necesario encontrar los valores posibles de la variable aleatoria y sus probabilidades correspondientes.
Para la condición $0.05 < |D-4| < 0.08$ se debe cumplir que $D-4 < -0.05$ o $D-4 > 0.05$ y $-0.08 < D-4 < 0.08$ es decir $D < 3.95$ o $D > 4.05$ y $3.92 < D < 4.08$ graficando estas desigualdades en una recta numérica se obtiene que la solución a la condición se da en el intervalo $3.92 < D < 3.95$ o en el intervalo $4.05 < D < 4.08$.



Representación gráfica de la condición $0.05 < |D-4| < 0.08$

La probabilidad de tener una pérdida de \$0.5 está dada por:

$$P(L=0.5)=P(0.05<|D-4|<0.08)=P(3.92<D<3.95)+P(4.05<D<4.08)$$

$$=P\left(\frac{3.92-4}{0.1}<Z<\frac{3.95-4}{0.1}\right)+P\left(\frac{4.05-4}{0.1}<Z<\frac{4.08-4}{0.1}\right)$$

$$=P(-0.8<Z<-0.5)+P(0.5<Z<0.8)$$

$$=2[\phi(0.8)-\phi(0.5)]$$

$$=2[0.7881-0.6915]=0.1932$$

La probabilidad de que la pérdida sea de \$1.00 está dada por:

$$P(L=1)=P(|D-4|>0.08)$$

$$=P(D-4<-0.08)+P(D-4>0.08)$$

$$=P(D<3.92)+P(D>4.08)$$

$$=P\left(Z<\frac{3.92-4}{0.1}\right)+P\left(D>\frac{4.08-4}{0.1}\right)$$

$$=P(Z<-0.8)+P(Z>0.8)$$

$$=2\phi(-0.8)$$

$$=2(0.2119)=0.4238$$

Observa que sólo hay pérdida en esos casos, entonces la pérdida será nula si $|D-4|<0.05$. Esto lo podemos resumir en cuadro siguiente:

Función de probabilidad de la pérdida del fabricante.

$L=\ell$	$P(L=\ell)$
0	0.383
0.5	0.1932
1.0	0.4238
suma	1.000

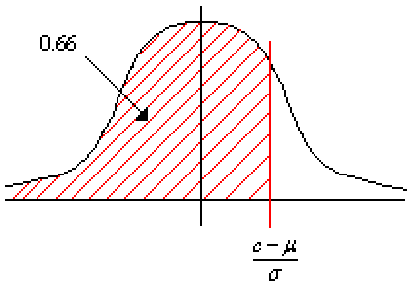
La pérdida esperada será entonces:

$$E(L)=\sum_{\ell}\ell P(L=\ell)=0(0.383)+0.5(0.1932)+1(0.4238)=\$0.5204$$

Problema 15._ Supón que *X* tiene una distribución Normal con media μ y varianza σ^2 . Determina *c*, como función de la media y la varianza, tal que $P(X \le c) = 2P(X > c)$.

Solución

$X \sim N(x;\ \mu,\ \sigma^2)$
$P(X \le c) = 2P(X > c) = 2[1 - P(X \le c)]$
$P(X \le c) + 2P(X \le c) = 2$
$P(X \le c) = \frac{2}{3}$
Estandarizando ambos lados de la desigualdad



$$P\left(\frac{X-\mu}{\sigma} \le \frac{c-\mu}{\sigma}\right) = \frac{2}{3}$$
$$P\left(Z \le \frac{c-\mu}{\sigma}\right) = 0.66$$

Buscando en las tablas de distribución Normal el valor en el cual se acumula 0.6666 de la distribución se obtiene que $\frac{c-\mu}{\sigma} = 0.43 \qquad \therefore\ c = \mu + 0.43\sigma$

Problema 16._ El diámetro de un cable eléctrico está distribuido normalmente con media 0.8 y varianza 0.0004. El cable se considera defectuoso si el diámetro se diferencia de su promedio en más de 0.025 ¿Cuál es la probabilidad de obtener un cable defectuoso?

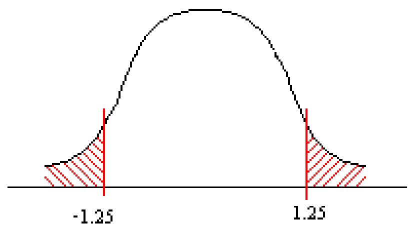
Solución

X: diámetro del cable.
$X \sim N(x;\ \mu = 0.8,\ \sigma^2 = 0.0004)$

Si $|X - \mu| > 0.025$ el cable se considera defectuoso.
Entonces:
$P(|X - \mu| > 0.025) = P(X - \mu < -0.025) + P(X - \mu > 0.025)$

Estandarizando se tiene que:

$$P(|X-\mu| > 0.025) = P\left(\frac{X-\mu}{\sigma} < \frac{-0.025}{\sigma}\right) + P\left(\frac{X-\mu}{\sigma} > \frac{0.025}{\sigma}\right)$$
$$= P\left(Z < \frac{-0.025}{0.02}\right) + P\left(Z > \frac{0.025}{0.02}\right)$$
$$= P(Z < -1.25) + P(Z > 1.25)$$
$$= 2\phi(-1.25) = 2(0.1056) = 0.2112$$



La probabilidad de que el cable se considere defectuoso es 0.2112.

Problema 17._ La distribución del peso de paquetes enviados de cierto modo es Normal con valor medio de 10 libras y desviación estándar de 2 libras. El servicio de paquetería desea establecer un valor de peso c, más allá del cual habrá cargo extra. ¿Cuál es el valor de c tal que 99% de todos los paquetes pesen por lo menos 1 libra abajo del peso con cargo extra?

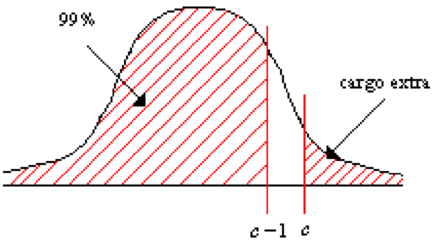
<u>Solución</u>

X: Peso del paquete.

$X \sim N(x; \mu = 10, \sigma^2 = 4)$

$P(X \leq c-1) = 0.99$

$P\left(Z \leq \frac{c-1-10}{2}\right) = P\left(Z \leq \frac{c-11}{2}\right) = 0.99$



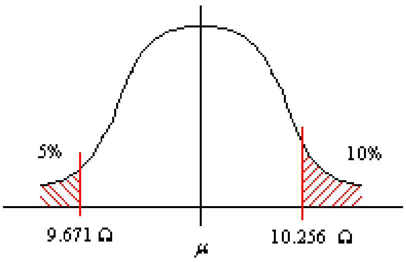
De tablas tenemos que $\frac{c-11}{2}$=2.33 $\quad \therefore c = 15.66$

Problema 18._ Se sabe que la distribución de resistencia para resistores de cierto tipo es normal, 10% de todos los resistores tiene una resistencia que excede $10.256\,\Omega$ y 5% tienen una resistencia menor de $9.671\,\Omega$. ¿Cuáles son el valor de la media y de la desviación estándar de la distribución de la resistencia?

<u>Solución</u>

X: Resistencia para resistores.

En la siguiente gráfica se ilustran las condiciones de las resistencias.



$P(X < 9.671) = 0.05$ y $P(X > 10.256) = 0.10$ Si normalizamos se tiene que:

$$P\left(Z < \frac{9.671-\mu}{\sigma}\right) = 0.05 \text{ y } P\left(Z > \frac{10.256-\mu}{\sigma}\right) = 0.10$$

Buscando en la tabla de distribución normal estándar se obtienen los valores de la v.a que cumplen las condiciones siguientes:

$$\frac{9.671-\mu}{\sigma} = -1.645 \quad \text{y} \quad \frac{10.256-\mu}{\sigma} = 1.28$$

De estas ecuaciones se forma el sistema de dos ecuaciones con dos incógnitas:

$$9.671 - \mu = -1.645\sigma \qquad (1)$$
$$10.256 - \mu = 1.28\sigma \qquad (2)$$

Restando (2) de (1)

$$-0.585 = -2.925\sigma \qquad \therefore \ \sigma = \frac{0.585}{2.925} = 0.2$$

Sustituyendo en la ecuación (1) se tiene que:

$$9.671 + 1.645(0.2) = \mu = 10.$$

Entonces la media de la distribución normal es 10 y la varianza es 0.04.

Problema 19._ Si una población de datos crudos posee una distribución normal, con media de 80 y una desviación estándar de 8, determina la media y la desviación estándar de la distribución muestral de la media para los siguientes tamaños de muestra; n = 16, n = 35 y n = 50. ¿Qué pasa cuando n crece?

<u>**Solución**</u>

La distribución de la población es:
$$X \sim N\left(x;\ \mu = 80,\ \sigma^2 = 8^2\right)$$

El Teorema Central del Límite (T.C.L) garantiza que la media muestral tiene una distribución aproximadamente normal, con media igual a la media de la población de donde salió y varianza igual a la varianza de la población de donde salió divida por el tamaño de la muestra. Lo cual se denota de la manera siguiente:

$$\overline{X} \approx N\left(\bar{x};\ \mu_{\bar{x}} = \mu,\ \sigma_{\bar{x}}^{\ 2} = \frac{\sigma^2}{n}\right)$$

Para encontrar la varianza de la media muestral tenemos que dividir 64 entre el tamaño de la muestra, para cada uno de los tres tamaños de muestra se tiene que:

Esperanza y varianza de la media muestral según el tamaño de la muestra.

n	$\mu_{\bar{x}}$	$\sigma_{\bar{x}}^2$	$\sigma_{\bar{x}}$
16	80	4	2
35	80	1.83	1.35
50	80	1.28	1.13

A medida que n crece la varianza decrece, lo cual se escribe: $\lim\limits_{n \to \infty} \sigma_{\bar{x}}^{\ 2} = 0.$

Problema 20._ Para el conjunto de datos poblacionales 3, 4, 5, 6, 7
a) Determina la distribución muestral de la media para muestras de tamaño 2. Supón que el muestreo es de uno a la vez, con reemplazo.

b) Demuestra que $\mu_{\bar{x}} = \mu$ y que $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$.

<u>Solución</u>

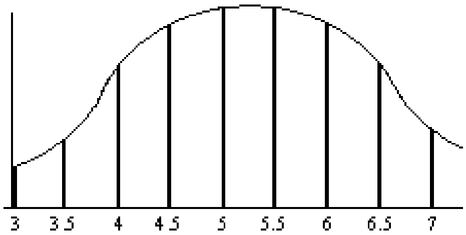
a) Para encontrar la distribución de la media muestral necesitamos encontrar los valores posibles de la media muestral y sus probabilidades asociadas. Para ello obtenemos todas las posibles muestras con reemplazo de tamaño dos que se pueden sacar de la población de interés {3, 4, 5, 6, 7} y encontramos la media muestral para cada una de estas muestras.

	Posible muestra	$\bar{x}$
1	(3,3)	3
2	(4,4)	4
3	(5,5)	5
4	(6,6)	6
5	(7,7)	7
6	(3,4)	3.5
7	(3,5)	4
8	(3,6)	4.5
9	(3,7)	5
10	(4,5)	4.5
11	(4,6)	5
12	(4,7)	5.5
13	(5,6)	5.5
14	(5,7)	6
15	(6,7)	6.5

Función de probabilidad de la media muestral.

$\bar{X} = \bar{x}$	$P(\bar{X} = \bar{x})$
3.0	1/15
3.5	1/15
4.0	2/15
4.5	2/15
5.0	3/15
5.5	2/15
6.0	2/15
6.5	1/15
7.0	1/15
Suma	15/15 = 1

Al graficar esta distribución observamos que tiene la forma acampanada de una distribución Normal.



Gráfica de la distribución de la media muestral.

b) La esperanza de la población está dada por:

$$\mu = \frac{1}{5}(3+4+5+6+7) = \frac{25}{5} = 5$$

La esperanza de la media muestral está dada por:

$$\mu_{\bar{x}} = 3\left(\frac{1}{15}\right) + 3.5\left(\frac{1}{15}\right) + 4\left(\frac{2}{15}\right) + 4.5\left(\frac{2}{15}\right) + 5\left(\frac{3}{15}\right) + 5.5\left(\frac{2}{15}\right) + 6\left(\frac{2}{15}\right) + 6.5\left(\frac{1}{15}\right) + 7\left(\frac{1}{15}\right) = 5$$

Se confirma el hecho de que $\mu_{\bar{x}} = \mu.$

La varianza de la población es:

$$\sigma^2 = \sum_x x^2 - \mu^2 = \frac{1}{5}\left(3^2+4^2+5^2+6^2+7^2\right) - 5^5 = \frac{135}{5} - 25 = 27 - 25 = 2$$

Para encontrar la varianza de la media muestral nos apoyamos en el cuadro siguiente:

$\bar{X} = \bar{x}$	$P(\bar{X} = \bar{x})$	$\bar{x}P(\bar{X} = \bar{x})$	$\bar{x}^2 P(\bar{X} = \bar{x})$
3	1/15	0.20	0.60
3.5	1/15	0.23	0.817
4	2/15	0.53	2.133
4.5	2/15	0.60	2.70
5	3/15	1.0	5.0
5.5	2/15	0.73	4.033
6	2/15	0.80	4.80
6.5	1/15	0.43	2.816
7	1/15	0.467	3.267
	Suma	5	26.167

La varianza de la media muestral es $V(\bar{X}) = E(\bar{X}^2) - E^2(\bar{X}) = 26.167 - 5^2 = 1.167$

Así que la desviación estándar de la media muestral es $\sigma_{\bar{x}} = \sqrt{V(\bar{X})} = \sqrt{1.16} = 1.08$

Y por otro lado $\frac{\sigma}{\sqrt{n}} = \frac{\sqrt{2}}{\sqrt{2}} = 1.00$ Por lo tanto se tiene que $\sigma_{\bar{X}} \approx \frac{\sigma}{\sqrt{n}}$.

Problema 21._ Un ensamble consta de 3 componentes colocados uno al lado de otro. La longitud de cada componente se distribuye normalmente con media de 2 pulgadas y desviación estándar 0.2 pulgadas. Las especificaciones requieren que todos los ensambles estén entre 5.7 y 6.3 pulgadas de longitud. ¿Cuántos ensambles cumplirán con estos requerimientos?

Solución

Sean las variables aleatorias:
X_i :longitud del $i-$ésimo componente. Con $i=1,\ 2,\ 3.\ \ X_i \sim N\left(x_i;\ \mu=2,\ \sigma^2=0.2^2\right)$

X: longitud del ensamble.

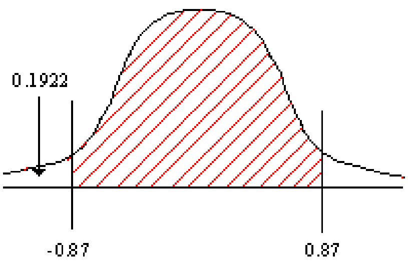
La longitud del ensamble se puede ver como la suma de las longitudes de los tres componentes, es decir: $X = X_1 + X_2 + X_3$

Sabemos que la suma de variables aleatorias normales se distribuye como una normal con media igual a la suma de las medias de las variables que se están sumando y varianza igual a la suma de las varianzas de las variables que se están sumando, entonces:

$$X \sim N\left(x;\ \mu=3(2)=6,\ \sigma^2=3\left(0.2^2\right)=0.12\right)$$

La proporción de ensambles que cumplirán los requerimientos es igual a la probabilidad de que un ensamble cumpla con los requerimientos.

$$P(5.7 \le X \le 6.3) = P\left(\frac{5.7-6}{\sqrt{0.12}} \le Z \le \frac{6.3-6}{\sqrt{0.12}}\right) = P(-0.86 \le Z \le 0.86) = 0.6102$$



$\therefore$ el 61.02% de los ensambles cumplen con los requerimientos.

Problema 22._ En el muestreo de un proceso de producción de artículos de los cuales 20% son defectuosos se selecciona una m.a de tamaño 100. ¿Cuál es la probabilidad de que a lo más 15 artículos sean defectuosos en la muestra?

Solución

X: Número de artículos defectuosos en la muestra.
$p = 0.2 \quad n = 100$

La variable aleatoria de interés se puede ver como la suma de 100 variables aleatorias Bernoulli con probabilidad de éxito 0.2 y usando la aproximación de la distribución Normal a la distribución binomial se sabe que:

$$X = \sum_{i=1}^{100} X_i \approx N\left(x; \mu_X = np = 20, \ \ \sigma_X^2 = npq = 16\right)$$

La probabilidad pedida sin corrección por finitud es:

$$P(X \le 15) = P\left(Z \le \frac{15-20}{4}\right) = P(Z \le -1.25) = 0.1056$$

Y la probabilidad con corrección por finitud es:

$$P(X \le 15) = P(X \le 15.5) = P\left(Z \le \frac{15.5-20}{4}\right) = P(Z \le -1.125) = 0.1313$$

Problema 23._ Un encuestador considera que el 20% de los votantes en cierta área está a favor de una emisión de valores bursátiles. Si se seleccionan 64 votantes al azar de un gran número de votantes en esta área, aproxima la probabilidad de que la fracción de votantes en la muestra a favor de la emisión de los valores no difiera en más de 0.06 de la fracción que él supone es la correcta.

Solución

$p = 0.2$ (proporción de votantes a favor de la emisión), $pq = 0.16$ y $n = 64$
X_i : Posición del l $i-$ésimo respecto a la emisión.
Si está a favor la variable vale 1 y si no vale cero.

Se trata de una población Bernoulli con media $\mu = p = 0.2$ y varianza $\sigma^2 = pq = 0.2(0.8) = 0.16$. Sea $\overline{X}$: la fracción de votantes en la muestra a favor de la emisión. Entonces por el Teorema Central del Límite se tiene que:

$$\overline{X} \approx N\left(\overline{x}; \mu_{\overline{x}} = \mu = 0.2, \ \ \sigma_{\overline{x}}^2 = \frac{\sigma^2}{n} = 0.0025\right)$$

La probabilidad pedida es:

$$P\left(\left|\overline{X} - 0.2\right| \le 0.06\right) = P(-0.06 \le \overline{X} - 0.2 \le 0.06) = P\left(\frac{-0.06}{0.05} \le \frac{\overline{X}-0.2}{0.05} \le \frac{0.06}{0.05}\right)$$

$$= P\left(-1.2 \le Z \le 1.2\right) = 0.7698$$

Problema 24._ Una línea aérea se da cuenta de que 5% de las personas que hacen sus reservaciones para cierto vuelo no se presentan. Si la aerolínea vende 160 boletos para un vuelo con solamente 155 asientos ¿Cuál es la probabilidad de que haya un asiento disponible para cada persona con reservación que se presenta para el vuelo?

Solución

P(La persona no llegue) = 0.05
P(La persona llegue) = 0.95

X_i : Condición del $i-$ésimo pasajero. Si la persona llega vale uno y si no llega vale cero.

Se trata de una población Bernoulli, siendo el éxito que la persona sí llegue al vuelo, entonces la media es $\mu = p = 0.95$ y la varianza $\sigma^2 = pq = 0.05(0.95) = 0.0475$.

La variable aleatoria X: Número de personas que sí llegan, se puede ver como la suma de 160 variables aleatorias Bernoulli $X = \sum_{i=1}^{160} X_i$.

Dado que $p = 0.95$, $q = 0.05$ y $n = 160$ se tiene que la v.a X tienen una distribución aproximadamente normal con parámetros 152 y 7.6 , lo cual se denota como:

$$X \sim N\left(x; \mu_X = n\mu = 152, \quad \sigma_X^2 = n\sigma^2 = 7.6\right)$$

Entonces la probabilidad de que haya un asiento disponible para cada persona con reservación que se presenta para el vuelo está dada por:

$$P(X \leq 155) = P\left(\frac{X - \mu_X}{\sigma_X} \leq \frac{155 - 152}{\sqrt{7.6}}\right) = P(Z \leq 1.088) = 0.8559$$

Problema 25._ Como un control de la abundancia relativa de cierta especie de pez en dos lagos, se hacen 50 observaciones con respecto a los resultados de la captura mediante trampas para cada lago. Para cada observación el experimento solamente anota si está o no la especie deseada. La experiencia previa ha mostrado que ésta especie aparecerá en las trampas del lago A aproximadamente 10% de las veces y en las trampas del lago B en aproximadamente 20% de las veces. Utiliza estos resultados para aproximar la probabilidad de que la diferencia entre las proporciones de las muestras difiera en a lo más 0.1 de la diferencia de las proporciones reales.

Solución

$p_A = 0.1$ proporción de peces en el lago A.
$p_B = 0.2$ proporción de peces en el lago B.
$\overline{X}_A$: proporción muestral de peces en el lago A.
$\overline{X}_B$: proporción muestral de peces en el lago B.
X : diferencia entre las proporciones muestrales.

Se tiene que la distribución de las proporciones es aproximadamente normal con los siguientes parámetros:

$$\overline{X}_A \approx N\left(\overline{x}_A; \mu_{\overline{x}_A} = 0.1,\ \sigma^2_{\overline{x}_A} = \frac{0.1(0.9)}{50}\right)$$
$$\overline{X}_B \approx N\left(\overline{x}_B; \mu_{\overline{x}_B} = 0.2,\ \sigma^2_{\overline{x}_B} = \frac{0.2(0.8)}{50}\right)$$

Suma de distribuciones normales también es normal con media igual a la suma de las medias de las variables que se están sumando y varianza igual a la suma de las varianzas de las variables aleatorias que se están sumando.

$$X = \overline{X}_A - \overline{X}_B$$

$$X \sim N\left(x; \mu = 0.1 - 0.2 = -0.1,\ \sigma^2 = \frac{0.1(0.9) + 0.2(0.8)}{50} = 0.005\right)$$

$$P(|X - (-0.1)| \le 0.1) = P(-0.1 \le X - (-0.1) \le 0.1)$$
$$= P\left(\frac{-0.1}{0.07071} \le \frac{X - (-0.1)}{0.07071} \le \frac{0.1}{0.07071}\right)$$
$$= P(-1.4142 \le Z \le 1.4142) = 0.8414$$

Problema 26._ La gerente de un taller de reparaciones no conoce la distribución de probabilidad del tiempo que se requiere para completar un trabajo. Sin embargo, de acuerdo con el desempeño pasado, ella ha podido estimar la media y la varianza como 14 días y 2 (días)^2, respectivamente. Encuentra un intervalo en el que la probabilidad de que un trabajo se termine en ese tiempo sea de 0.75.

Solución

X: tiempo para completar el trabajo.
$X \sim f_X(x; \mu = 14,\ \sigma^2 = 2)$

Como no se conoce la distribución de la variable aleatoria se utiliza el Teorema de Chebyshev, que establece que: $P(\mu - k\sigma < X < \mu + k\sigma) \ge 1 - \frac{1}{k^2}$

$$1 - \frac{1}{k^2} = 0.75 \ \therefore\ k^2 = 4 \ \text{ y } \ k = 2$$

El intervalo pedido está dado por la expresión $\mu \pm k\sigma = 14 \pm 2\sqrt{2}$. El intervalo es [11, 17] días.

Problema 27._ El servicio postal requiere, en promedio, 2 días para entregar una carta en una ciudad. La varianza se estima como 0.4 (días)^2. Si un ejecutivo desea que el 99% de sus cartas se entreguen a tiempo, ¿Con cuánta anticipación las debe depositar en el correo?

Solución

X : tiempo de entrega de la carta.
$\mu = 2$ días.
$\sigma^2 = 0.4$ días^2

Dado que no se conoce la distribución de la variable aleatoria se utiliza el Teorema de Chebyshev:

$$P(\mu - k\sigma < X < \mu + k\sigma) \geq 1 - \frac{1}{k^2}$$

$$1 - \frac{1}{k^2} = 0.99 \quad \therefore \quad k = 10$$

$$\mu \mp k\sigma = 2 \mp 10\sqrt{0.4} = [-4.3, 8.3]$$

Para garantizar que las cartas lleguen a tiempo debe mandarlas con 8 días de anticipación.

UNIDAD IV Distribución de varias variables aleatorias.

I) CASO DISCRETO.

Problema 1._ Sea X el número de caras y Y el número de caras menos el número de cruces cuando se lanzan 3 monedas. Encuentra la distribución de probabilidad conjunta de X y Y.

Solución

Para resolver este problema podemos hacer un cuadro indicando todos los resultados posibles del experimento e identificando el valor de las variables aleatorias X y Y.

Posible resultado.	X No. de caras	Y No. Caras - No. De cruces
ccc	3	3
ccx	2	1
cxc	2	1
xcc	2	1
xxc	1	-1
xcx	1	-1
cxx	1	-1
xxx	0	-3

La variable aleatoria X tiene una distribución binomial con parámetros 3 y 0.5, es decir:

$$X \sim bin\left(x; n=3, p=\frac{1}{2}\right)$$

$$P(X=0)=\binom{3}{0}p^0q^3=\frac{1}{8}$$

$$P(X=1)=\binom{3}{1}p^1q^2=\frac{3}{8}$$

$$P(X=2)=\binom{3}{2}p^1q^2=\frac{3}{8}$$

$$P(X=3)=\binom{3}{3}p^3q^0=\frac{1}{8}$$

La función de distribución de probabilidad conjunta o bidimensional, se expresa mediante el cuadro siguiente:

Función de probabilidad conjunta de (X,Y).

X	Y			
	-3	-1	1	3
0	1/8	0	0	0
1	0	3/8	0	0
2	0	0	3/8	0
3	0	0	0	1/8

Problema 2._ Dos líneas de producción fabrican cierto tipo de artículos. Supón que la capacidad (en cualquier día dado) es de 5 artículos para la línea I y de 3 artículos para la línea II y que él número verdadero de artículos producidos por cada una de las líneas es una v. a.
Sea (X,Y) la representación de la v.a bidimensional que proporciona el número de artículos producidos por la línea I y por la línea II respectivamente.

Función de probabilidad conjunta $P(X,Y)$.

Y \ X	0	1	2	3	4	5
0	0.00	0.01	0.03	0.05	0.07	0.09
1	0.01	0.02	0.04	0.05	0.06	0.08
2	0.01	0.03	0.05	0.05	0.05	0.06
3	0.01	0.02	0.04	0.06	0.06	0.05

a) Encuentra la probabilidad de que la línea I produzca 2 artículos y la línea II produzca 3 artículos.
b) Encuentra la probabilidad de que la línea I produzca más artículos que la línea II.
c) ¿Cuál es la probabilidad de que la línea I produzca exactamente 4 artículos?
d) ¿Cuál es el número esperado de artículos producidos por la línea II?

Solución

Sean las variables unidimensionales:
X : Número de artículos producidos por la línea I.
Y : Número de artículos producidos por la línea II.

a) La probabilidad de interés se observa directamente de la tabla de la f.d.p conjunta, es el número que está en la intersección de la columna encabezada por 2 y la fila correspondiente al número 3.

$$P(X=2,Y=3)=0.04$$

b) $P(X>Y)=\sum_{x>y}\sum P(X=x,Y=y)$

$$=P(X=1,Y=0)+P(X=2,Y=0)+P(X=2,Y=1)+P(X=3,Y=0)+P(X=3,Y=1)$$

$$+P(X=3,Y=2)+P(X=4,Y=0)+P(X=4,Y=1)+P(X=4,Y=2)+P(X=4,Y=3)$$

$$+P(X=5,Y=0)+P(X=5,Y=1)+P(X=5,Y=2)+P(X=5,Y=3)$$

$$=0.01+(0.03+0.04)+(0.05)3+(0.07+0.06+0.05+0.06)+(0.09+0.08+0.06+0.05)=0.75$$

c) $P(X=4)=\sum_{y=0}^{3}P(X=4,Y=y)$

$$=P(X=4,Y=0)+P(X=4,Y=1)+P(X=4,Y=2)+P(X=4,Y=3)$$
$$=0.07+0.06+0.05+0.06=0.24$$

d) Para la esperanza de Y, primero se encuentra la f.d.p de la v.a Y y después se usa la definición de la esperanza para el caso discreto unidimensional.

Función de probabilidad del número de artículos producidos por la línea II.

$Y = y$	$P(Y = y)$
0	0.25
1	0.26
2	0.25
3	0.24

Finalmente la esperanza de Y está dada por $E(Y) = \sum_{y=0}^{3} yP(Y = y) = 0.26 + 2(0.25) + 3(0.24) = 1.48$

Problema 3._ La f.d.p conjunta de (X,Y) está dada por $f_{X,Y}(x,y) = \frac{1}{n^2}$ para $x = 1,2,...,n$ y $y = 1,2,\ldots,n.$

a) Verifica que la función satisface las condiciones necesarias para ser una f.d.p discreta.
b) Calcula las densidades marginales X y Y.
c) ¿Son independientes X y Y?

Solución

a) Para que la función sea una f.d.p conjunta debe cumplir dos propiedades; la primera es que sea una función no negativa lo cual se cumple ya que n^2 es un número positivo.
La segunda propiedad es que la suma sobre todos los valores posibles de las variables (X,Y) sea uno.

$$f_{X,Y}(x,y) = \sum_{y=1}^{n}\sum_{x=1}^{n}\frac{1}{n^2} = \sum_{y=1}^{n} n\left(\frac{1}{n^2}\right) = n(n)\left(\frac{1}{n^2}\right) = \frac{n^2}{n^2} = 1$$

b) La f.d.p marginal para X se obtiene sumando la f.d.p conjunta sobre todos los valores posibles de la v.a Y y la f.d.p marginal para Y se obtiene sumando la f.d.p conjunta sobre todos los valores posibles de la v.a X.

$$f_X(x) = \sum_{y=1}^{n}\frac{1}{n^2} = n\left(\frac{1}{n^2}\right) = \frac{1}{n} \quad \text{Para } x = 1,2,...,n.$$

$$f_Y(y) = \sum_{x=1}^{n}\frac{1}{n^2} = n\left(\frac{1}{n^2}\right) = \frac{1}{n} \quad \text{Para } y = 1,2,...,n.$$

c) X y Y sí son independientes cuando el producto de las marginales coincide con la f.d.p conjunta.

$$f_X(x)f_Y(y) = \left(\frac{1}{n}\right)\left(\frac{1}{n}\right) = \frac{1}{n^2} = f_{X,Y}(x,y)$$

Por lo tanto las variables aleatorias X y Y sí son independientes.

Problema 4._ Sea X el número de errores sintácticos, y Y el de errores lógicos, en la primera ejecución de un programa escrito en C^{++}.

Función de probabilidad conjunta de (X,Y).

X/Y	**0**	**1**	**2**	**3**
0	0.400	0.100	0.020	0.005
1	0.300	0.040	0.010	0.004
2	0.040	0.010	0.009	0.003
3	0.009	0.008	0.007	0.003
4	0.008	0.007	0.005	0.002
5	0.005	0.002	0.002	0.001

a) ¿Son independientes las variables X y Y?
b) Calcula $E(XY)$ y $E(X+Y)$.

Solución

a) Para averiguar la independencia de las variables aleatorias se debe verificar que para todos los valores de ambas variables se cumple que la probabilidad de la intersección es igual al producto de las probabilidades, es decir:

$$P(X=x,Y=y)=P(X=x)P(Y=y)$$

a)

Función de probabilidad conjunta de (X,Y).

X/Y	**0**	**1**	**2**	**3**	**Suma**
0	0.400	0.100	0.020	0.005	0.525
1	0.300	0.040	0.010	0.004	0.354
2	0.040	0.010	0.009	0.003	0.062
3	0.009	0.008	0.007	0.003	0.027
4	0.008	0.007	0.005	0.002	0.022
5	0.005	0.002	0.002	0.001	0.010
Suma	0.762	0.167	0.053	0.018	1.000

$P(X=0,Y=1)=0.10\neq 0.087675=(0.525)(0.167)=P(X=0)P(Y=1)$
Por lo tanto las variables X y Y no son independientes.

b) La $E(XY)=\sum_{y=0}^{3}\sum_{x=0}^{5}xyP(X=x,Y=y)$

$$\begin{aligned}&=(1)(1)(0.04)+(1)(2)(0.01)+(1)(3)(0.004)+\\&\quad(2)(1)(0.01)+(2)(2)(0.009)+(2)(3)(0.003)+\\&\quad(3)(1)(0.008)+(3)(2)(0.007)+(3)(3)(0.003)+\\&\quad(4)(1)(0.007)+(4)(2)(0.005)+(4)(3)(0.002)+\\&\quad(5)(1)(0.002)+(5)(2)(0.002)+(5)(3)(0.001)+\\&\therefore\ E(XY)=0.376\end{aligned}$$

La $E(X+Y)=\sum_{y=0}^{3}\sum_{x=0}^{5}(x+y)P(X=x,Y=y)$

$$\begin{aligned}
&=(0+0)(0.4)+(0+1)(0.1)+(0+2)(0.02)+(0+3)(0.005)+\\
&\quad(1+0)(0.3)+(1+1)(0.04)+(1+2)(0.01)+(1+3)(0.004)+\\
&\quad(2+0)(0.04)+(2+1)(0.01)+(2+2)(0.009)+(2+3)(0.003)+\\
&\quad(3+0)(0.009)+(3+1)(0.008)+(3+2)(0.007)+(3+3)(0.003)+\\
&\quad(4+0)(0.008)+(4+1)(0.007)+(4+2)(0.005)+(4+3)(0.002)+\\
&\quad(5+0)(0.005)+(5+1)(0.002)+(5+2)(0.002)+(5+3)(0.001)+
\end{aligned}$$

$E(X+Y)=1.024$

Problema 5._ Cuando un automóvil es detenido por una patrulla, se revisa el desgaste de cada neumático y cada faro delantero se verifica para ver si está correctamente alineado. Denotemos por *X* el número de faros delanteros que necesitan ajuste y por *Y* el numero de neumáticos defectuosos.

a) Si X y Y son independientes con $f_X(0)=0.5$, $f_X(1)=0.1$, $f_X(2)=0.4$ y $f_Y(0)=0.6$, $f_Y(1)=0.1$. $f_Y(2)=f_Y(3)=0.05$, $f_Y(4)=0.2$. Escribe la función de probabilidad conjunta de la v. a. bidimensional (*X, Y*) mediante una tabla.

b) Calcula $P(X\le 1, Y\le 1)$ y verifica que es igual al producto $P(X\le 1)P(Y\le 1)$.

c) ¿Cuál es la probabilidad de no violaciones $[P(X+Y=0)]$?

Solución

X: Número de faros delanteros que necesitan ajuste.
Y: Número de neumáticos defectuosos.

a) Primero escribimos la f.d.p marginal para cada una de las variables aleatorias X y Y.

Función de probabilidad de la v.a X.

$X=x$	$P(X=x)$
0	0.5
1	0.1
2	0.4

Función de probabilidad de la v.a Y.

$Y=y$	$P(Y=y)$
0	0.60
1	0.10
2	0.05
3	0.05
4	0.20

Para llenar cada una de las celdas de la f.d.p conjunta se usa el hecho de que hay independencia entre las variables, es decir $P(X=x,Y=y)=P(X=x)P(Y=y)$.

Función de probabilidad conjunta de la v.a bidimensional (X,Y).

$Y \backslash X$	**0**	**1**	**2**	$P(Y=y)$
0	0.3	0.06	0.24	0.6
1	0.05	0.01	0.04	0.1
2	0.025	0.005	0.02	0.05
3	0.025	0.005	0.02	0.05
4	0.1	0.02	0.08	0.2
$P(X=x)$	0.5	0.1	0.4	1.0

b) $P(X \leq 1, Y \leq 1) = \sum_{y=0}^{1}\sum_{x=0}^{1} P(X = x, Y = y) = 0.3 + 0.06 + 0.05 + 0.01 = 0.42$

Por otro lado tenemos que:

$$P(X \leq 1) = P(X = 0) + P(X = 1) = 0.5 + 0.1 = 0.6$$

$$P(Y \leq 1) = P(Y = 0) + P(Y = 1) = 0.6 + 0.1 = 0.7$$

Entonces $P(X \leq 1)P(Y \leq 1) = 0.6(0.7) = 0.42$

Por lo tanto $P(X \leq 1, Y \leq 1) = P(X \leq 1)P(Y \leq 1)$

c) La probabilidad de no violaciones es $P(X + Y = 0) = P(X = 0, Y = 0) = 0.3$

Problema 6._ Un maestro acaba de entregar un largo artículo a una mecanógrafa y otro, un poco más corto a otra. Sea *X* el número de errores de mecanografía del primer artículo y *Y* el número de errores de mecanografía del segundo artículo. Supón que X tiene una distribución de Poisson con parámetro λ y *Y* tiene una distribución de Poisson con parámetro μ y que *X* y *Y* son independientes.

a) ¿Cuál es la función de probabilidad conjunta para la v. a. bidimensional (*X, Y*)?

b) ¿Cuál es la probabilidad de que a lo sumo se cometa un error en ambos artículos combinados?

c) Obtén una expresión general para la probabilidad de que el número total de errores de los dos artículos sea m, donde m es un entero no negativo.

Sugerencia: $A = \{(x, y) \mid x + y = m\} = \{(m,0), (m-1,1), \ldots, (1, m-1), (0, m)\}$

Ahora suma la f.d.p conjunta sobre $(x, y) \in A$ y utiliza el Teorema del binomio que dice

$\sum_{k=0}^{m} \binom{m}{k} a^k b^{m-k} = (a+b)^m$ para cualquier (a, b).

__Solución__

X: Número de errores del artículo I. $X \sim P(x;\lambda)$

Y: Número de errores del artículo II. $Y \sim P(y;\mu)$ $X \perp Y$

a) $$f_{X,Y}(x, y) = f_X(x) f_Y(y) = e^{-\lambda}\frac{\lambda^x}{x!} e^{-\mu}\frac{\mu^y}{y!} = e^{-(\lambda+\mu)}\frac{\lambda^x \mu^y}{x!\,y!}$$

b) $$P(X + Y \leq 1) = P(X = 0, Y = 1) + P(X = 1, Y = 0) + P(X = 0, Y = 0)$$

$$= f_X(0) f_Y(1) + f_X(1) f_Y(0) + f_X(0) f_Y(0)$$

$$= e^{-\lambda}\frac{\lambda^0}{0!} e^{-\mu}\mu + e^{-\lambda}\lambda e^{-\mu} + e^{-\lambda}e^{-\mu} = \mu e^{-(\lambda+\mu)} + \lambda e^{-(\lambda+\mu)} + e^{-(\lambda+\mu)}$$

$$= (1 + \lambda + \mu)e^{-(\lambda+\mu)}$$

c) Para calcular la probabilidad de que el número total de errores de los dos artículos sea m es necesario conocer todas las combinaciones de valores para las variables X y Y, para tal propósito resulta de gran ayuda el cuadro siguiente.

Valores de las variables que suman m.

X	Y
m	0
m-1	1
m-2	2
$\vdots$	$\vdots$
1	m-1
0	m

$$P(X+Y=m)=f_X(m)f_Y(0)+f_X(m-1)f_Y(1)+f_X(m-2)f_Y(2)$$

$$+\cdots+f_X(1)f_Y(m-1)+f_X(0)f_Y(m)$$

$$=\sum_{i=0}^{m}=e^{-\lambda}\frac{\lambda^i}{i!}e^{-\mu}\frac{\mu^{m-i}}{(m-i)!}=e^{-(\lambda+\mu)}\sum_{i=0}^{m}\frac{\lambda^i\mu^{m-i}}{i!(m-i)!}=\frac{e^{-(\lambda+\mu)}}{m!}\sum_{i=0}^{m}\frac{m!}{i!(m-i)!}\lambda^i\mu^{m-i}$$

$$=\frac{e^{-(\lambda+\mu)}}{m!}\sum_{i=0}^{m}\binom{m}{i}\lambda^i\mu^{m-i}=\frac{e^{-(\lambda+\mu)}}{m!}(\lambda+\mu)^m=e^{-(\lambda+\mu)}\frac{(\lambda+\mu)^m}{m!}$$

Problema 7._ Supón que una máquina se usa para un trabajo especifico en la mañana y para otro diferente en la tarde. X es el número de veces que la máquina falla en la mañana y Y el número de veces que la máquina falla en la tarde. En la tabla se muestra la distribución de probabilidad conjunta. ¿Son independientes las variables aleatorias X y Y?

Función de probabilidad conjunta $P(X,Y)$.

Y \ X	0	1	2	Suma $P(Y=y)$
0	0.10	0.20	0.20	0.5
1	0.04	0.08	0.08	0.2
2	0.06	0.12	0.12	0.3
Suma $P(X=x)$	0.20	0.40	0.40	1.00

Solución

Para que sean independientes se debe cumplir que la conjunta sea el producto de las marginales, entonces:

$P(X=0,Y=0)=P(X=0)P(Y=0)$

$0.1=0.5(0.2)$

$P(X=0,Y=1)=P(X=0)P(Y=1)$

$0.04=0.2(0.2)$

$P(X=0,Y=2)=P(X=0)P(Y=2)$

$0.06=0.2(0.3)$

Continuando así por columnas se tiene que:

$0.2 = 0.4(0.5)$ $\quad$ $0.2 = 0.4(0.5)$
$0.08 = 0.4(0.2)$ $\quad$ $0.08 = 0.4(0.2)$
$0.12 = 0.4(0.3)$ $\quad$ $0.12 = 0.4(0.3)$

Por lo tanto X y Y sí son independientes.

Problema 8: La distribución conjunta para Y_1 , el número de contratos asignados a la empresa *A*, y Y_2 , el número de contratos asignados a la empresa *B*, está dado por las entradas de la tabla siguiente.

Función de probabilidad de la (Y_1, Y_2).

Y_2	Y_1			Suma
	0	1	2	
0	1/9	2/9	1/9	4/9
1	2/9	2/9	0	4/9
2	1/9	0	0	1/9
Suma	4/9	4/9	1/9	1

a) Demuestra que la marginal para Y_1 es $bin\left(y_1, n = 2, p = \frac{1}{3}\right)$.

b) Encuentra $E(Y_1 - Y_2)$.

<u>Solución</u>

a)

Función de probabilidad de Y_1.

$Y_1 = y_1$	$P(Y_1 = y_1)$
0	4/9
1	4/9
2	1/9

La binomial con parámetros *n* y *p* está dada por $P(Y_1 = y_1) = \binom{n}{y_1} p^{y_1} q^{n-y_1}$

De la tabla se tiene que $P(Y_1 = 0) = \frac{4}{9}$ y de la fórmula se tiene que

$\frac{4}{9} = \binom{n}{0} p^0 q^n$ $\quad$ con $\quad n = 2$

$\frac{4}{9} = q^2$ entonces $q = \sqrt{\frac{4}{9}} = \frac{2}{3}$ y por lo tanto $p = \frac{1}{3}$

Si $Y_1 \sim bin\left(y_1;\ n=2,\ p=\frac{1}{3} \right)$ tendremos que:

$$P(Y_1 = 1)=\binom{2}{1}\ pq = 2\left(\frac{1}{3}\right)\left(\frac{2}{3}\right) = \frac{4}{9}$$

$$P(Y_1 = 2)=\binom{2}{2}\ p^2 q^0 = p^2 = \left(\frac{1}{3}\right)^2 = \frac{1}{9}$$

b) $E(Y_1 - Y_2) = \sum_{y_2=0}^{2}\sum_{y_1=0}^{2}(y_1 - y_2)P(Y_1 = y_1,\ Y_2 = y_2)$

$$= (0-0)\frac{1}{9} + (1-0)\frac{2}{9} + (2-0)\frac{1}{9} + (0-1)\frac{2}{9} + (1-1)\frac{2}{9}$$

$$+ (2-1)0 + (0-2)\frac{1}{9} + (1-2)0 + (2-2)0$$

$$= \frac{2}{9} + \frac{2}{9} - \frac{2}{9} - \frac{2}{9} = 0$$

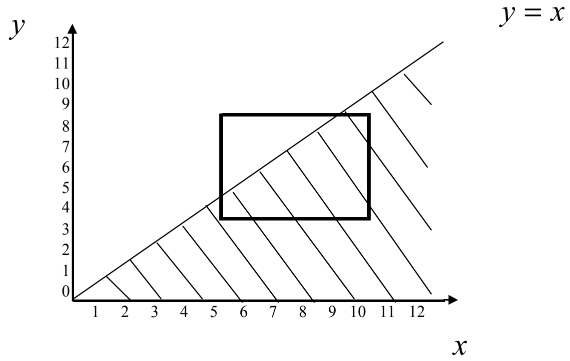
II) CASO CONTINUO.

Problema 1._ Supón que un fabricante de bombillas está interesado en el número de éstas que le han sido pedidas durante los meses de Enero y Febrero. X y Y indican el número de bombillas ordenadas durante esos dos meses, respectivamente. Supón que (X,Y) es una v. a. bidimensional con la siguiente f.d.p conjunta $f_{X,Y}(x,y) = C$, $5000 \le x \le 10{,}000$ y $4000 \le y \le 9000$.

a) Encuentra el valor de la constante *C*.

b) Calcula $P(X \ge Y)$.

<u>Solución</u>



a) Sabemos que:

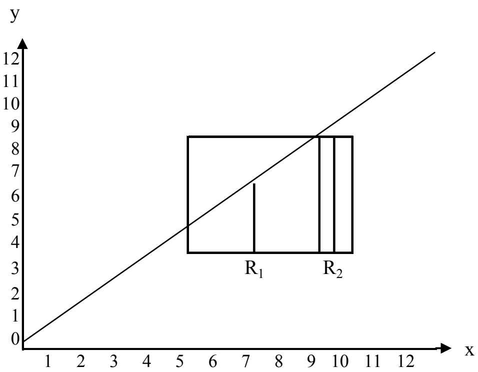
$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)dxdy = 1$$

$$\therefore \int_{4mil}^{9mil}\int_{5mil}^{10mil} Cdxdy = 1$$

$$C\left[x\Big|_{5mil}^{10mil} \quad y\Big|_{4mil}^{9mil} \right] = C\left[(5mil)(5mil)\right] = 1$$

$$\therefore C = \frac{1}{25{,}000{,}000}$$

b) Se tiene que integrar en la parte inferior de la recta $y = x$, para ello se divide la región de interés en dos partes R_1 y R_2 debido a que los límites de integración cambian en estas dos regiones. Tomando el elemento diferencial vertical se tiene que en la primera región y corre de 4,000 a la recta a 45 grados $y = x$ mientras que x corre libremente de 5,000 a 9,000. Dado que y queda en términos de x, primero se integra respecto a y y luego respecto a x. En la región 2 la variable x corre de 9,000 a 10,000 y la variable y toma valores de 4,000 a 9,000, como se indica en la figura.



$$P(X \geq Y) = c\left[\int_{x=5{,}000}^{9{,}000} \int_{y=4{,}000}^{x} dydx + \int_{y=4{,}000}^{9{,}000} \int_{x=9{,}000}^{10{,}000} dxdy \right]$$

$$= c\left[\int_{x=5{,}000}^{9{,}000} \left(y \Big|_{4{,}000}^{x} \right) dx + \left(x \Big|_{9{,}000}^{10{,}000} \right)\left(y \Big|_{4{,}000}^{9{,}000} \right) \right]$$

$$= c\left[\int_{x=5{,}000}^{9{,}000} (x - 4{,}000)dx + 5{,}000{,}000 \right]$$

$$= c\left[\left(\frac{x^2}{2} - 4{,}000x \right) \Bigg|_{5{,}000}^{9{,}000} + 5{,}000{,}000 \right]$$

$$= \frac{1}{25{,}000{,}000}\left[\frac{81{,}000{,}000 - 25{,}000{,}000}{2} - 16{,}000{,}000 + 5{,}000{,}000 \right]$$

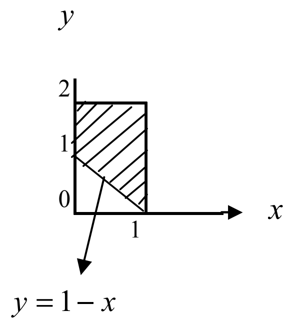
$$= \frac{17}{25} = 0.68$$

Problema 2._ Si la v. a. bidimensional continua (X,Y) tiene una f.d.p conjunta dada por

$$f_{X,Y}(x,y) = x^2 + \frac{xy}{3} \qquad 0 \le x \le 1,\ 0 \le y \le 2$$

a) Verifica que el volumen bajo la curva en la región definida es uno.
b) Encuentre $P(X+Y \ge 1)$.

<u>Solución</u>



a) $$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dxdy = \int_0^2\int_0^1 (x^2 + \frac{y}{3}x)\,dxdy$$

$$= \int_0^2 \left[\frac{x^3}{3} + \frac{yx^2}{6}\right]_0^1 dy$$

$$= \int_0^2 \left(\frac{1}{3} + \frac{1}{6}y\right) dy$$

$$= \left[\frac{1}{3}y + \frac{1}{12}y^2\right]\Bigg|_0^2$$

$$= \frac{2}{3} + \frac{1}{3} = 1$$

b) Si $x + y = 1 \;\therefore\; y = 1 - x$

La probabilidad pedida es la integral en la región sombreada.

$$P(X+Y \ge 1) = \int_{x=0}^{x=1}\left[\int_{y=1-x}^{2}\left(x^2 + \frac{y}{3}x\right)dy\right]dx$$

$$= \int_{x=0}^{x=1}\left(x^2 y + \frac{x}{3}\frac{y^2}{2}\right)\Bigg|_{y=1-x}^{2} dx$$

$$= \int_{x=0}^{x=1}\left(2x^2 - (1-x)x^2 + \frac{4}{6}x - \frac{x}{6}(1-x)^2\right)dx$$

$$= \int_{x=0}^{x=1} \left(2x^2 - x^2 + x^3 + \frac{2}{3}x - \frac{x}{6}\left[1 - 2x + x^2\right] \right) dx$$

$$= \int_{x=0}^{x=1} \left(2x^2 - x^2 + x^3 + \frac{2}{3}x - \frac{x}{6} + \frac{1}{3}x^2 - \frac{x^3}{6} \right) dx$$

$$= \int_{x=0}^{x=1} \left(\frac{5}{6}x^3 + \frac{4}{3}x^2 + \frac{1}{2}x \right) dx$$

$$= \left[\frac{5}{6}\frac{x^4}{4} + \frac{4}{3}\frac{x^3}{3} + \frac{1}{2}\frac{x^2}{2} \right]_0^1 = \left[\frac{5}{24}x^4 + \frac{4}{9}x^3 + \frac{1}{4}x^2 \right]_0^1$$

$$= \frac{45 + 96 + 54}{216} = \frac{195}{216} = \frac{65}{72}$$

Problema 3._ Y_1 y Y_2 tienen la función de densidad conjunta dada por
$f_{Y_1,Y_2}(y_1, y_2) = ky_1 y_2$ si $0 \le y_1 \le 1$, $0 \le y_2 \le 1$

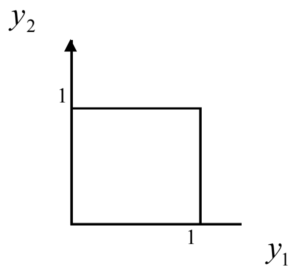
a) Determina el valor de k que la convierte en una función de densidad de probabilidad.

b) Encuentre la función $F_{Y_1,Y_2}(y_1, y_2)$.

c) Calcula $P\left(Y_1 \le \frac{3}{4}, Y_2 \ge \frac{1}{2}\right)$.

Solución

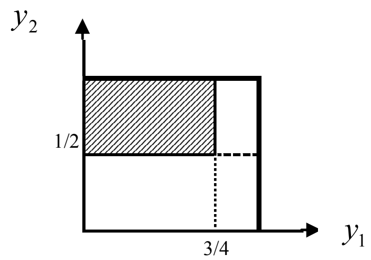
La f.d.p conjunta está definida en el cuadrado unitario siguiente:



a) $k\int_0^1\int_0^1 y_1 y_2 dy_1 dy_2 = k \left.\frac{y_1^{\,2}}{2}\right|_0^1 \left.\frac{y_2^{\,2}}{2}\right|_0^1 = \frac{k}{4} = 1 \quad \therefore \quad k = 4$

b) $F_{Y_1,Y_2}(y_1, y_2) = 4\int_0^{y_2}\int_0^{y_1} y_1 y_2 dy_1 dy_2 = 4\frac{y_1^{\,2}}{2}\frac{y_2^{\,2}}{2} = y_1^{\,2} y_2^{\,2}$ dentro del cuadrado $0 \le y_1 \le 1$, $0 \le y_2 \le 1$

c)



$$P\left(Y_1 \le \frac{3}{4}, Y_2 \ge \frac{1}{2}\right) = 4\int_0^{\frac{3}{4}} y_1 dy_1 \int_{\frac{1}{2}}^{1} y_2 dy_2$$

$$= 4\left.\frac{y_1^{\,2}}{2}\right|_0^{\frac{4}{3}} \left.\frac{y_2^{\,2}}{2}\right|_{\frac{1}{2}}^{1}$$

$$= \frac{4}{4}\left[\left. y_1^{\,2}\right|_0^{\frac{3}{4}} \left. y_2^{\,2}\right|_{\frac{1}{2}}^{1}\right]$$

$$= \left[\left(\frac{3}{4}\right)^2\left(1-\left(\frac{1}{2}\right)^2\right)\right]$$

$$= \left[\left(\frac{9}{16}\right)\left(\frac{3}{4}\right)\right] = \frac{27}{64} = 0.422$$

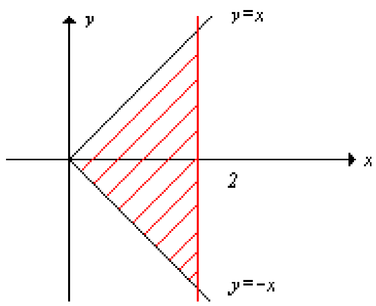
Problema 4._ Supón que la v. a. bidimensional (X, Y) tienen f.d. p conjunta dada por

$$f_{X,Y}(x,y) = kx\ \ (x-y) \qquad 0<x<2 \quad y \qquad -x<y<x$$

a) Encuentra el valor de k.
b) Encuentra las f. d. p marginales.

__Solución__

a) De las condiciones para x y y se tiene que la región en donde existe la f.d.p conjunta es:



Región donde existe la f.d.p conjunta.

La integral de la función en la región donde está definida debe ser la unidad, entonces:

$$1 = k\int_{x=0}^{x=2}\left[\int_{y=-x}^{y=x}(x^2 - xy)dy\right]dx$$

$$= k\int_{x=0}^{x=2}\left(x^2 y - x\frac{y^2}{2}\right)\Bigg|_{y=-x}^{y=x} dx$$

$$= k\int_0^2\left[x^3 - \frac{x^3}{2} - \left(-x^3 - \frac{x^3}{2}\right)\right]dx$$

$$= k\int_0^2 2x^3 dx = k2\frac{x^4}{4} = \frac{1}{2}kx^4\Big|_0^2 = 8k \quad \therefore \quad k = \frac{1}{8}$$

b) ***Marginal para*** X ._ Para encontrar la marginal de X se tiene que integrar la distribución conjunta respecto a la variable Y.

$$f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)dy = \frac{1}{8}\int_{y=-x}^{x}\left(x^2 - xy\right)dy = \frac{1}{8}\left[x^2 y - \frac{xy^2}{2}\right]_{-x}^{x}$$

$$= \frac{1}{8}\left[x^3 - \frac{1}{2}x^3 - \left(-x^3 - \frac{x^3}{2}\right)\right] = \frac{1}{8}\left[x^3 - \frac{1}{2}x^3 + x^3 + \frac{x^3}{2}\right] = \frac{1}{4}x^3$$

Entonces la función de densidad marginal para X está dada por:

$$f_X(x) = \frac{1}{4}x^3 \qquad \text{para} \quad 0 < x < 2$$

Marginal para Y ._ Para encontrar la marginal de Y se tiene que integrar la conjunta respecto a la variable X. Observa que para valores positivos de Y los valores de X corren de y a 2 y para valores negativos de Y los valores de X corren de $-y$ a 2. Por lo tanto la función de densidad marginal para Y es una función por partes.

i) Para $0 < y < 2$

$$f_Y(y) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)dx = \frac{1}{8}\int_{x=y}^{2}\left(x^2 - xy\right)dx = \frac{1}{8}\left[\frac{x^3}{3} - y\frac{x^2}{2}\right]\Bigg|_{x=y}^{2} = \frac{1}{8}\left[\frac{8}{3} - 2y - \left(\frac{y^3}{3} - \frac{y^3}{2}\right)\right]$$

$$= \frac{1}{8}\left[\frac{8}{3} - 2y + \frac{y^3}{6}\right] = \frac{1}{3} - \frac{1}{4}y + \frac{y^3}{48} \quad 0 < y < 2$$

ii) Para $-2 < y < 0$

$$f_Y(y) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)dx = \frac{1}{8}\int_{x=-y}^{2}\left(x^2 - xy\right)dx = \frac{1}{8}\left[\frac{x^3}{3} - y\frac{x^2}{2}\right]\Bigg|_{x=-y}^{2}$$

$$= \frac{1}{8}\left[\frac{8}{3} - 2y - \left(-\frac{y^3}{3} - \frac{y^3}{2}\right)\right] = \frac{1}{3} - \frac{y}{4} + \frac{5y^3}{48} \qquad \text{para} \qquad -2 \le y \le 0$$

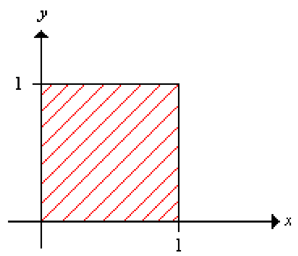
Finalmente la densidad marginal para Y se expresa como la siguiente función por partes:

$$f_Y(y) = \begin{cases} \dfrac{1}{3} - \dfrac{y}{4} + \dfrac{y^3}{48} & 0 < y < 2 \\ \\ \dfrac{1}{3} - \dfrac{y}{4} + \dfrac{5y^3}{48} & -2 \le y \le 0 \end{cases}$$

Problema 5._ Para que valor de k es $f_{X,Y}(x,y) = ke^{-(x+y)}$ una f. d. p conjunta de (*X, Y*) en la región $0 < x < 1$, $0 < y < 1$?

Solución

La región en donde está definida la función de densidad conjunta es el cuadrado siguiente:



Región donde existe la f.d.p conjunta.

Por propiedades de función de densidad conjunta la integral de la función en la región donde existe debe ser uno. Es decir:

$$1 = k\int_0^1\int_0^1 e^{-(x+y)}dxdy$$

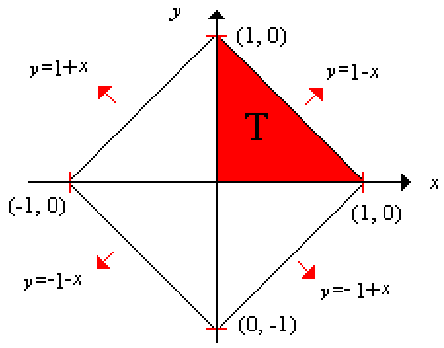
$$= k\int_0^1 e^{-x}dx\int_0^1 e^{-y}dy$$

$$= k\left(-e^{-x}\right)\Big|_0^1\left(-e^{-y}\right)\Big|_0^1 = k\left(1 - e^{-1}\right)^2 \quad \therefore \; k = \frac{1}{\left(1 - e^{-1}\right)^2}$$

Problema 6._ Supón que la v. a. bidimensional (*X, Y*) está distribuida uniformemente en el cuadrado cuyos vértices son (1,0), (0,1), (-1,0), (0,-1). Encuentra las densidades marginales $f_X(x)$ y $f_y(y)$.

Solución

La región donde existe la uniforme bidimensional es la siguiente:



Región donde existe la f.d.p conjunta.

Observa que el área de la región donde existe la conjunta es cuatro veces el área del triángulo (T). Primero se calcula el área en T y luego se multiplica por 4 para sacar el área de la región total.

$$\text{Area de T} = \int_{y=0}^{1}\int_{x=0}^{1-y} dxdy = \int_{y=0}^{1} x\Big|_{0}^{1-y} dy = \int_{y=0}^{1}(1-y)dy = \left(y-\frac{y^2}{2}\right)\Bigg|_0^1 = 1-\frac{1}{2} = \frac{1}{2}$$

$$\therefore \text{ Area de la región} = 4\left(\frac{1}{2}\right) = 2 \quad \therefore k = \frac{1}{2}$$

La distribución uniforme bidimensional es igual al inverso del área de la región en donde está definida así que la distribución que estamos buscando es:

$$f_{X,Y}(x,y) = \frac{1}{2}$$

Marginal para X._ Para encontrar la marginal de X se tiene que integrar la distribución conjunta respecto a la variable Y.

$$f_X(x) = \begin{cases} \displaystyle\int_{y=-1-x}^{y=1+x} \frac{1}{2}dy = \frac{1}{2}y\Big|_{-1-x}^{1+x} = 1+x & \text{si} \quad x \in [-1,0] \\ \displaystyle\int_{y=-1+x}^{y=1-x} \frac{1}{2}dy = \frac{1}{2}y\Big|_{-1+x}^{1-x} = 1-x & \text{si} \quad x \in [0,1] \end{cases}$$

Esta función también se puede escribir como $f_X(x) = 1 - |x| \quad \text{si} \quad x \in [-1, 1]$

Ya que $1 - |x| = \begin{cases} 1 - x & \text{si } x \in [0, 1] \\ 1 + x & \text{si } x \in [-1, 0] \end{cases}$

Marginal para Y ._ Para encontrar la marginal de Y se tiene que integrar la conjunta respecto a la variable X.

$$f_Y(y) = \begin{cases} \displaystyle\int_{x=y-1}^{x=1-y} \frac{1}{2}\,dx = \frac{1}{2}x \Big|_{y-1}^{1-y} = 1 - y & \text{si} \quad y \in [0, 1] \\ \displaystyle\int_{x=-1-y}^{x=y+1} \frac{1}{2}\,dx = \frac{1}{2}x \Big|_{-1-y}^{y+1} = 1 + y & \text{si} \quad y \in [-1, 0] \end{cases}$$

La densidad marginal para la v.a Y se puede escribir como $f_Y(y) = 1 - |y|$ para $-1 < y < 1$.

Problema 7._ Sean X y Y la duración de dos dispositivos electrónicos, con f. d. p conjunta dada por:

$$f_{X,Y}(x, y) = e^{-(x+y)} \qquad y \geq 0 \quad , \quad x \geq 0$$

¿Son X y Y variables aleatorias independientes?

<u>Solución</u>

Si la f.d.p conjunta se puede escribir como el producto de las marginales entonces sí son independientes, es decir si $f_{X,Y}(x, y) = f_X(x) f_Y(y)$ entonces $X \perp Y$.

Las f.d.p marginales son:

$$f_X(x) = \int_0^\infty e^{-x} e^{-y}\,dy = e^{-x}\left(-e^{-y}\right)\Big|_0^\infty = e^{-x}(1) = e^{-x}$$

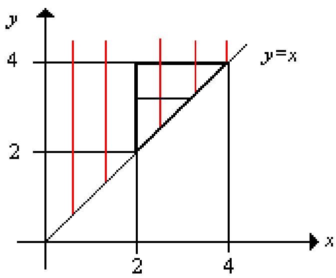
$$f_Y(y) = \int_0^\infty e^{-x} e^{-y}\,dx = e^{-y}\left(-e^{-x}\right)\Big|_0^\infty = e^{-y}$$

$f_{X,Y}(x, y) = e^{-(x+y)} = e^{-x}e^{-y} = e^{-x}\ e^{-y}$ entonces X y Y sí son independientes.

Problema 8._ Supón que la f. d. p conjunta de (*X*, *Y*) está dada por $f_{X,Y}(x,y)=e^{-y}$ para $x>0, y>x$.
Encuentra:
a) Las marginales para *X* y para *Y*.
b) $P(X>2 \,|\, y<4)$.

Solución

La región en donde existe la densidad conjunta es la parte que está por arriba de la recta a 45 grados en la gráfica siguiente:



a) La densidad marginal para X se obtiene integrando la conjunta respecto a Y y la marginal para Y se obtiene integrando la conjunta respecto a X.

$$f_X(x)=\int_{y=x}^{\infty} e^{-y}dy=-e^{-y}\Big|_x^{\infty}=e^{-x} \qquad x>0$$

$$f_Y(y)=\int_{x=0}^{y} e^{-y}dx=e^{-y}x\Big|_0^{y}=ye^{-y} \qquad y>x>0$$

b) Se trata de una probabilidad condicional así que usa la definición:

$$P(X>2|Y<4)=\frac{P(X>2,Y<4)}{P(Y<4)}$$

Para encontrar la probabilidad del numerador se tiene que integrar sobre el triángulo marcado con borde negrito de la figura de arriba.

$$P(X>2,>Y<4)=\int_{y=2}^{y=4}\left[\int_{x=2}^{x=y} e^{-y}dx\right]dy$$

$$=\int_{y=2}^{y=4} e^{-y}x\Big|_2^{y}dy$$

$$=\int_{y=2}^{y=4}(y-2)e^{-y}dy$$

$$=\int_{y=2}^{y=4} ye^{-y}dy-2\int_{y=2}^{y=4}e^{-y}dy$$

La primera integral se hace por partes tomando

$$u = y \qquad du = dy$$
$$dv = e^{-y}dy \qquad v = -e^{-y}$$

Se tiene que:

$$\int_{y=2}^{y=4} ye^{-y}dy = -ye^{-y}\Big|_{y=2}^{y=4} - \int_{y=2}^{y=4} -e^{-y}dy$$
$$= -4e^{-4} + 2e^{-2} - e^{-y}\Big|_{y=2}^{y=4}$$
$$= -4e^{-4} + 2e^{-2} - e^{-4} + e^{-2}$$
$$= -5e^{-4} + 3e^{-2}$$

Y la segunda integral es directa:

$$\int_{y=2}^{y=4} e^{-y}dy = -e^{-y}\Big|_{y=2}^{y=4} = -e^{-4} + e^{-2}$$

Así que $P(X > 2, Y < 4) = -5e^{-4} + 3e^{-2} - 2\left(-e^{-4} + e^{-2}\right) = -3e^{-4} + e^{-2}$

La integral del denominador es:

$$P(Y < 4) = \int_0^4 f_Y(y)dy = \int_0^4 ye^{-y}dy$$

Integrando por partes con : $u = y \qquad du = dy$

$$dv = e^{-y}dy \qquad v = -e^{-y}$$

$$P(Y < 4) = -ye^{-y}\Big|_0^4 + \int_0^4 e^{-y}dy$$
$$= \left[-ye^{-y} - e^{-y}\right]\Big|_0^4 = -4e^{-4} - e^{-4} + e^{-0}$$
$$= 1 - 5e^{-4}$$

$$\therefore P(X > 2 | Y < 4) = \frac{e^{-2} - 3e^{-4}}{1 - 5e^{-4}}$$

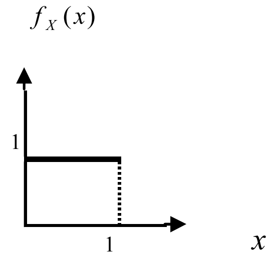
Problema 9._ El empuje X y la razón de la mezcla Y son dos características del funcionamiento de un motor a reacción. Si (X,Y) es una v. a. bidimensional con f. d. p $f_{X,Y}(x,y)=2(x+y-2xy)$ en $0\le x\le 1$, $0\le y\le 1$. Encuentra las distribuciones marginales para X y Y.

__**Solución**__

$$f_X(x)=\int_{y=0}^{1} f_{X,Y}(x,y)dy \quad \text{con} \quad 0\le x\le 1$$
$$=\int_0^1 2(x+y-2xy)dy$$
$$=2\left[xy+\frac{y^2}{2}-\frac{2xy^2}{2}\right]_0^1=2\left[x+\frac{1}{2}-x\right]=1$$

Entonces la f.d.p marginal para X es una distribución uniforme $X\sim U(0,1)$ dada por $f_X(x)=1$. Cuya gráfica es la siguiente:

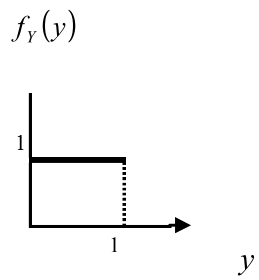
Gráfica de la v.a $X\sim U(0,1)$



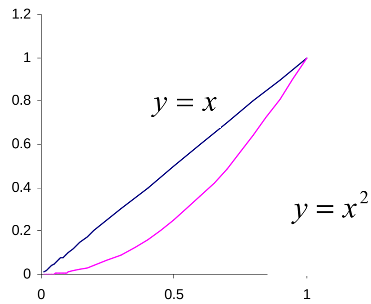
Para encontrar la marginal de Y se integra la conjunta respecto a X como sigue:

$$f_Y(y)=2\int_0^1 (x+y-2xy)dx \ \text{con} \ 0\le y\le 1$$
$$=2\left[\frac{x^2}{2}+xy-\frac{2x^2y}{2}\right]_0^1$$
$$=2\left[\frac{1}{2}+y-y\right]=1$$

Por lo tanto Y también tiene una f.d.p uniforme de 0 a 1, lo cual se escribe $Y\sim U(0,1)$ con $f_Y(y)=1$. Su representación gráfica es:



Problema 10._ Supón que la v. a. bidimensional (X,Y) está distribuida uniformemente en la región R comprendida entre la recta $y = x$ y la curva $y = x^2$ encuentra su f. d. p y las marginales.



Región en donde está definida la f.d.p conjunta.

Solución

Primero se encuentra el área de la región en donde está definida la uniforme bidimensional.

$$\text{Area de } R = \int_{x=0}^{1}\int_{y=x^2}^{y=x} dydx = \int_0^1 y\Big|_{y=x^2}^{y=x} dx$$

$$= \int_0^1 \left(x - x^2\right)dx = \left[\frac{x^2}{2} - \frac{x^3}{3}\right]_0^1 = \frac{1}{2} - \frac{1}{3} = \frac{1}{6}$$

Entonces la f.d.p conjunta está dada por $f_{X,Y}(x,y) = 6 \qquad (x,y) \in R$

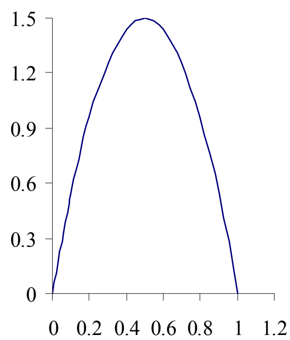
Cálculo de las marginales:

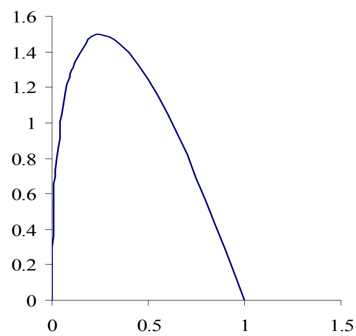
$$f_X(x) = \int_{y=x^2}^{x} 6dy = 6\left[y\Big|_{x^2}^{x}\right] = 6\left[x - x^2\right] \qquad 0 \le x \le 1$$

$$f_Y(y) = \int_{x=y}^{\sqrt{y}} 6dx = 6\left[x\Big|_{y}^{\sqrt{y}}\right] = 6\left[\sqrt{y} - y\right] \qquad 0 \le y \le 1$$

La representación gráfica de las f.d.p marginales para las variables X y Y respectivamente es:

$f_X(x)$ $\qquad$ $f_Y(y)$



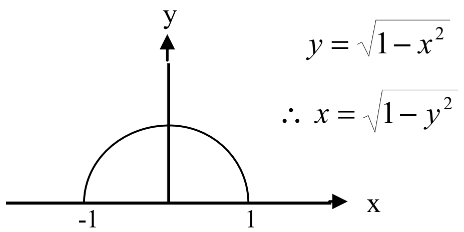


Problema 11._Supón que (X,Y) se distribuye de manera uniforme sobre el semicírculo del diagrama. De tal modo que $f_{X,Y}(x,y)=\dfrac{2}{\pi}$ si (x,y) está en el semicírculo. Encuentra:

a) Las distribuciones marginales de X y Y.

b) Las distribuciones de probabilidad condicional.

c) Las esperanzas condicionales.



Región donde existe la f.d.p conjunta.

Solución

a) $$f_X(x)=\int_{y=0}^{\sqrt{1-x^2}}\frac{2}{\pi}dy=\frac{2}{\pi}y\Big|_0^{\sqrt{1-x^2}}=\frac{2}{\pi}\sqrt{1-x^2}\qquad -1<x<1$$

$$f_Y(y)=2\left[\frac{2}{\pi}\int_{x=0}^{\sqrt{1-y^2}}dx\right]=\frac{4}{\pi}x\Big|_0^{\sqrt{1-y^2}}=\frac{4}{\pi}\sqrt{1-y^2}\qquad 0<y<1$$

b) $$f_{X|y}(x|y)=\frac{f_{X,Y}(x,y)}{f_Y(y)}=\frac{\frac{2}{\pi}}{\frac{4}{\pi}\sqrt{1-y^2}}=\frac{1}{2\sqrt{1-y^2}}\quad \text{para -1} \le x \le 1$$

$$f_{Y|x}(y|x)=\frac{f_{X,Y}(x,y)}{f_X(x)}=\frac{\frac{2}{\pi}}{\frac{2}{\pi}\sqrt{1-x^2}}=\frac{1}{\sqrt{1-x^2}}\quad \text{para } 0\le y\le 1$$

c) $$E(X|y)=\int_{-1}^{1}xf_{X|y}(x|y)dx=\int_{-1}^{1}x\frac{1}{2\sqrt{1-y^2}}dx=\frac{1}{2\sqrt{1-y^2}}\frac{x^2}{2}\Big|_{-1}^{1}=0$$

$$E(Y|x)=\int_0^1 yf_{Y|x}(y|x)dy=\int_0^1 y\frac{1}{\sqrt{1-x^2}}dy=\frac{1}{\sqrt{1-x^2}}\frac{y^2}{2}\Big|_0^1=\frac{1}{2\sqrt{1-x^2}}$$

Problema 12._ Supón que X_1 y X_2 son calificaciones codificadas en dos pruebas de inteligencia y la función de densidad de probabilidad conjunta está dada por

$$f_{X_1,X_2}(x_1,x_2) = 6x_1^2 x_2 \qquad 0 \le x_1 \le 1,\ \ 0 \le x_2 \le 1$$

a) Encuentra el valor esperado de la calificación de la prueba número 2 dada la calificación de la prueba número 1.
b) Además, obtén el valor esperado de la calificación de la prueba número 1 dada la calificación de la prueba número 2.

Solución

Primero se encuentran las marginales

$$f_{X_1}(x_1) = 6\int_0^1 x_1^2 x_2 dx_2 = 6x_1^2 \left.\frac{x_2^2}{2}\right|_0^1 = 3x_1^2$$

$$f_{X_2}(x_2) = 6\int_0^1 x_1^2 x_2 dx_1 = 6x_2 \left.\frac{x_1^3}{3}\right|_0^1 = 2x_2$$

Luego las f.d.p condicionales son:

$$f_{X_1|x_2} = \frac{f_{X_1,X_2}(x_1,x_2)}{f_{X_2}(x_2)} = \frac{6x_1^2 x_2}{2x_2} = 3x_1^2$$

$$f_{X_2|x_1} = \frac{f_{X_1,X_2}(x_1,x_2)}{f_{X_1}(x_1)} = \frac{6x_1^2 x_2}{3x_1^2} = 2x_2$$

Finalmente las esperanzas condicionales son:

$$a)\ \ E(X_1|x_2) = \int_0^1 x_1 3x_1^2 dx_1 = 3\int_0^1 x_1^3 dx_1 = \left.\frac{3x_1^4}{4}\right|_0^1 = \frac{3}{4}$$

$$b)\ \ E(X_2|x_1) = \int_0^1 x_2 f_{X_2|x_1}(x_2 \mid x_1) dx_2 = 2\int_0^1 x_2^2 dx_2 = 2\left.\frac{x_2^3}{3}\right|_0^1 = \frac{2}{3}$$

UNIDAD V Estadística paramétrica usando estimación y prueba de hipótesis.

I) ESTIMACIÓN PUNTUAL

Problema 1._ Si X es una v.a con media μ y varianza σ^2 y $X_1, X_2, \ldots, X_n$ es una m.a de tamaño n de X, se tiene que la media muestral $\bar{x}$ y la varianza muestral S^2 son estimadores puntuales de la media y la varianza poblacional respectivamente. ¿Son insesgados estos estimadores?

<u>Solución</u>

Recordemos que la media muestral está dada por la expresión $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} X_i$.

Entonces su esperanza es: $E(\bar{x}) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \mu \;\therefore\; \bar{x}$ sí estima de manera insesgada a la media poblacional.

La varianza muestral está dada por:

$$
\begin{aligned}
S^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 \\
&= \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i^2 - 2x_i\bar{x} + \bar{x}^2\right) \\
&= \frac{1}{n-1}\left[\sum_{i=1}^{n} x_i^2 - 2\bar{x}\sum_{i=1}^{n} x_i + \sum_{i=1}^{n}\bar{x}^2\right] \\
&= \frac{1}{n-1}\left[\sum_{i=1}^{n} x_i^2 - 2\bar{x}n\bar{x} + \sum_{i=1}^{n}\bar{x}^2\right] \\
&= \frac{1}{n-1}\left[\sum_{i=1}^{n} x_i^2 - 2n\bar{x}^2 + n\bar{x}^2\right] \\
&= \frac{1}{n-1}\left[\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right]
\end{aligned}
$$

Entonces: $$E(S^2) = \frac{1}{n-1}\left[\sum_{i=1}^{n} E(x_i^2) - nE(\bar{x}^2)\right]$$

Recuerda que $\bar{x} \sim N\left(\bar{x}; \mu_{\bar{x}} = \mu, \sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}\right)$ y que $V(X) = E(X^2) - (E(X))^2$ y por lo tanto el segundo momento de la v.a está dado por: $E(X^2) = V(X) + (E(X))^2$.

Utilizando toda esta información tenemos que:

$$E\left(S^2\right)=\frac{1}{n-1}\left[n\left(\sigma^2+\mu^2\right)-n\left(\frac{\sigma^2}{n}+\mu^2\right)\right]$$

$$=\frac{1}{n-1}\left(n\sigma^2+n\mu^2-\sigma^2-n\mu^2\right)$$

$$E(S^2)=\frac{\sigma^2(n-1)}{n-1}=\sigma^2$$

Entonces S^2 estima de manera insesgada a la varianza poblacional.

Problema 2._ Demuestra que la media muestral es el estimador de varianza mínima para la media poblacional de una distribución normal con media μ y varianza σ^2.

Solución

Para ver si $\bar{x}$ es el estimador de varianza mínima utilizaremos la Cota de Cramér-Rao que dice que:

$$V(\hat{\theta})=\frac{1}{nE\left[\frac{\partial}{\partial\theta}\ln f_X(x;\theta)\right]^2}$$

Ya vimos que la media muestral es un estimador insesgado de la media poblacional μ.

La f.d.p de la normal está dada por la expresión $f_X\left(x;\mu,\sigma^2\right)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$

$$\ln f_X\left(x;\mu,\sigma^2\right)=\ln\frac{1}{\sigma\sqrt{2\pi}}-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2$$

$$\frac{\partial}{\partial\mu}\ln f_X\left(x;\mu,\sigma^2\right)=\frac{1}{2\sigma^2}\left[2(x-\mu)\right]=\frac{x-\mu}{\sigma^2}$$

$$E\left[\frac{\partial}{\partial\mu}\ln f_X\left(x;\mu,\sigma^2\right)\right]^2=E\left[\frac{1}{\sigma^4}(x-\mu)^2\right]=\frac{1}{\sigma^4}E(x-\mu)^2=\frac{1}{\left(\sigma^2\right)^2}V(X)=\frac{1}{\sigma^2}$$

Entonces $CCR=\frac{\sigma^2}{n}=V\left(\bar{x}\right)$ por lo tanto $\bar{x}$ es el estimador insesgado de varianza mínima para μ.

Problema 3._ Sea $X \sim N\left(x;\mu,\sigma^2\right)$ estima $\underline{\theta}=\left(\mu,\sigma^2\right)$ en base a una muestra aleatoria de tamaño n usando el método de los momentos.

<u>Solución</u>

Tenemos dos parámetros a estimar por lo tanto necesitamos los dos primeros momentos muestrales y los dos primeros momentos poblacionales.

$\mu_1 = E(X) = \mu$ y $m_2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 = \bar{x}$

Se igualan los primeros momentos, y tenemos que: $\hat{\mu} = \bar{x}$

Se igualan también los segundos momentos: $\mu_2 = E\left(X^2\right) = \sigma^2 + \mu^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 = m_2$

Sustituyendo el estimador de la media población y despejando al estimador de la varianza:

$$\tilde{\sigma}^2 + \tilde{\mu}^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2$$

$$\tilde{\sigma}^2 + \left(\bar{x}\right)^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2$$

$$\tilde{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \left(\bar{x}\right)^2$$

$$= \frac{1}{n}\left[\sum_{i=1}^{n} x_i^2 - n\left(\bar{x}\right)^2\right]$$

$$\therefore \tilde{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2$$

Problema 4._ Determina los estimadores por momentos para los parámetros α, β de la distribución gamma, en base a una m.a de tamaño *n*.

<u>Solución</u>

Igualando los dos primeros momentos poblacionales con los muestrales tenemos un sistema de dos ecuaciones con dos incógnitas que se tiene que resolver.

$$\mu_1 = \frac{\tilde{\alpha}}{\tilde{\beta}} = \bar{x} = m_1$$

$$\mu_2 = \frac{\tilde{\alpha}}{\tilde{\beta}^2} + \frac{\tilde{\alpha}^2}{\tilde{\beta}^2} = \frac{1}{n}\sum_{i=1}^{n} x_i^2 = m_2$$

Despejando $\tilde{\alpha}$ de la primera ecuación y sustituyendo en la segunda tenemos que:

$$\tilde{\alpha} = \tilde{\beta}\ \bar{x}$$

$$\frac{\tilde{\beta}\ \bar{x}}{\tilde{\beta}^2} + \frac{\tilde{\beta}^2 \bar{x}^2}{\tilde{\beta}^2} = \frac{1}{n}\sum_{i=1}^{n} x_i^2$$

$$\frac{\bar{x}}{\tilde{\beta}} + \bar{x}^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2$$

$$\frac{\bar{x}}{\tilde{\beta}} = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{x}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2$$

Finalmente se tiene que los estimadores por el método de los momentos para los parámetros de la distribución Gamma son:

$$\tilde{\beta} = \frac{n\bar{x}}{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2} \quad \text{y} \quad \tilde{\alpha} = \tilde{\beta}\,\bar{x} = \frac{n\bar{x}^2}{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}$$

Problema 5._ Encuentra el estimador máximo verosímil de p en base a una m.a de tamaño n de una población Bernoulli.

<u>Solución</u>

La f.d.p de una distribución Bernoulli está dada por:
$f_X(x;p) = p^x(1-p)^{1-x}$ si $x = 0,\ 1$

i) La función de verosimilitud de una m.a de tamaño n está dada por:

$$L(p) = \prod_{i=1}^{n} f\left(x_i, p\right) = \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i} = p^{\sum_{i=1}^{n} x_i}(1-p)^{n-\sum_{i=1}^{n} x_i}$$

ii) Sacamos el logaritmo natural de la función de verosimilitud.

$$\ln L(p) = \left(\sum_{i=1}^{n} x_i\right)\ln(p) + \left(n - \sum_{i=1}^{n} x_i\right)\ln(1-p)$$

iii) Obtenemos la derivada parcial respecto a p.

$$\frac{\partial}{\partial p}\ln L(p) = \frac{\sum_{i=1}^{n} x_i}{p} - \frac{n - \sum_{i=1}^{n} x_i}{1-p}$$

iv) Igualamos a cero y resolvemos la ecuación $\frac{\partial}{\partial p}\ln L(p) = 0$. En el momento en que se iguala a cero se le pone un "gorrito" al parámetro para indicar que se trata de su estimador máximo verosímil.

$$\frac{\sum_{i=1}^{n} x_i}{\hat{p}} - \frac{n - \sum_{i=1}^{n} x_i}{1-\hat{p}} = 0$$

$$\frac{\sum_{i=1}^{n} x_i}{\hat{p}} = \frac{n - \sum_{i=1}^{n} x_i}{1-\hat{p}}$$

$$\frac{1-\hat{p}}{\hat{p}}=\frac{n-\sum_{i=1}^{n}x_i}{\sum_{i=1}^{n}x_i}$$

$$\frac{1}{\hat{p}}-1=\frac{n}{\sum_{i=1}^{n}x_i}-1$$

$$\frac{1}{\hat{p}}=\frac{n}{\sum_{i=1}^{n}x_i}\ \therefore\ \hat{p}=\bar{x}$$

Problema 6._ Se toma una muestra aleatoria de tamaño n de una población con f.d.p de Poisson con media λ.

a) Encuentra el estimador máximo verosímil para λ.
b) Encuentra el valor esperado y la varianza del estimador.
c) ¿Es un estimador consistente?

Solución

a) La f.d.p de una distribución Poisson está dada por:

$$P(x;\lambda)=\frac{e^{-\lambda}\lambda^{x}}{x!}\quad \text{si } x=0,1,2,\ldots$$

i) La función de verosimilitud está dada por:

$$L(\lambda)=\prod_{i=1}^{n}\frac{\lambda^{x_i}e^{-\lambda}}{x_i!}=e^{-\lambda n}\frac{\lambda^{\sum_{i=1}^{n}x_i}}{\prod_{i=1}^{n}x_i!}$$

$$\ln L(\lambda)=-\lambda n+\sum_{i=1}^{n}x_i\ln(\lambda)-\ln\prod_{i=1}^{n}x_i!$$

ii) Igualando a cero y resolviendo encontramos el estimador máximo verosímil.

$$n=\sum_{i=1}^{n}\frac{x_i}{\lambda}$$

$$\hat{\lambda}=\frac{1}{n}\sum_{i=1}^{n}x_i=\bar{x}\quad\therefore\quad\hat{\lambda}=\bar{x}$$

b) Valor esperado y varianza del estimador: $E(\bar{x})=\lambda$ y $V(\bar{x})=\frac{\lambda}{n}$

c) Para saber si el estimador es consistente debemos probar que $\lim_{n\to\infty}ECM(\hat{\theta}_n)=0$

Pero como es un estimador insesgado sabemos que $ECM(\hat{\theta}_n)=V(\hat{\theta}_n)$

por lo tanto $\lim_{n\to\infty}V(\bar{x})=\lim_{n\to\infty}\frac{\lambda}{n}=0$ y entonces el estimador sí es consistente.

Problema 7._ Sea $Y_1, Y_2, \ldots, Y_n$ una m.a de una distribución uniforme con función de densidad de probabilidad $f_Y(y) = \dfrac{1}{2\theta + 1}$ $\quad 0 \le y \le 2\theta + 1$. Encuentra el estimador máximo verosímil para θ.

<u>Solución</u>

La función de verosimilitud está dada por:

$$L(\theta) = \prod_{i=1}^{n} \frac{1}{2\theta + 1} = \frac{1}{(2\theta + 1)^n}$$

$$\ln L(\theta) = -n \ln(2\theta + 1)$$

$$\frac{\partial}{\partial \lambda} \ln L(\theta) = \frac{-2n}{2\theta + 1}$$

Observaciones de la muestra ordenada.

Vemos que la función $L(\theta)$ no se hará cero para ningún valor de θ, sin embargo $L(\theta)$ crece cuando θ decrece, por lo tanto alcanza su valor máximo cuando (2θ+1) es el más pequeño posible, pero de acuerdo a la muestra lo menos que puede valer (2θ+1) es $Y_{(n)}$ por lo tanto $\hat{\theta} = \frac{1}{2}\left(Y_{(n)} - 1\right)$.

Problema 8._ Cierto tipo de componente electrónico tiene una duración *Y* en horas, con f.d.p. $f_Y(y;\theta) = \dfrac{1}{\theta^2} y e^{-y/\theta}$ con $y > 0$. Sea $\hat{\theta}$ el estimador máximo verosímil de θ. Supón que tres componentes al probarlos de manera independiente presentan duración de 120, 130 y 128 hrs.

a) ¿Cuál es la estimación máximo verosímil de θ?
b) ¿Cuánto valen la esperanza y la varianza del estimador?
c) ¿Es un estimador consistente?

<u>Solución</u>

a) Primero obtendremos el estimador en forma general y después se lo aplicaremos a la muestra.

i) La función de verosimilitud está dada por:

$$L(\theta) = \prod_{i=1}^{n} \frac{1}{\theta^2} y_i e^{-y_i/\theta} = \frac{1}{\theta^{2n}} \left(\prod_{i=1}^{n} y_i \right) e^{-\frac{1}{\theta}\sum_{1}^{n} y_i}$$

$$\ln L(\theta) = -2n \ln \theta + \ln \prod_{i=1}^{n} y_i - \frac{1}{\theta} \sum_{i=1}^{n} y_i$$

$$\frac{\partial}{\partial \lambda} \ln L(\theta) = \frac{-2n}{\theta} + \frac{1}{\theta^2} \sum_{i=1}^{n} y_i$$

ii) Igualando a cero y resolviendo encontramos el estimador máximo verosímil.

$$\frac{-2n}{\hat{\theta}}+\frac{1}{\hat{\theta}^2}\sum_{i=1}^{n} y_i = 0 \quad \therefore \ \hat{\theta}=\frac{\sum_{i=1}^{n} y_i}{2n}$$

Ahora sean $y_1 = 120,\ y_2 = 130$ y $y_3 = 128$.

Sustituyendo tenemos que: $\hat{\theta}=\dfrac{120+130+128}{2(3)}=63$

b) Para calcular la esperanza y la varianza del estimador primero debemos obtener la media y la varianza de la f.d.p.

Para la esperanza se tiene: $E(Y)=\displaystyle\int_0^{\infty}\frac{1}{\theta^2}y^2e^{-y/\theta}dy$

Integrando por partes: $E(Y)=\left[-\dfrac{1}{\theta}y^2e^{-y/\theta}-2ye^{-y/\theta}-2\theta\ e^{-y/\theta}\right]_0^{\infty} \ \therefore \ E(Y)=2\theta$

Para la varianza: $E(Y^2)=\displaystyle\int_0^{\infty}\frac{1}{\theta^2}y^3e^{-y/\theta}dy$

$$E(Y^2)=\left[-\frac{1}{\theta}y^3e^{-y/\theta}-3y^2e^{-y/\theta}-6\theta ye^{-y/\theta}-6\theta^2e^{-y/\theta}\right]_0^{\infty} \ \therefore \ E(Y^2)=6\theta^2$$

$$V(Y)=E(Y^2)-(E(Y))^2=6\theta^2-4\theta^2=2\theta^2 \quad \therefore \ V(Y)=2\theta^2$$

Finalmente tenemos que la media y la varianza del estimador están dados por:

$$E\left(\hat{\theta}\right)=E\left(\frac{\sum_{i=1}^{n} y_i}{2n}\right)=\frac{1}{2n}\sum_{i=1}^{n}E\left(y_i\right)=\frac{1}{2n}\sum_{i=1}^{n}2\theta=\theta \quad \text{por lo tanto } E\left(\hat{\theta}\right)=\theta$$

$$V\left(\hat{\theta}\right)=V\left(\frac{\sum_{i=1}^{n} y_i}{2n}\right)=\frac{1}{4n^2}\sum_{i=1}^{n}V\left(y_i\right)=\frac{1}{4n^2}\sum_{i=1}^{n}2\theta^2=\frac{\theta^2}{2n}$$

por lo tanto $V\left(\hat{\theta}\right)=\dfrac{\theta^2}{2n}$

c) $\lim\limits_{n\to\infty}ECM\left(\hat{\theta}\right)=\lim\limits_{n\to\infty}V\left(\hat{\theta}\right)=\lim\limits_{n\to\infty}\dfrac{\theta^2}{2n}=0$ entonces $\hat{\theta}$ sí es un estimador consistente para θ.

II) INTERVALOS DE CONFIANZA.

Problema 1._ Se desea un intervalo de confianza (IC) para el promedio verdadero de pérdida de carga μ (watts) para cierto tipo de motor de inducción cuando la corriente de línea se mantiene en 10 Amps. Para una velocidad de 1500 rpm. Supón que la pérdida de carga está normalmente distribuida con $\sigma = 30$.

a) Calcula un IC de 95% para μ cuando $n = 25$ y $\bar{x} = 58.3$.

b) Calcula un IC de 95% para μ cuando $n = 100$ y $\bar{x} = 58.3$.

c) Calcula un IC de 99% para μ cuando $n = 100$ y $\bar{x} = 58.3$.

d) Calcula un IC de 82% para μ cuando $n = 100$ y $\bar{x} = 58.3$.

<u>**Solución**</u>

El IC para la media de una distribución normal con varianza conocida, es de la forma:

$$\left[\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}},\ \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$$

a) La confianza es del 95%, es decir, $1-\alpha = 0.95$ por lo tanto $\alpha = 0.05$ y $\alpha/2 = 0.025$, buscando en la tabla de la distribución normal tenemos que $z_{0.025} = 1.96$. Sustituyendo los datos en la expresión del IC se tiene que:

$$\left[58.3 - (1.96)\frac{30}{\sqrt{25}},\ 58.3 + (1.96)\frac{30}{\sqrt{25}}\right] = [46.54,\ 70.06]$$

Podemos asegurar con una confianza del 95% que el promedio de pérdida de carga para cierto tipo de motor de inducción cuando la corriente de línea se mantiene en 10 Amps está entre 46.54 y 70.06 watts.

b) Sustituyendo en la expresión del IC se tiene que:

$$\left[58.3 - (1.96)\frac{30}{\sqrt{100}},\ 58.3 + (1.96)\frac{30}{\sqrt{100}}\right] = [52.42,\ 64.18]$$

c) Para una confianza de 99% tenemos que $z_{\alpha/2} = z_{0.005} = 2.575$, sustituyendo los datos para este inciso en la expresión del IC se tiene que:

$$\left[58.3 - (2.575)\frac{30}{\sqrt{100}},\ 58.3 + (2.575)\frac{30}{\sqrt{100}}\right] = [50.575,\ 66.025]$$

d) Para una confianza de 82% tenemos que $z_{\alpha/2} = z_{0.09} = 1.34$. Entonces el intervalo de confianza es de la forma.

$$\left[58.3 - (1.34)\frac{30}{\sqrt{100}},\ 58.3 + (1.34)\frac{30}{\sqrt{100}}\right] = [54.28, 62.82]$$

Problema 2._ Para cada una de 18 muestras de perforación de depósitos de carbonato impregnado de petróleo, se midió la cantidad de saturación de gas residual después de una inyección de solvente de un flujo de agua. Las observaciones, en porcentaje de volumen de poros, fueron:

23.5	31.5	34.0	46.7	45.6	32.5
41.4	37.2	42.5	46.9	51.5	36.4
44.5	35.7	33.5	39.3	22.0	51.2

Calcula un intervalo de confianza al 98% para el verdadero promedio de cantidad de saturación de gas residual.

<u>Solución</u>

Para el caso de una muestra pequeña con varianza desconocida el intervalo de confianza es de la forma:

$$\left[\bar{x} - t_{\alpha/2,n-1}\frac{s}{\sqrt{n}},\ \bar{x} - t_{\alpha/2,n-1}\frac{s}{\sqrt{n}}\right]$$

Primero calculamos la media muestral y la varianza muestral.

$$\bar{x} = \frac{1}{18}\sum_{i=1}^{18} x_i = \frac{1}{18}(695.9) = 38.66$$

$$s^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right]$$

$$= \frac{1}{17}\left[1220.52 - 18(38.66)^2\right]$$

$$s^2 = 71.795$$

$$s = 8.47$$

La confianza del intervalo es del 98% así que $t_{\alpha/2,n-1} = t_{0.01,17} = 2.567$, sustituyendo los datos en la expresión para el IC tenemos que:

$$\left[38.66 - 2.567\frac{8.47}{\sqrt{18}}, 38.661 + 2.567\frac{8.47}{\sqrt{18}}\right] = [33.53, 43.79]$$

Problema 3._ Considera los siguientes 1000 intervalos de confianza al 95% para μ que un consultor en estadística obtendrá para varios clientes. Supón que los conjuntos de datos sobre los que están basados los intervalos se seleccionan de manera independiente entre sí. ¿Cuántos de estos 1000 intervalos esperas que capturen el valor correspondiente de μ? ¿Cuál es la probabilidad de que entre 940 y 960 de estos intervalos contengan el valor correspondiente de μ? (sugerencia: sea Y el número entre los 1000 intervalos que contienen μ. ¿Qué clase de variable aleatoria es Y?)

Solución

Si la variable aleatoria Y es el número entre los 1000 intervalos que contienen a μ, con probabilidad de contenerla igual a 0.95 sabiendo que hay independencia podemos asegurar que la distribución de esta v.a es binomial con parámetros 1000 y 0.95, es decir, $Y \sim bin(y; n = 1000, p = 0.95)$.

El número de intervalos que esperamos capturen el valor de μ es el valor esperado de la v.a, es decir:
$E(Y) = np = 1000(0.95) = 950$

Para encontrar la probabilidad de que entre 940 y 960 de los intervalos contengan el valor de la media poblacional, usamos el Teorema Central del límite que garantiza que:

$$Y \sim N\left(y; \mu_y = np, \quad \sigma_Y^2 = npq\right)$$
$$Y \sim N\left(x; \mu_y = 1000(0.95), \quad \sigma_Y^2 = 1000(0.95)(0.05)\right)$$
$$Y \sim N\left(x; \mu_y = 950, \quad \sigma_Y^2 = 47.5\right)$$

Calculando la probabilidad y aplicando la corrección por continuidad tenemos que:

$$\begin{aligned}
P(940 \le Y \le 960) &\approx P(939.5 \le Y \le 960.5) \\
&= P\left(\frac{939.5 - 950}{\sqrt{47.5}} \le Z \le \frac{960.5 - 950}{\sqrt{47.5}}\right) \\
&= P(-1.52 \le Z \le 1.52) \\
&= \phi(1.52) - \phi(-1.52) \\
&= 0.9357 - 0.0643 = 0.8714
\end{aligned}$$

Problema 4._ Supón que la porosidad del helio de muestras de carbón tomadas de cualquier veta en particular está normalmente distribuida con una desviación estándar verdadera de 0.75.
a) Calcula un IC de 95% para el verdadero promedio de porosidad de cierta veta, si el promedio de porosidad para 20 especímenes de la veta es de 4.85.
b) Calcula un IC al 98% para el verdadero promedio de porosidad de otra veta con base en 16 especímenes con un promedio de porosidad muestral de 4.56.
c) Di cual estimación *a)* o *b)* es más precisa.
d) ¿Qué tan grande se necesita una muestra si la longitud del intervalo con una confianza del 95% debe ser 0.40?

<u>Solución</u>

a) Los datos disponibles para el problema son $\sigma = 0.75$, $\bar{x} = 4.85$, $n = 20$, confianza del 95%, es decir, $1-\alpha = 0.95$ entonces $z_{\alpha/2} = z_{0.025} = 1.96$. El intervalo de confianza es de la forma:

$$\left[\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}},\ \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$$

Sustituyendo tenemos que: $\left[4.85-(1.96)\dfrac{0.75}{\sqrt{20}},\ 4.85+(1.96)\dfrac{0.75}{\sqrt{20}}\right] = [4.52,\ 5.18]$

b) Para una confianza del 98% se tiene que $z_{\alpha/2} = z_{0.01} = 2.33$. Sustituyendo valores se tiene que el intervalo de confianza es:

$$\left[4.56-(2.33)\frac{0.75}{\sqrt{16}},\ 4.56-(2.33)\frac{0.75}{\sqrt{16}}\right] = [4.12,\ 5.00]$$

c) Para saber cual estimación es más precisa tenemos que encontrar la longitud de cada intervalo y compararlas, cuanto más angosto sea el intervalo más precisa es la estimación.

Longitud para el intervalo del inciso *a)*: $5.18 - 4.52 = 0.66$
Longitud para el intervalo del inciso *b)*: $5.00 - 4.12 = 0.87$

Por lo tanto la estimación más precisa es la del inciso *a)*.

c) Sabemos que $\sigma = 0.75$ y que para una confianza de 95% $z_{\alpha/2} = z_{0.025} = 1.96$
Por otra parte, si el ancho del intervalo es de 0.40 entonces podemos considerar un error máximo de 0.20 y el tamaño de la muestra estará dado por:

$$n = \left[\frac{\sigma\ z_{\alpha/2}}{\varepsilon}\right]^2$$

$$n = \left[\frac{0.75(1.96)}{0.20}\right]^2$$

$$n = 54.02$$

En caso de no ser exacto el cociente se acostumbra tomar el entero inmediato en este caso el tamaño de muestra recomendado sería de 55 observaciones.

Problema 5._ Supón que μ_1 y μ_2 son las verdaderas distancias medias de parada a 50 mph para automóviles de cierto tipo equipados con dos tipos diferentes de sistemas de frenos. Si se sabe que $n_1 = 6$, $\bar{x}_1 = 115.7$, $s_1 = 5$ $n_2 = 6$, $\bar{x}_2 = 129.3$ y $s_2 = 5.38$. Calcular un intervalo de confianza de 95% para la diferencia entre el verdadero promedio de distancia de parada para automóviles equipados con el sistema 1 y automóviles equipados con el sistema 2.

<u>Solución</u>

Dado que no se conocen las varianzas poblacionales primero necesitamos averiguar si éstas son estadísticamente diferentes o no. Para ello construimos el intervalo de confianza para el cociente de las varianzas poblacionales, si tal intervalo contiene al 1 se concluye que las varianzas aunque desconocidas se pueden considerar estadísticamente iguales.

La forma del IC para $\dfrac{\sigma_1^2}{\sigma_2^2}$ es $\left[\left(\dfrac{s_1}{s_2}\right)^2 F_{1-\frac{\alpha}{2},n_2-1,n_1-1} \quad , \quad \left(\dfrac{s_1}{s_2}\right)^2 F_{\frac{\alpha}{2},n_2-1,n_1-1}\right]$

Dado que la confianza es del 95% se tiene que:

$F_{1-\frac{\alpha}{2},n_2-1,n_1-1} = F_{0.975,5,5} = 0.14$

$F_{\frac{\alpha}{2},n_2-1,n_1-1} = F_{0.025,5,5} = 7.146$

Sustituyendo valores: $\left[\left(\dfrac{5}{5.38}\right)^2 (0.14) \quad , \quad \left(\dfrac{5}{5.38}\right)^2 (7.146)\right] = [0.12, 6.17]$

El intervalo de confianza para $\dfrac{\sigma_1^2}{\sigma_2^2}$ sí contiene al 1, entonces se concluye que las varianzas poblacionales son estadísticamente iguales.

El intervalo de confianza para una diferencia de medias con varianzas iguales pero desconocidas es de la forma: $\left[\bar{x}_1 - \bar{x}_2 - S_p\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}\, t_{\alpha/2,n_1+n_2-2} \ , \ \bar{x}_1 - \bar{x}_2 - S_p\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}\, t_{\alpha/2,n_1+n_2-2}\right]$

Con $S_p^2 = \dfrac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2} = \dfrac{5(5^2)+5(5.38^2)}{6+6-2} = 26.9722$

Entonces $S_p = \sqrt{26.9722} = 5.19$ y $t_{\alpha/2,n_1+n_2-2} = t_{0.025,10} = 2.228$.

Sustituyendo valores:

$$\left[115.7 - 129.3 - 5.19\sqrt{\frac{1}{6}+\frac{1}{6}}(2.228), 115.7 - 129.3 + 5.19\sqrt{\frac{1}{6}+\frac{1}{6}}(2.228)\right] = [-20.28, -6.92]$$

Dado que el IC contiene sólo números negativos se concluye que la diferencia entre las distancias medias de parada es menor cuando se usa el sistema de frenos tipo 1 que cuando se usa el sistema de frenos tipo 2.

Problema 6._ En un estudio de 277 compradoras adultas seleccionadas al azar, 69 dijeron que siempre que un artículo anunciado no se encontraba en su supermercado local, solicitaban vale. Obtén un intervalo de confianza de 99% para la verdadera proporción de compradoras adultas que solicitan vale en tales situaciones.

Solución

El estimador puntual para la proporción de compradoras adultas que solicitan vale es $p = \frac{69}{277} = 0.249$ y $z_{\alpha/2} = z_{0.005} = 2.575$.

Utilizamos la expresión:

$$\left[p - z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}},\ p + z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}} \right] = \left[0.249 - 2.575\sqrt{\frac{0.249(0.751)}{277}}, 0.249 + 2.575\sqrt{\frac{0.249(0.751)}{277}} \right]$$

El intervalo de confianza es $[0.18, 0.32]$. Se puede asegurar con una confianza del 99% que la verdadera proporción de compradoras adultas que solicitan vale está entre 0.18 y 0.32.

Problema 7._ Un artículo de *Los Angeles Times* reporta que la técnica gráfica de teletermometría de esfuerzo (GST) detectó con precisión 23 de 29 casos conocidos de cáncer de pecho. Construye un intervalo de confianza de 90% para la verdadera proporción de cánceres de pecho que serían detectados por la técnica GST (como *n* es pequeña, el intervalo será muy amplio).

Solución

El intervalo de confianza es de la forma $\left[p - z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}},\ p + z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}} \right]$

Con $p = \frac{23}{29} = 0.793$ y $z_{\alpha/2} = z_{0.05} = 1.645$. Sustituyendo valores se tiene que:

$$\left[0.793 - 1.645\sqrt{\frac{0.793(0.207)}{29}}, 0.793 + 1.645\sqrt{\frac{0.793(0.207)}{29}} \right] = [0.67, 0.92]$$

Es decir, podemos estar seguros en un 90% de que la verdadera proporción de cánceres de pecho que serían detectados por la técnica GST está entre el 67 % y el 92%.

Problema 8._ Se determinó la cantidad de expansión lateral (mils) para una muestra de 9 soldaduras de arco de metal y gas accionado por pulsos, que se emplean en tanques contenedores de gas licuado natural en barcos. La desviación estándar muestral resultante fue $s = 2.81$ mils. Si se supone normalidad, deriva un intervalo de confianza de 95% para σ^2 y para σ.

__Solución__

El intervalo de confianza para la varianza poblacional tiene la forma:

$$\left[\frac{(n-1)s^2}{\chi^2_{\alpha/2,n-1}}, \frac{(n-1)s^2}{\chi^2_{1-\alpha/2,n-1}}\right]$$

Sustituyendo los valores de $n = 9$, $s = 2.81$ junto con los valores de la tabla de la distribución χ^2 $\chi^2_{\alpha/2,n-1} = \chi^2_{0.025,8} = 17.534$ y $\chi^2_{1-\alpha/2,n-1} = \chi^2_{0.975,8} = 2.18$ en la expresión obtenemos:

$$\left[\frac{(9-1)(2.81)^2}{17.534}, \frac{(9-1)(2.81)^2}{2.18}\right] = [3.60, 28.98]$$

El intervalo de confianza para la desviación estándar se calcula sacando la raíz cuadrada de los extremos del IC para la varianza. Entonces el intervalo de confianza de 95% para σ es $[1.9, 5.38]$.

Problema 9._ Se hicieron las siguientes 22 observaciones de resistencia a la fractura de placas base 18% de acero marginizado al níquel:

69.5	71.9	72.6	73.1	73.3	73.5	75.5	79.6
75.8	76.1	76.2	76.2	77.0	77.9	78.1	
79.7	79.9	80.1	82.2	83.7	93.7	75.7	

Calcula el intervalo de confianza de 99% para la desviación estándar de la distribución de la resistencia a la fractura. ¿Es válido este intervalo, cualquiera que sea la naturaleza de la distribución? Explica.

__Solución__

Los valores para la media y la varianza de la muestra anterior son $\bar{x} = 77.33$ y $s^2 = 25.36$.

Suponiendo que las observaciones anteriores son una muestra aleatoria de una distribución normal, el intervalo de confianza para la varianza tiene la forma:

$$\left[\frac{(n-1)s^2}{\chi^2_{\alpha/2,n-1}}, \frac{(n-1)s^2}{\chi^2_{1-\alpha/2,n-1}}\right]$$

Los valores críticos de la distribución ji-cuadrada para una confianza del 99% son: $\chi^2_{\alpha/2,n-1} = \chi^2_{0.005,21} = 41.399$ y $\chi^2_{1-\alpha/2,n-1} = \chi^2_{0.995,21} = 8.033$.

Sustituyendo valores tenemos que el intervalo de confianza para la varianza poblacional es:

$$\left[\frac{(22-1)(25.36)}{41.399} \le \sigma^2 \le \frac{(22-1)(25.36)}{8.033}\right] = [12.86, 66.30]$$

Sacando raíz cuadrada a los extremos del IC anterior se encuentra el intervalo de confianza para la desviación estándar, $[3.59, 8.14]$.

Para que este intervalo sea válido debe tratarse de una muestra aleatoria de una distribución normal de cualquier tamaño o de una distribución cualquiera pero con tamaño de muestra suficientemente grande. Por lo tanto dado que la muestra es pequeña este intervalo es válido solo si la población de donde se saco la muestra tiene una distribución normal.

III) PRUEBAS DE HIPÓTESIS.

Problema 1._ Una tienda departamental tiene un vendedor de quien sospecha comete más errores que el promedio de todos sus vendedores.
a) Si la tienda decide retirar al vendedor si se confirma la sospecha, ¿Qué hipótesis nula y qué hipótesis alternativa deberá utilizar?
b) Si la tienda decide retirar al vendedor a menos que en realidad comenta menos errores que el promedio de todos sus vendedores, ¿Qué hipótesis nula y qué hipótesis alternativa deberá emplear?

Solución

a) Si μ es el número promedio real de errores cometidos por el vendedor y μ_0 es el número promedio de errores cometidos por todos los vendedores, el contraste de hipótesis es el siguiente:

$$H_0 : \mu \leq \mu_0$$
$$H_a : \mu > \mu_0$$

Si la evidencia muestral nos lleva a rechazar la hipótesis nula entonces la tienda retirará al vendedor.

b) El contraste de hipótesis para este inciso es:

$$H_0 : \mu \geq \mu_0$$
$$H_a : \mu < \mu_0$$

El vendedor será despedido si la evidencia estadística apoya a la hipótesis nula.

Problema 2._ ¿Es menos probable que alguien que cambia de marcas por inducción financiera permanezca leal que alguien que cambia sin inducción? Si p_1 y p_2 son las verdaderas proporciones de quienes cambian a una cierta marca con y sin inducción, respectivamente, y que posteriormente repiten una compra. Prueba $H_0 : p_1 - p_2 = 0$ vs. $H_a : p_1 - p_2 < 0$ usando $\alpha = 0.01$ y los siguientes datos:

$n_1 = 200$, número de éxitos igual a 30.

$n_2 = 600$, número de éxitos igual a 180.

Solución

Las proporciones muestrales de quienes cambian a una cierta marca con y sin inducción, respectivamente, y que posteriormente repiten una compra están dadas por:

$$\hat{p}_1 = \frac{30}{200} = 0.15, \; \hat{p}_2 = \frac{180}{600} = 0.3$$

Contraste de hipótesis: $H_0 : p_1 - p_2 \geq 0$

$H_a : p_1 - p_2 < 0$

Estadístico de prueba:

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}} = \frac{0.15 - 0.3}{\sqrt{\frac{0.15(0.85)}{200} + \frac{0.3(0.7)}{600}}} = -4.77$$

Criterio de decisión: Rechazar H_0 si $z < -z_\alpha$

Nivel de significancia: $\alpha = 0.01$

Valor crítico: $-z_\alpha = -z_{0.01} = -2.33$

El estadístico de prueba cae en la región de rechazo y por lo tanto se rechaza la hipótesis nula.
Lo cual apoya la hipótesis de que es menos probable que alguien que cambia de marcas por inducción financiera permanezca leal que alguien que cambia sin inducción.

Problema 3._ Una muestra aleatoria de 5,726 números telefónicos de cierta región tomada en marzo de 1992 dio 1,105 que no estaban en el directorio y, un año después, una muestra de 5,384 dio 980 números que no estaban en el directorio. Prueba un nivel de 0.10 para ver si hay una diferencia en las verdaderas proporciones de números que no aparecen en el directorio entre los dos años.

Solución

X_1 : Número de teléfonos que no se encontraron en el directorio en 1992.
X_2 : Número de teléfonos que no se encontraron en el directorio en 1993.

Datos: $n_1 = 5726$, $n_2 = 5384$

$$\hat{p}_1 = \frac{1105}{5726} = 0.193, \ \hat{p}_2 = \frac{980}{5384} = 0.182$$

Aplicando el Teorema Central de Límite (TCL) para la diferencia de proporciones muestrales tenemos que: $P_1 - P_2 \sim N\left(p_1 - p_2; \mu = p_1 - p_2, \sigma^2 = \frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2} \right)$

Contraste de hipótesis: $H_0 : p_1 - p_2 = 0$
$H_a : p_1 - p_2 \neq 0$

Estadístico de prueba:

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}} = \frac{0.193 - 0.182}{\sqrt{\frac{0.193(0.807)}{5726} + \frac{0.182(0.818)}{5384}}} = 1.48$$

Criterio de decisión: Rechazar H_0 si $|z| > z_{\alpha/2}$

Nivel de significancia: $\alpha = 0.10$

Valor crítico: $z_{\alpha/2} = z_{0.05} = 1.645$

Entonces no se debe rechazar la hipótesis nula.

Por lo tanto se concluye con base a la muestra que no existe una diferencia significativa entre las proporciones de números que no aparecen en el directorio en los dos años considerados.

Problema 4._ En su incansable búsqueda de un sistema de llenado adecuado, cierta empresa prueba dos máquinas. Automat -fill se usa para llenar 21 frascos que dan una desviación estándar de 2.1 onzas. Con Robo-fill se llenan 16 frascos y da una desviación estándar de 1.9 onzas en el llenado. Si la empresa tiene que elegir uno de estos sistemas en función de la uniformidad de llenado. ¿Cuál deberá seleccionar? Usa $\alpha = 0.10$.

Solución

X_1 : Número de onzas de llenado con Automat –fill.
X_2 : Número de onzas de llenado con Robo-fill.

Datos: $n_1 = 21$, $n_2 = 16$

$s_1 = 2.1$, $s_2 = 1.9$

Contraste de hipótesis: $H_0 : {\sigma_1}^2 = {\sigma_2}^2$
$H_a : {\sigma_1}^2 > {\sigma_2}^2$

O equivalentemente: $H_0 : {\sigma_1}^2 - {\sigma_2}^2 = 0$
$H_a : {\sigma_1}^2 - {\sigma_2}^2 > 0$

Estadístico de prueba:

$$F = \frac{{s_1}^2}{{s_2}^2} = \frac{(2.1)^2}{(1.9)^2} = 1.22$$

Criterio de decisión: Rechazar H_0 si $F > F_{\alpha, n_1 - 1, n_2 - 1}$

Nivel de significancia: $\alpha = 0.10$

Valor crítico: $F_{\alpha, n_1 - 1, n_2 - 1} = F_{0.10, 20, 15} = 1.92$

No se rechaza la hipótesis nula. Por lo que se concluye con un nivel de significancia del 10% que no existe diferencia significativa entre la variación de llenado de la máquina Automat-Fill y de la máquina Robo-Fill, por lo que se puede seleccionar cualquiera de las dos máquinas.

Problema 5._ ¿Los estudiantes universitarios hombres se aburren más fácilmente que sus compañeras mujeres? Esta pregunta se examino en el artículo "Boredom in Young Adult – Gender and cultural Comparisons". Los autores aplicaron la escala de propensión al aburrimiento a 97 estudiantes hombres y 148 mujeres de universidades de Estados Unidos. ¿La siguiente información apoya la hipótesis de investigación de que la tasa de aburrimiento es más alta para hombres que para mujeres? Prueba las hipótesis apropiadas usando un nivel de significancia de 0.05.

Género	**Tamaño muestral**	**Media muestral**	**Desviación estándar**
Hombres	97	10.40	4.83
Mujeres	148	9.26	4.68

Solución

X_1 : Número de hombres que se aburren.
X_2 : Número de mujeres que se aburren.

Datos: $n_1 = 97$, $n_2 = 148$
$\bar{x}_1 = 10.40$, $s_1 = 4.83$
$\bar{x}_2 = 9.26$, $s_2 = 4.68$

Contraste de hipótesis: $H_0 : \mu_1 = \mu_2$
$H_a : \mu_1 > \mu_2$

Estadístico de prueba:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^{\,2}}{n_1} + \frac{s_2^{\,2}}{n_2}}} = \frac{10.40 - 9.26}{\sqrt{\frac{(4.83)^2}{97} + \frac{(4.68)^2}{148}}} = 1.82$$

$$gl = \frac{\left[\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right]^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}} = \frac{\left[\frac{(4.83)^2}{97} + \frac{(4.68)^2}{148}\right]^2}{\frac{\left(\frac{(4.83)^2}{97}\right)^2}{97 - 1} + \frac{\left(\frac{(4.68)^2}{148}\right)^2}{148 - 1}} = 200.83 \approx 201$$

Nivel de significancia: $\alpha = 0.05$

Criterio de decisión: Rechazar H_0 si $t > t_{\alpha, n-1}$

Valor crítico: $t_{0.05,200} = 1.645$

Entonces se debe rechazar la hipótesis nula. Por lo tanto se concluye que no hay condiciones para afirmar que los estudiantes universitarios hombres se aburren más fácilmente que sus compañeras mujeres.

Problema 6._ El artículo "Statistical Evidence of Discrimination" (J. Amer. Star. Assoc. 1982, pp, 773-783) analiza el caso judicial Swain vs Alabama (1965), en el cual se dijo que había discriminación contra negros en la selección del gran jurado. Los datos de un censo sugirieron que 25% de los elegibles para prestar servicio como gran jurado eran negros, pero una muestra aleatoria del 1,050 llamados para presentarse para un posible servicio dio por resultado solo 177 negros. Mediante el uso de una prueba con un nivel de significancia de 0.01, ¿se concluye que hay discriminación?

__Solución__

X : Número de negros llamados para un posible servicio.

Datos: $n = 1050$

$$p = 0.25\,,\ p_0 = \frac{177}{1050} = 0.168\,.$$

Contraste de hipótesis:

$H_0 : p = p_0$ $\qquad$ $H_0 : p - p_0 = 0$

$H_a : p > p_0$ $\qquad$ $H_a : p - p_0 > 0$

Estadístico de prueba:

$$z = \frac{\hat{p} - \hat{p}_0}{\sqrt{\dfrac{\hat{p}_0 \hat{q}_0}{n}}} = \frac{0.25 - 0.168}{\sqrt{\dfrac{0.168(0.832)}{1050}}} = 7.10$$

Criterio de decisión: Rechazar H_0 si $z > z_\alpha$

Nivel de significancia: $\alpha = 0.01$

Valor crítico: $z_\alpha = z_{0.01} = -2.33$

Tenemos que $z = 7.10 > -2.33 = z_\alpha$ entonces se debe rechazar la hipótesis nula y se concluye con un nivel de significancia de $\alpha = 0.01$ que sí existe discriminación contra los negros para la selección del gran jurado.

Problema 7._ El *Instituto Eléctrico Edison* publica cifras del número anual de Kilowatt-hora que gastan varios aparatos electrodomésticos. Se afirma que una aspiradora gasta un promedio de 46 kilowatt-hora al año. Si una muestra aleatoria de 12 hogares que se incluye en un estudio planeado indica que las aspiradoras gastan un promedio de 42 kilowatt-hora al año con una desviación estándar de11.9 kilowatt-hora, ¿esto sugiere con un nivel de significancia de 0.05 que las aspiradoras gastan, en promedio, menos de 46 kilowatt-hora anualmente? Supón que la población de kilowatt-hora es normal.

Solución

Datos: $n = 12$

$\mu_0 = 46$, $\bar{x} = 42$, $s = 11.9$

Contraste de hipótesis:

$H_0 : \mu = \mu_0$ $\quad\quad$ $H_0 : \mu - \mu_0 = 0$

$H_a : \mu < \mu_0$ $\quad\quad$ $H_a : \mu - \mu_0 < 0$

Estadístico de prueba:

$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{42 - 46}{\frac{11.9}{\sqrt{12}}} = -1.16$$

Criterio de decisión: Rechazar H_0 si $t < -t_{\alpha,n-1}$

Nivel de significancia: $\alpha = 0.05$

Valor crítico: $-t_{\alpha,n-1} = -t_{0.05,12-1} = -t_{0.05,11} = -1.796$

Se tiene que $t = -1.16 > -1.796 = -t_{\alpha,n-1}$ por lo tanto no se rechaza la hipótesis nula.

Se concluye con un nivel de significancia de $\alpha = 0.05$ que el número promedio de kilowatt-hora gastadas al año por las aspiradoras no es significativamente menor que 46.

Problema 8._ Una empresa empacadora de azúcar está considerando una máquina nueva para reemplazar su máquina actual. Los pesos de una muestra de 21 paquetes de 5 libras empacados por la máquina vieja producen una varianza de 0.16, mientras que los pesos de 20 paquetes de 5 libras empacados por la máquina nueva dan una varianza de 0.09. En base a estos datos, ¿aconsejaría usted al gerente comprar la máquina nueva? Use un $\alpha = 0.05$.

Solución

X_1 : Número de libras en los paquetes empacados por la máquina vieja.
X_2 : Número de libras en los paquetes empacados por la máquina nueva.

Datos: $n_1 = 21$, $n_2 = 20$

$s_1^2 = 0.16$, $s_2^2 = 0.09$

Contraste de hipótesis:

$H_0 : \sigma_1^2 = \sigma_2^2 \qquad H_0 : \sigma_1^2 - \sigma_2^2 = 0$

$H_a : \sigma_1^2 > \sigma_2^2 \qquad H_a : \sigma_1^2 - \sigma_2^2 > 0$

Estadístico de prueba:

$$F = \frac{s_1^2}{s_2^2} = \frac{0.16}{0.09} = 1.77$$

Criterio de decisión: Rechazar H_0 si $F > F_{\alpha, n_1-1, n_2-1}$

Nivel de significancia: $\alpha = 0.05$

Valor crítico: $F_{\alpha, n_1-1, n_2-1} = F_{0.05, 21-1, 20-1} = F_{0.05, 20, 19} = 2.16$

Como $1.77 < 2.16$ no se rechaza H_0 y se concluye que no es conveniente comprar la máquina nueva.

Bibliografía.

1) De Groot, Morris H. *Probabilidad y estadística*. México: Addison Wesley Iberoamericana, 694 p. ISBN 0-201-64405-3.

2) Devore, Jay L. *Probabilidad y estadística para ingeniería y ciencias.* 5ª edición. México, D.F: Thomson, Learning, 2001. 762 p. ISBN 0-534-37281-3

3) Hines, William H; Douglas C. Montgomery. *Probabilidad y estadística para ingeniería y administración.* 3ª edición. México, D.F: CECSA, 834 p.

4) Mendenhall, William; et al. *Estadística matemática con aplicaciones*. 2ª edición, México, D.F: Gpo. Editorial Iberoamérica, 1991. 772 p. ISBN 0-87150-939-3

5) Meyer. *Probabilidad y aplicaciones estadísticas.* Edición revisada. México, D.F: Addison Wesley Longman, 480 p. ISBN 968-4444-2211

6) Milton, J.S; Arnold, J.C. *Probabilidad y estadística con aplicaciones para ingeniería y ciencias computacionales*. México, D.F: McGraw Hill, 2004. 804 p. ISBN 0-07-246836-X

7) Navidi, Willian. *Estadística para ingenieros y científicos*. México, D.F: McGraw Hill, 2006. 868 p. ISBN 970-10-5629-9

8) Quevedo, Héctor; Pérez Blanca R. *Estadística para ingeniería y ciencias*. México, D.F: Grupo Editorial Patria, 2008. 440 p. ISBN 978-970-817-232-5

9) Velasco, Gabriel; Wisniewski, P.M. *Probabilidad y estadística para ingeniería y ciencias.* Edición. México, D.F: Thomson, 2001. 326 p. ISBN 970-686-136-X

10) Walpole, Myers & Myers. *Probabilidad y estadística para ingeniería y ciencias.* 8ª edición. México, D.F: Pearson, Prentice Hall, 2007. 816 p. ISBN 10: 970-26-0936-4

11) Weimer, Richard *C. Estadística.* 4 ª reimpresión. México, D.F: CECSA, 2002. 839 p. ISBN 968-26-1261-6