INSTITUTO POLITÉCNICO NACIONAL
ESCUELA SUPERIOR DE CÓMPUTO
DEPARTAMENTO DE FORMACIÓN BÁSICA

NOTAS DE PROBABILIDAD Y ESTADÍSTICA

PROFESORA: LETICIA CAÑEDO SUÁREZ

1 INTRODUCCIÓN A LA PROBABILIDAD.

OBJETIVO: Aprender a calcular la probabilidad de un evento a partir de las probabilidades de otros eventos relacionados, esto se conoce como Regla de Bayes, para lo cual necesitamos aprender los conceptos básicos.

1.1 GENERALIDADES.

La teoría de la probabilidad es la rama de las matemáticas que estudia los fenómenos que dependen de sucesos fortuitos conocidos también con el nombre de azarosos o aleatorios.

Un poco de historia:

Existe evidencia de la existencia de los juegos de azar desde el año 3500 AC en Egipto, pero el sustento matemático de la teoría de la probabilidad es iniciado por los matemáticos franceses Blaise Pascal (1623-1662) y Pierre Fermat (1601-1665) cuando logran obtener probabilidades exactas para ciertos problemas relacionados con los juegos de dados.
A partir del siglo XVII la teoría de la probabilidad ha sido constantemente desarrollada y ampliamente aplicada en diversos campos de estudio tales como las ciencias sociales, política, publicidad y mercadotecnia, ingeniería, administración, medicina, meteorología, predicción de terremotos, comportamiento humano, diseño de sistemas de computadores, discriminación laboral, etc.

1.1.1 MODELOS PROBABILÍSTICOS Y DETERMINÍSTICOS.

Cada vez que utilizamos las matemáticas con el objeto de estudiar fenómenos observables es indispensable empezar por construir un modelo matemático para tales fenómenos. Estos pueden ser de dos tipos:

i) **Modelo determinístico**._ aquél en que las condiciones en las que se verifica un experimento determinan el resultado del mismo. Por ejemplo, supón que la relación entre una variable de respuesta y con una variable explicativa x es de la forma $y = a + bx$ con a y b desconocidas. Esto es un modelo determinístico porque no permite error en la predicción de y como función de x. Por ejemplo, la distancia recorrida verticalmente sobre el suelo por un objeto está dada por $S = -16t^2 + bV_0t$, donde t es el tiempo y V_0 es la velocidad inicial.

ii) **Modelo probabilístico o estocástico**._ aquél que involucra cierto error que depende de cosas fortuitas por ejemplo $Y = A + Bx + \varepsilon$, donde ε es un componente aleatorio. Ejemplo: Se desea determinar cuánta lluvia caerá debido a una tormenta que pasa por una zona específica.

1.1.2 ESPACIO MUESTRAL Y EVENTOS.

Definición:

i) Un **experimento** es cualquier proceso cuyos resultados no se conocen de antemano con certeza.

ii) El **espacio muestral** S del experimento es la colección o conjunto de **todos** los posibles resultados del experimento.

iii) Un **evento** o **suceso** es cualquier subconjunto del espacio muestral.

Ejemplos:

1._ **Experimento**: Se lanza una moneda 3 veces.
Notación: c indica que la moneda cae en cara y x indica que la moneda cae en cruz.
Espacio muestral: S ={ccc, ccx, cxc, xcc, cxx, xxc, xcx, xxx}
Eventos: A: caen dos caras A ={ccx, cxc, xcc}
B: cae al menos una cara B ={ccc, ccx, cxc, xcc, cxx, xxc, xcx}
C: a lo más una cruz C ={ccc, ccx, cxc, xcc}

2._ **Experimento**: Un vehículo llega a una intersección de carreteras y puede dar vuelta a la derecha, a la izquierda o seguir de frente.
Notación: D indica que el vehículo da vuelta a la derecha, I vuelta a la izquierda y F indica que sigue de frente.
Espacio muestral: S ={D, I, F}
Eventos: A: el vehículo sigue de frente A ={F}
B: El vehículo se detiene B ={}=$\varnothing$

3._ **Experimento**: Se eligen al azar tres artículos de un proceso de manufactura, se inspecciona cada uno de ellos y se clasifica como defectuoso D o bueno B.
Espacio muestral: S ={DDD, DDB, DBD, BDD, DBB, BDB, BBD, BBB}
Eventos: A: exactamente dos artículos buenos A ={DBB, BDB, BBD}
B: a lo más un artículo defectuoso B ={DBB, BDB, BBD, BBB}

4._ **Experimento**: Se lanza una moneda hasta obtener una cruz (x) o tres caras (c).
Espacio muestral: S ={x, cx, ccx, ccc}
Evento: A: al menos dos caras A ={ccx, ccc}

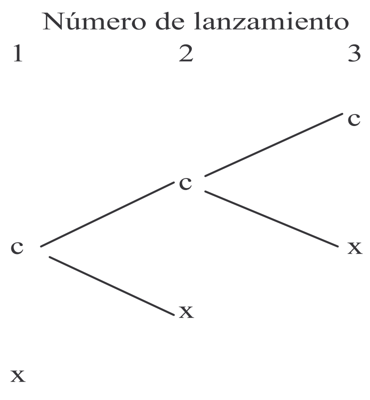


Diagrama de árbol para el problema 4.

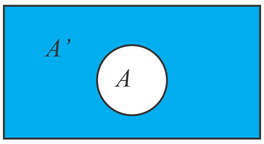
5._ Una puerta de automóvil se ensambla con un gran número de puntos de soldadura. Después del ensamblado, se inspecciona cada punto y se cuenta el número total de defectos.
Notación: N es el número total de puntos de soldadura.
Espacio muestral: $S=\{0, 1, 2, ..., N\}$
Eventos: A: El trabajo es impecable $A=\{0\}$
B: Se tuvo un número de defectos mayor al 50% del total de puntos

$$B=\begin{cases}\{N/2+1, N/2+2, ..., N\} & \text{con } N \text{ par}\\ \{(N+1)/2, (N+1)/2+1, ..., N\} & \text{con } N \text{ impar}\end{cases}$$

6._ Se seleccionan al azar cuatro estudiantes de una clase de química y se clasifican como masculino o femenino. Lista los elementos del espacio muestral asociado a este experimento. Define un segundo espacio muestral donde los elementos representen el número de mujeres seleccionadas.
Notación: f denota a una mujer y m a un hombre.
Espacio muestral 1: S ={ffff, fffm, ffmf, fmff, mfff, ffmm,mmff, fmmf, mffm, mfmf, fmfm, mmmf, mmfm, mfmm, fmmm, mmmm}
Espacio muestsral 2: $S=\{0, 1, 2, 3, 4\}$

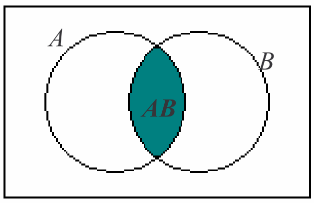
Operaciones entre eventos.
Dado que los eventos son subconjuntos de elementos, todas las operaciones que conocemos para los conjuntos aplican en los eventos.

Definición:
i) El **complemento de un evento** A con respecto al espacio muestral S es el conjunto de todos los elementos de S que no están en A. Se denota como A', A^c o $\overline{A}$. $\boxed{A' = \{x \mid x \in S \text{ y } x \notin A\}}$



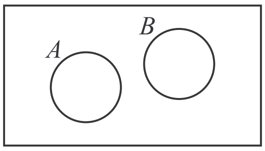
El círculo representa el evento A y la parte sombreada el complemento de A, el rectángulo Es el espacio muestral S.

ii) La **intersección de dos eventos** es el conjunto de elementos que pertenecen a los dos eventos al mismo tiempo y se escribe $\boxed{AB = \{x \mid x \in A \text{ y } x \in B\}}$

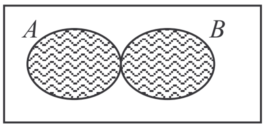


El círculo de la izquierda representa el evento A y el de la derecha el evento B, la parte sombreada es la intersección de ambos eventos y el rectángulo es el espacio muestral.

iii) Dos eventos son **excluyentes** o **disjuntos** si no se intersectan, es decir $\boxed{AB = \varnothing}$



iv) La **unión de dos eventos** es el conjunto de elementos que están en cualquiera de los dos eventos o en ambos y se denota $\boxed{AUB = \{x \mid x \in A \text{ o } x \in B\}}$



Dos leyes que serán de gran ayuda a lo largo de este curso son las llamadas Leyes de D'Morgan:

i) $(AUB)' = A'B'$ El complemento de la unión es la intersección de los complementos.

ii) $(AB)' = A'UB'$ El complemento de la intersección es la unión de los complementos.

1.1.3 Interpretación De La Probabilidad.

Frecuentista._ La probabilidad de un evento es interpretada como la frecuencia relativa con la que se obtendría ese resultado si el proceso se repitiera un gran número de veces, en condiciones similares. Por ejemplo si se lanza una moneda N veces y en exactamente n lanzamientos la moneda cae con la cara hacia arriba se tendría que la probabilidad de obtener cara está dada por $\frac{n}{N}$.

Clásica._ está basada en el concepto de resultados igualmente verosímiles, es decir se asume que cada uno de los resultados tiene la misma oportunidad de ocurrir. En el ejemplo de la moneda tenemos dos posibles resultados, cara o cruz, por lo tanto si la probabilidad de que ocurra una cara es p la probabilidad de que ocurra una cruz también es p, cumpliendo ciertas propiedades que veremos más adelante.

Subjetiva._ tiene que ver con el grado de conocimiento de la persona sobre el evento y en base a esta experiencia la persona asigna una probabilidad conveniente.

1.2 Elementos De Análisis Combinatorio.

Hay ocasiones en que no necesitamos conocer en detalle los elementos del espacio muestral más bien necesitamos saber cuantos son. Así que aprenderemos un poco sobre métodos de conteo.

1.2.1 Permutaciones.

Definición: Una **permutación** es un arreglo de objetos distintos donde el orden es importante.

Una permutación difiere de otra si el orden o el contenido del arreglo es diferente.

Ejemplo._ Considerando las letras a, b, c, d:

i) Hay 4 permutaciones distintas tomando una sola letra a la vez: a, b, c, d.
ii) Hay 12 permutaciones distintas tomando 2 letras a la vez: ab, ba, ac, ca, ad, da, bc, cb, bd, db, cd, dc.
iii) Hay 24 permutaciones distintas tomando 3 letras a la vez: abc, acb, bac, bca, cab, cba, abd, adb, bad, bda, dab, dba, bcd, bdc, cbd, cdb, dbc, dcb, acd, adc, cad, cda, dac, dca.

También se puede utilizar el método de "las casitas" para contar, si quiero ver de cuántas maneras puedo permutar 3 de 4 letras posibles, tendría que la primer letra puede ser cualquiera de las 4 originales, la siguiente puede ser cualquiera de las 3 que quedan y la última puede ser cualquiera de las 2 letras restantes, dibujo tres casitas que corresponden a las tres letras que deseo tomar y las lleno de la manera descrita, quedando

4	3	2

4x3x2 = 24 permutaciones diferentes.

En general si quiero permutar k de n elementos tendría

n	n-1	n-2	...	n-(k-1)

= $n(n\text{-}1)(n\text{-}2)...(n\text{-}k\text{+}1)$

Si multiplicamos por 1 no pasa nada, así el número de permutaciones diferentes está dado por

$$P_k^n = n(n-1)(n-2)...(n-k+1)\frac{(n-k)(n-k-1)(n-k-2)...(2)(1)}{(n-k)(n-k-1)(n-k-2)...(2)(1)}$$

$$\boxed{P_k^n = \frac{n!}{(n-k)!}}$$

1.2.2 Combinaciones.

Definición: Una **combinación** es un arreglo de objetos distintos donde no importa el orden.

Una combinación difiere de otra sólo si el contenido es diferente.
Ejemplo._ Considerando nuevamente las cuatro letras a, b, c, d se tiene que:

i) Hay 4 combinaciones distintas tomando una sola letra a la vez: a, b, c, d.
ii) Hay 6 combinaciones distintas tomando 2 letras a la vez: ab = ba, ac = ca, ad = da, bc = cb, bd = db, cd = dc.
iii) Hay 4 combinaciones distintas tomando 3 letras a la vez: abc = acb = bac = bca = cab = cba, abd = adb = bad = bda = dab = dba, bcd = bdc = cbd = cdb = dbc = dcb, acd = adc = cad = cda = dac = dca.
iv) Sólo 1 manera de sacar las 4 letras a la vez: abcd.

Observa que teníamos 12 permutaciones tomando dos letras a la vez, pero dado que dos letras pueden cambiar el orden 2 veces, al no importar el orden para nosotros es lo mismo el par (ab) que el par (ba), dos letras dos ordenamientos, quitando las repeticiones nos quedamos con 6 combinaciones.
Teníamos 24 permutaciones tomando 3 de 4 letras, pero con las mismas 3 letras podemos hacer 3(2)(1) es igual a 6 ordenamientos diferentes, dado que ahora ya no importa el orden estos 6 arreglos son uno solo, por lo tanto quitando las repeticiones nos quedan 4 combinaciones diferentes. Siguiendo con este razonamiento llegamos a que el número de combinaciones al tomar k de n elementos esta dado por:

$$\boxed{C_k^n = \binom{n}{k} = \frac{n!}{k!(n-k)!}}$$

Ejemplos:

1._ ¿Cuántos equipos que incluyan 2 ingenieros en electrónica y 1 ingeniero en sistemas computacionales se pueden formar si se dispone de 4 ingenieros en electrónica y 3 ingenieros en sistemas computacionales?

<u>**Solución**</u>

El orden no es importante por lo tanto necesitamos combinaciones

$$\binom{4}{2}\binom{3}{1} = \frac{4!}{2!(4-2)!}\frac{3!}{1!(3-1)!} = \frac{2!(3)(4)}{2!(2!)}\frac{2!(3)}{2!} = \underline{18 \text{ equipos.}}$$

2._ ¿Cuántas fichas tiene un domino si se usan n números y cada ficha es de la forma (x, y) siendo x y y no necesariamente diferentes y $(x, y) = (y, x)$?

<u>**Solución**</u>

El orden no importa, entonces primero necesito encontrar el número de combinaciones que tengo al tomar 2 de los n números y luego sumo n pares con números iguales, es decir los pares (1,1), (2,2), ... ,(n,n).

$$C_2^n = \binom{n}{2} = \frac{n!}{2!(n-2)!} = \frac{n(n-1)(n-2)!}{2!(n-2)!} = \frac{n(n-1)}{2}$$

El número total de fichas es entonces

$$\frac{n(n-1)}{2} + n = \frac{n(n-1)+2n}{2} = \frac{n(n-1+2)}{2} = \underline{\frac{n(n+1)}{2}}$$

3._ ¿De cuántas manera se pueden formar 6 personas para subirse a un autobús si ...
a) No hay restricción alguna?
b) Si 3 personas insisten en seguir cada una de ellas a las otras?
c) Si en un determinado conjunto de 2 personas, cada una de ellas se niega a seguir a la otra?

Solución

a) Si no hay restricción alguna, cualquiera de las 6 personas puede ir en la primera posición de la fila, cualquiera de las 5 que quedan en la siguiente posición y así sucesivamente, por lo tanto el número de maneras diferentes de formar a las 6 personas es 6! = 720.

b) Consideramos a las 3 personas que quieren estar juntas como si fuera una sola y entonces tendríamos que acomodar en la fila a 4 personas (las tres que no tienen preferencia y el paquete de tres que desean estar juntas), se pueden acomodar de 4! = 24 maneras diferentes, pero dentro de cada una de estas 24 maneras las 3 personas que quieren estar juntas se pueden acomodar de 3! = 6 maneras diferentes, entonces en total hay 4!(3!) = 144 maneras diferentes de acomodar a las 6 personas con la restricción pedida.

c) Se procede similar al inciso anterior considerando que las dos personas quieren estar siempre juntas, entonces tendríamos 4 personas sin preferencia alguna y un paquete de las 2 que quieren estar juntas, pero estas dos personas se pueden acomodar de 2 maneras diferentes, entonces hay 5!(2!) = 240 maneras en que siempre están juntas, si quitamos estas al número total de maneras en que se pueden acomodar las 6 personas en la fila, me quedo con el número de maneras en que nunca están juntas estas dos personas, es decir, la solución al problema esta dada por, 720-240 = 480 maneras.

4._ El número de formas en que pueden acomodarse n objetos distintos en un círculo es $(n-1)!$

a) Presenta un argumento que justifique esta fórmula.

b) ¿De cuántas maneras pueden montarse 5 turquesas en un brazalete de plata?

Solución

a) Se fija un objeto cualquiera en una posición del circulo y de ahí se acomodan los $(n-1)$ objetos restantes.

b) 4! = 4(3)(2)(1) = 24 maneras de montar las 5 turquesas.

5._ ¿Cuántos números diferentes de 3 dígitos pueden formarse con los números 3, 4, 7, 8 y 9 si se desea que sean mayores que 500 y que no se repitan los dígitos?

Solución

Para que el número sea mayor que 500 debe comenzar con 7, 8 o 9 y como no se pueden repetir, el siguiente dígito puede ser cualquiera de los 4 números que quedan y el tercer dígito cualquiera de los 3 números restantes.

(3)(4)(3) = 36 números.

6._ ¿Cuántos números pares de tres dígitos se pueden formar con los números 1, 2, 5, 6, 9 si sólo es posible utilizar cada uno de estos una sola vez?

Solución

Para que el número sea par debe terminar con 0 o con par, es decir sólo hay dos opciones: 2 o 6. Los dos primeros números pueden ser entonces cualquiera de los 4 y luego cualquiera de los 3 que quedan. (4)(3)(2) = 24 números pares con tres dígitos no repetidos.

7._ ¿Cuántos números telefónicos diferentes de siete dígitos se pueden formar si el primer dígito no puede ser cero?

Solución

No hay restricción sobre las repeticiones así que el número pedido es $9(10)^6$

1.3 AXIOMAS DE PROBABILIDAD.

Definición:
Sea S un espacio muestral asociado a un experimento y A un evento cualquiera de S, **la probabilidad del evento** A es una función que asigna a A un número real $P(A)$ con las siguientes propiedades:

i) $0 \leq P(A) \leq 1$

ii) $P(S) = 1$

iii) Si $A_1, A_2, ..., A_n$ son eventos mutuamente excluyentes en S entonces:

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$$

iv) Si $A_1, A_2, ...$ es una secuencia numerable de eventos mutuamente excluyentes en S, entonces:

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

Teorema._ Sean A, B y C tres eventos cualesquiera de un espacio muestral S asociado a un experimento, entonces:

i) $P(\varnothing) = 0$

ii) $P(A') = 1 - P(A)$

iii) $P(A \cup B) = P(A) + P(B) - P(AB)$

iv) $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC)$

v) $P\left(\bigcup_{i=1}^{k} A_i\right) = \sum_{i=1}^{k} P(A_i) - \sum_{i<j=2}^{k} P(A_i A_j) + \sum_{i<j<r=3}^{k} P(A_i A_j A_r) + \cdots + (-1)^{k-1} P\left(\bigcap_{i=1}^{k} A_i\right)$

vi) Si $A \subset B$ entonces $P(A) \leq P(B)$

Demostración:

i) $S \cup \varnothing = S$

$P(S \cup \varnothing) = P(S)$ pero $S\varnothing = \varnothing$

$$P(S \cup \varnothing) = P(S) + P(\varnothing) = P(S)$$
$$\therefore P(\varnothing) = P(S) - P(S) = 0 \quad \blacklozenge$$

ii) $S = A \cup A'$ con $AA' = \varnothing$
$$P(S) = P(A \cup A') = P(A) + P(A')$$
$$P(A') = P(S) - P(A)$$
$$\therefore \; P(A') = 1 - P(A) \quad \blacklozenge$$

iii) Observa que $A = AB' \cup AB$ y $(AB')(AB) = \varnothing$

$B = AB \cup A'B$ y $(AB)(A'B) = \varnothing$

entonces
$$P(A) = P(AB') + P(AB) \qquad (1)$$
$$P(B) = P(AB) + P(A'B) \qquad (2)$$

También $A \cup B = AB' \cup AB \cup A'B$ y los tres eventos son mutuamente excluyentes
$$\therefore P(A \cup B) = P(AB') + P(AB) + P(A'B) \qquad (3)$$

Despejando $P(AB')$ de (1) y $P(A'B)$ de (2) y sustituyendo en (3)

$$P(A \cup B) = P(A) - P(AB) + P(AB) + P(B) - P(AB)$$
$$P(A \cup B) = P(A) + P(B) - P(AB) \quad \blacklozenge$$

iv) Consideramos la unión de A y B como un solo evento y utilizamos ii) las veces que sea necesario.
$$\begin{aligned} P[(A \cup B) \cup C] &= P(A \cup B) + P(C) - P[(A \cup B)C] \\ &= P(A) + P(B) - P(AB) + P(C) - P[(AC) \cup (BC)] \\ &= P(A) + P(B) - P(AB) + P(C) - [P(AC) + P(BC) - P(ABC)] \\ &= P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC) \quad \blacklozenge \end{aligned}$$

v) Tareita. Sugerencia: Utilizar el método de inducción matemática.

vi) Observa que si $A \subset B$ entonces $B = A \cup A'B$ pero $A \cap A'B = \varnothing$

$$P(B) = P(A) + P(A'B) \quad \text{como } P(A'B) \geq 0$$

Se tiene que si $P(A'B) = 0 \;\; \therefore \;\; P(A) = P(B)$

si $P(A'B) > 0 \;\; \therefore P(A) < P(B) \quad \blacklozenge$

Recuerda: Si el conjunto A tiene a elementos se dice que la cardinalidad de A es a y se escribe $\#(A) = a$.

Definición: La probabilidad del evento A definido en el espacio muestral S está dada por

$$\boxed{P(A) = \frac{\#(A)}{\#(S)}}$$

Ejemplos:

1._ La probabilidad de que Juan apruebe el curso de probabilidad es $\frac{2}{3}$ y la probabilidad de que apruebe el curso de estadística es $\frac{4}{9}$. Si la probabilidad de aprobar ambos cursos es $\frac{1}{4}$. ¿Cuál es la probabilidad de que Juan ...

a) Apruebe cuando menos uno de estos cursos?
b) Repruebe probabilidad?
c) Repruebe las dos materias?

Solución

Sean los eventos p: Juan aprueba el curso de probabilidad.
e: Juan aprueba el curso de estadística.

Datos: $P(p)=\frac{2}{3}$, $P(e)=\frac{4}{9}$ y $P(pe)=\frac{1}{4}$

a) $P(p\cup e)=P(p)+P(e)-P(pe)=\frac{2}{3}+\frac{4}{9}-\frac{1}{4}=\underline{\frac{31}{36}}$

b) $P(p')=1-P(p)=1-\frac{2}{3}=\underline{\frac{1}{3}}$

c) $P(p'e')=P(p\cup e)'=1-P(p\cup e)=1-\frac{31}{36}=\underline{\frac{5}{36}}$

2._ Se construye un dado de manera que el 1 y el 2 ocurran con el doble de la frecuencia que se presenta el 5, el cual ocurre con una frecuencia 3 veces superior al 3, al 4 o al 6. Si se lanza una vez, encuentra la probabilidad de que:

a) El número sea par.
b) El número sea 4.
c) El número sea mayor que 4.

Solución

Si a es el número de veces que cae el número 5 tendríamos que la ocurrencia de cada cara del dado está dada por

1	2	3	4	5	6
$2a$	$2a$	$a/3$	$a/3$	a	$a/3$

Al multiplicar por 3 para evitar los quebrados, tenemos que el número total de lanzamientos es $6a+6a+a+a+3a+a=18a$.

a) $P(par)=P(2)+P(4)+P(6)=\frac{6a+a+a}{18a}=\frac{8a}{18a}=\underline{\frac{4}{9}}$

b) $P(4)=\dfrac{a}{18a}=\underline{\dfrac{1}{18}}$

c) $P(5\text{ o }6)=P(5)+P(6)=\dfrac{3a}{18a}+\dfrac{a}{18a}=\underline{\dfrac{2}{9}}$

Observa que también pudimos razonar de la manera siguiente:

Suponemos que la probabilidad de ocurrencia del 3, del 4 y del 6 es p, entonces la probabilidad de ocurrencia del 5 es $3p$ y por lo tanto el 1 y el 2 tienen probabilidad de ocurrencia de $6p$.
Sumando la probabilidad de ocurrencia del uno al seis se obtiene la unidad, es decir

$$\begin{aligned}1&=P(1)+P(2)+P(3)+P(4)+P(5)+P(6)\\&=6p+6p+p+p+3p+p=18p\end{aligned}$$

Entonces $p=\dfrac{1}{18}$. Así por ejemplo para resolver el inciso a) tenemos que

$$P(par)=P(2)+P(4)+P(6)=6p+p+p=8p=8\left(\frac{1}{18}\right)=\frac{4}{9}$$

3._ Se está realizando la inspección final de aparatos de televisión después del ensamble. Se identifican tres tipos de defectos como críticos, mayores y menores y una empresa de envíos por correo los clasifica en C, M o m respectivamente. Se analizan los datos con los siguientes resultados:
Aparatos que sólo tienen defectos críticos 2%.
Aparatos que sólo tienen defectos mayores 5%.
Aparatos que sólo tienen defectos menores 7%.
Aparatos que sólo tienen defectos críticos y mayores 3%.
Aparatos que sólo tienen defectos críticos y menores 4%.
Aparatos que sólo tienen defectos mayores y menores 3%.
Aparatos que tienen los tres tipos de defectos 1%.

a) ¿Qué porcentaje de los aparatos no tienen defectos?
Los aparatos con defectos críticos o mayores (o ambos) deben manufacturarse nuevamente.
b) ¿Qué porcentaje corresponde a esta categoría?

<u>Solución</u>

Hacer un diagrama de Venn y poner mucho cuidado con la palabra “sólo”.
a) <u>75% no tienen defectos</u>.
b) <u>18% de los aparatos deben manufacturarse nuevamente</u>.

4._ Considera los eventos A y B tales que $P(A)=\frac{1}{3}$ y $P(B)=\frac{1}{2}$. Determina el valor de $P(BA')$ para cada una de las siguientes condiciones: a) A y B son disjuntos, b) A está contenido en B y c) $P(AB)=\frac{1}{8}$.

Solución

a) Si los eventos son disjuntos $BA'=B$ entonces $P(BA')=P(B)=\underline{\frac{1}{2}}$

b) Si $A\subset B$ entonces $P(BA')=P(B)-P(A)=\frac{1}{2}-\frac{1}{3}=\frac{3-2}{6}=\underline{\frac{1}{6}}$

c) Si los conjuntos se intersectan, entonces $P(BA')=P(B)-P(AB)=\frac{1}{2}-\frac{1}{8}=\frac{4-1}{8}=\underline{\frac{3}{8}}$

5._ Cuatro pelotas con los números 1, 2, 3 y 4 se colocan en una bolsa, se mezclan y se sacan una por una ¿Cuál es la probabilidad de que salgan justo en ese orden?

Solución

Primero necesito contar de cuantas manera pueden salir las 4 bolitas, como están numeradas esto ocurre de $4!=24$ maneras de las cuales sólo una de esas maneras sigue el orden deseado, por lo tanto

$$P(1,2,3,4)=\underline{\frac{1}{24}}$$

6._ En una montaña hay 5 rutas para subir y 5 rutas para bajar de la cima. ¿Cuál es la probabilidad de que dos conocidos se encuentren si uno sube y el otro baja?

Solución

El que sube puede hacerlo por cualquiera de las 5 rutas pero por cada una de estas, el que baja tiene 5 opciones también, por lo tanto el número total de maneras en que uno puede subir y el otro bajar es 25, de las cuales puede que ambos coincidan en la ruta 1, en la ruta 2, en la ruta 3, en la ruta 4 o bien en la ruta 5, por lo tanto:

$$P(coincidan)=\frac{5}{25}=\underline{\frac{1}{5}}$$

7._ Un ingeniero de una fábrica de microcircuitos inspeccionará un lote de obleas de silicio para tratar de encontrarles defectos. Supón que hay cuatro circuitos integrados defectuosos en un recipiente que contiene 20 obleas. Para esa inspección se seleccionan 2 obleas al azar. Calcula la probabilidad de que:
a) ninguna de ellas tenga defectos.
b) Por lo menos una de las 2 no tenga defectos.

Solución

a) Hay 4 defectuosas y 16 buenas de un total de 20 obleas.

El número total en que se pueden seleccionar 2 de las 20 obleas está dado por

$$\binom{20}{2}=\frac{20!}{2!18!}=\frac{18!(19)(20)}{2!18!}=19(10)=190$$

Queremos que ninguna tenga defecto, por lo tanto queremos contar las maneras en que podemos seleccionar 2 de las 16 obleas buenas

$$\binom{16}{2}=\frac{16!}{2!14!}=\frac{14!(15)(16)}{2!14!}=15(8)=120$$, entonces.

$$P(2\text{ buenas})=\frac{120}{190}=\frac{12}{19}=\underline{0.6316}$$

b) Por lo menos una de las 2 sin defecto, significa 1 buena o 2 buenas, ya tenemos la probabilidad de tener 2 buenas, ahora necesitamos la probabilidad de 1 buena y luego sumamos ya que estos eventos son disjuntos.

¿De cuántas maneras puedo tener 1 buena y 1 mala?

$$\binom{16}{1}\binom{4}{1}=16(4)=64$$

$$P(1\text{ buena})=\frac{64}{190}=\frac{32}{95}$$

Entonces $P(\text{almenos}\,1\text{ buenas})=P(1\text{ buena})+P(2\text{ buenas})=\frac{32}{95}+\frac{12}{19}=\frac{32+60}{95}=\frac{92}{95}=\underline{0.9684}$

1.4 Probabilidad Condicional.

La probabilidad de que ocurra un evento B cuando se sabe que ya ocurrió el evento A se denomina probabilidad condicional y se denota por $P(B\mid A)$.

Definición: La **probabilidad condicional de** B **dado** A está dada por $\boxed{P(B\mid A)=\frac{P(BA)}{P(A)}}$ con $P(A)>0$.

Observación:

1) Si $P(A)=0$ entonces la probabilidad condicional no está definida.
2) La probabilidad condicional cumple las propiedades que hemos visto para la probabilidad de un evento pero ahora el espacio muestral que antes era S se ha reducido al evento A. Así por ejemplo se tiene que:

$$\text{i) } P(B \mid A)+P(B' \mid A)=1$$

$$\text{ii) } P(A\cup B|C)=P(A|C)+P(B|C)-P(AB|C)$$

Demostración:

i) $$P(B|A)+P(B'|A)=\frac{P(BA)}{P(A)}+\frac{P(B'A)}{P(A)}=\frac{P(BA)+P(B'A)}{P(A)}=\frac{P(BA\cup B'A)}{P(A)}=\frac{P(A)}{P(A)}=1 \quad \blacklozenge$$

ii) $$P(A\cup B|C)=\frac{P[(A\cup B)C]}{P(C)}=\frac{P(AC\cup BC)}{P(C)}=\frac{P(AC)+P(BC)-P(ABC)}{P(C)}$$

$$=\frac{P(AC)}{P(C)}+\frac{P(BC)}{P(C)}-\frac{P(ABC)}{P(C)}=P(A|C)+P(B|C)-P(AB|C)$$

Ejemplos:

1._ S es la población de egresados de la primera generación de la ESCOM, se clasifican según su condición de sexo y estado de ocupación como se muestra en la tabla siguiente.

	Empleado	Desempleado	Total
Masculino	460	40	500
Femenino	140	260	400
Total	600	300	900

Se va a invitar a uno de estos egresados aleatoriamente para que dé un seminario sobre las oportunidades que los ingenieros egresados de la ESCOM tienen en el campo laboral.
a) ¿Cuál es la probabilidad de que el invitado sea hombre si se sabe que tiene empleo?
b) Dado que la elegida es una mujer ¿Cuál es la probabilidad de que esté desempleada?
c) ¿Cuál es la probabilidad de que el invitado no tenga empleo si se sabe que es un hombre?

Solución

Sean los eventos H: el elegido es un hombre, M: el elegido es una mujer, E: el elegido tiene empleo y D: el elegido está desempleado.
Observa que hay dos maneras de resolver el problema:

i) Criterio de espacio reducido._ Consiste en encontrar la probabilidad del evento de interés en base al espacio muestral reducido o condición.

a) El espacio muestral reducido se compone únicamente de los egresados empleados. De los cuales me interesan los hombres, en este espacio muestral reducido hay 460 hombres, por lo tanto,

$$P(H \mid E) = \frac{460}{600} = \frac{23}{30} = \underline{0.766}$$

ii) Definición de probabilidad condicional._ En base al espacio muestral original compuesto por todos los egresados de la primera generación de la ESCOM.

$$P(H \mid E) = \frac{P(HE)}{P(E)} = \frac{460/900}{600/900} = \frac{460}{600} = \frac{23}{30} = \underline{0.766}$$

b) $P(D \mid M) = \dfrac{P(DM)}{P(M)} = \dfrac{260/900}{400/900} = \dfrac{260}{400} = \dfrac{13}{20} = \underline{0.65}$

c) $P(D \mid H) = \dfrac{P(DH)}{P(H)} = \dfrac{40/900}{500/900} = \dfrac{40}{500} = \dfrac{2}{25} = \underline{0.08}$

2._ La población de la ciudad X es 40% masculina y 60% femenina. Si el 50% de los hombres fuman y 30% de las mujeres fuman. ¿Cuál es la probabilidad de que un fumador sea hombre?

<u>Solución</u>

Sean los eventos *M*: masculino, *F*: femenino y *s*: fumador

Datos: $P(M) = 0.4$, $P(F) = 0.6$ $P(s \mid M) = 0.5$ $P(s \mid F) = 0.3$

Observa que:

i) $P(s \mid M) = \dfrac{P(sM)}{P(M)} \therefore P(sM) = P(s \mid M)P(M)$

ii) $P(s \mid F) = \dfrac{P(sF)}{P(F)} \therefore P(sF) = P(s \mid F)P(F)$

iii) $s = sM \cup sF$ por ser disjuntos tenemos que

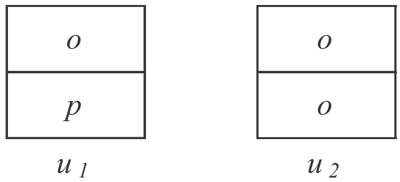
$P(s) = P(sM) + P(sF) = P(s \mid M)P(M) + P(s \mid F)P(F)$

La probabilidad que me piden es:

$$P(M \mid s) = \frac{P(Ms)}{P(s)} = \frac{P(s \mid M)P(M)}{P(s \mid M)P(M) + P(s \mid F)P(F)} = \frac{0.5(0.4)}{0.5(0.4) + 0.3(0.6)} = \underline{0.526}$$

3._ Se tienen 2 urnas cada una con 2 cajones. La urna 1 tiene una moneda de oro en un cajón y una de plata en el otro, mientras que la urna 2 tiene una moneda de oro en cada uno de los cajones. Se escoge una urna al azar y de esta se escoge un cajón al azar, se saca la moneda y se observa que es de oro ¿Cuál es la probabilidad de que la moneda provenga de la urna 2?

Solución



Sean los eventos o: la moneda es de oro, u_1: urna 1 y u_2:urna 2.

Datos:

$$P(u_1) = P(u_2) = \frac{1}{2}$$
$$P(o \mid u_1) = \frac{1}{2} \text{ y } P(o \mid u_2) = 1$$

Observa que el evento "la moneda es de oro" se puede ver como la unión de los eventos "es de oro de la urna 1" y "es de oro de la urna 2", lo cual representamos mediante la expresión:

$$o = ou_1 \cup ou_2 \quad \text{es decir} \quad P(o) = P(ou_1) + P(ou_2)$$
$$= P(o \mid u_1)P(u_1) + P(o \mid u_2)P(u_2)$$

Utilizamos el hecho de que $P(o \mid u_2) = \dfrac{P(ou_2)}{P(u_2)}$ despejando $P(ou_2) = P(o \mid u_2)P(u_2)$

Entonces

$$P(u_2 \mid o) = \frac{P(u_2 o)}{P(o)} = \frac{P(o \mid u_2)P(u_2)}{P(o \mid u_1)P(u_1) + P(o \mid u_2)P(u_2)}$$

$$= \frac{1\left(\frac{1}{2}\right)}{\frac{1}{2}\left(1 + \frac{1}{2}\right)} = \underline{\frac{2}{3}}$$

4._ A un sospechoso se le aplica un suero de la verdad que se sabe es confiable en 90% cuando la persona es culpable y en 99% cuando la persona es inocente. En otras palabras, el 10% de los culpables se considera inocente cuando se usa el suero y el 1% de los inocentes se juzga culpable. Si el sospechoso se escogió de un grupo del cual sólo 5% ha cometido alguna vez un crimen y el suero indica que la persona es culpable. ¿Cuál es la probabilidad de que sea inocente?

Solución

Sean los eventos c: culpable, c': inocente,
s: se dice inocente tomando el suero s': se dice culpable tomando el suero.

Datos: $P(s \mid c) = 0.1 \quad \therefore P(s' \mid c) = 0.9$

$P(s' \mid c') = 0.01$

$P(c) = 0.05 \quad \therefore P(c') = 0.95$

$$P(c' \mid s') = \frac{P(c's')}{P(s')} = \frac{P(s' \mid c')P(c')}{P(s' \mid c')P(c') + P(s' \mid c)P(c)} = \frac{0.01(0.95)}{0.01(0.95) + 0.9(0.05)} = \underline{0.174}$$

5._ Se conduce una investigación detallada de accidentes aéreos. La probabilidad de que un accidente por falla estructural se identifique correctamente es 0.9 y la probabilidad de que un accidente que no se debe a una falla estructural se identifique en forma incorrecta como un accidente por falla estructural es 0.2. Si el 25% de los accidentes aéreos se debe a fallas estructurales, determina la probabilidad de que un accidente aéreo que se ha identificado como debido a una falla estructural en realidad haya ocurrido por una falla de este tipo.

Solución

Sean los eventos *e*: accidente por falla estructural y *E*: se identifica como falla estructural.
Datos:

$P(E \mid e) = 0.9$

$P(E \mid e') = 0.2$

$P(e) = 0.25 \quad \therefore P(e') = 0.75$

Entonces

$$P(e \mid E) = \frac{P(eE)}{P(E)} = \frac{P(E \mid e)P(e)}{P(E \mid e)P(e) + P(E \mid e')P(e')} = \frac{0.9(0.25)}{0.9(0.25) + 0.2(0.75)} = \underline{0.6}$$

6._ Dos bolas se extraen de una urna que contiene *m* bolas numeradas de 1 a *m*. Si la primera bola tiene el número 1 se conserva y se regresa en caso contrario. ¿Cuál es la probabilidad de que la segunda bola extraída tenga el número 2?

Solución

Observa que la segunda bola puede tener el número 2 de dos maneras diferentes y excluyentes, puede que la primera tenga el número 1 y la segunda el número 2 o bien puede ser que la primera tenga un número diferente de 1 y la segunda bola tenga el número 2.

Sean los eventos:

B_1 : la primera bola tiene el número 1.

B_1' : la primera bola tiene un número diferente de 1.

b_2 :: la segunda bola tiene el número 2.

El evento de interés puede verse como la unión de dos eventos disjuntos;

$$
\begin{aligned}
b_2 &= B_1 b_2 \cup B_1' b_2 \\
P(b_2) &= P(B_1 b_2 \cup B_1' b_2) \\
&= P(B_1 b_2) + P(B_1' b_2) \\
&= P(B_1)P(b_2|B_1) + P(B_1')P(b_2|B_1') \\
&= \frac{1}{m}\frac{1}{m-1} + \frac{m-1}{m}\frac{1}{m} \\
&= \frac{1}{m(m-1)} + \frac{m-1}{m^2} \\
&= \frac{m+(m-1)^2}{m^2(m-1)} \\
&= \frac{m+m^2-2m+1}{m^2(m-1)} \qquad \therefore\ \underline{P(b_2) = \frac{m^2-m+1}{m^2(m-1)}}
\end{aligned}
$$

7._ Muestra que si A, B y C son tres eventos tales que $P(ABC) \neq 0$ y $P(C \mid AB) = P(C \mid B)$ entonces $P(A \mid BC) = P(A \mid B)$.

Solución

Se tienen que $P(C \mid AB) = P(C \mid B)$ usando la definición de probabilidad condicional

$$P(C \mid AB) = \frac{P(ABC)}{P(AB)} = \frac{P(BC)}{P(B)} = P(C \mid B)$$

$$\frac{P(ABC)}{P(BC)} = \frac{P(AB)}{P(B)}$$

$$P(A \mid BC) = P(A \mid B)$$

1.5 Eventos independientes y Regla de la multiplicación.

Ejemplos:

1._ Considera una caja con 4 bolitas etiquetadas del 1 al 4. Se saca una de ellas al azar y se observa el número de la etiqueta. Con los eventos A y B definidos a continuación.

$S = \{1, 2, 3, 4\}$

$A = \{1, 2\} \qquad P(A) = \frac{1}{2}$

$B = \{1, 3\} \qquad P(B) = \frac{1}{2}$

$$AB = \{1\} \qquad P(AB) = \frac{1}{4}$$

Entonces $\quad P(B \mid A) = \dfrac{P(BA)}{P(A)} = \dfrac{\frac{1}{4}}{\frac{1}{2}} = \underline{\dfrac{1}{2}}$

Observa que $\quad P(B) = P(B \mid A)$

2._ De los viajeros que llegan a un aeropuerto pequeño, 60% utilizan aerolíneas importantes, 30% viaja mediante aviones privados, y el resto usa aviones comerciales que no pertenecen a las aerolíneas importantes. De las personas que usan aerolíneas importantes 50% viaja por negocios, mientras que 60% de los pasajeros de los aviones privados y 90% de los que usan otras aeronaves comerciales también viaja por negocios. Supón que se selecciona al azar una persona que llega a este aeropuerto. ¿Cuál es la probabilidad de que la persona
a) viaje por negocios?
b) Viaje por negocios en avión privado?
c) Haya viajado en avión privado dado que lo hace por negocios?

Solución

Sean los eventos
I: viaja en aerolínea importante.
p: viaja en avión privado.
c: viaja en aerolínea comercial.
n: viaja por negocios.

Datos:

$$\begin{array}{ll} P(I) = 0.6 & P(n \mid I) = 0.5 \\ P(p) = 0.3 & P(n \mid p) = 0.6 \\ P(c) = 0.1 & P(n \mid c) = 0.9 \end{array}$$

a) La persona puede viajar por negocios en cualquiera de las tres aerolíneas, y no pude viajar en dos o más aerolíneas a la vez por lo tanto son eventos disjuntos y se tiene que:

$$\begin{aligned} P(n) &= P(nI) + P(np) + P(nc) \\ &= P(n \mid I)P(I) + P(n \mid p)P(p) + P(n \mid c)P(c) \\ &= 0.5(0.6) + 0.6(0.3) + 0.9(0.1) = \underline{0.57} \end{aligned}$$

Observa que $P(n) \neq P(n \mid I)$, $P(n) \neq P(n \mid p)$ y $P(n) \neq P(n \mid c)$

b) $P(np) = P(n \mid p)P(p) = 0.6(0.3) = \underline{0.18}$

c) $P(p \mid n) = \dfrac{P(pn)}{P(n)} = \dfrac{0.18}{0.57} = \dfrac{6}{19} = \underline{0.3158}$

Observa también que $P(p \mid n) \neq P(p)$.

Con los dos ejemplos previos hemos visto que hay casos en donde la ocurrencia de un evento no altera la ocurrencia del otro, como en el ejemplo 1, la probabilidad incondicional (en base al espacio muestral original *S*) y la probabilidad condicional (en base al espacio muestral reducido) son iguales [$P(B) = P(B \mid S) = P(B \mid A)$], si ocurre o no ocurre *A*, la probabilidad de *B* no se altera, es decir *B* no depende de *A,* cuando esto ocurre se dice que A y B son independientes.
Si en cambio la ocurrencia de uno altera la ocurrencia del otro, podemos pensar que hay una dependencia entre ellos como ocurre en el ejemplo 2, y en este caso la probabilidad incondicional y la probabilidad condicional son diferentes.

Definición: Los eventos *A* y *B* son **independientes** si $\boxed{P(B) = P(B \mid A)}$ y se escribe $A \perp B$.

Es decir, la ocurrencia de *A* no influye en la ocurrencia de *B*, pase o no pase *A* la probabilidad de *B* no cambia.

<u>Observación</u>: Si $A \perp B$ sabemos que $P(B) = P(B \mid A) = \dfrac{P(BA)}{P(A)} \therefore \boxed{P(BA) = P(B)P(A)}$ es decir, si *A* y *B* son eventos independientes entonces la probabilidad de la intersección de ellos es el producto de las probabilidades individuales.
Generalmente pensamos que si dos eventos son disjuntos o excluyentes entonces son independientes, pero esto no es verdad, considera el siguiente

Ejemplo:

3._ $S = \{1, 2, 3, 4\}$

$A = \{1, 2\} \quad \therefore \; P(A) = \dfrac{1}{2}$

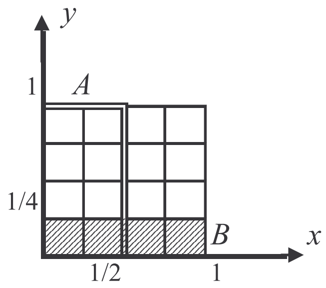
$B = \{3,4\} \quad \therefore \; P(B) = \dfrac{1}{2}$

$AB = \varnothing \;\therefore\; P(AB) = 0 \neq \dfrac{1}{4} = P(A)P(B)$ Los eventos *A* y *B* son disjuntos, pero no independientes.

Nota._ Se dice que los eventos *A*, *B* y *C* son mutuamente independientes si se cumplen todas las condiciones siguientes:
i) $P(AB) = P(A)P(B)$
ii) $P(AC) = P(A)P(C)$
iii) $P(BC) = P(B)P(C)$
iv) $P(ABC) = P(A)P(B)P(C)$

4._ Sea *S* el cuadrado $0 \le x \le 1,\; 0 \le y \le 1,$ en el plano. Considera el espacio de probabilidad uniforme sobre el cuadrado y sean *A* y *B* los eventos:
$A = \{(x, y) \mid 0 \le x \le 0.5,\; 0 \le y \le 1\}$
$B = \{(x, y) \mid 0 \le x \le 1,\; 0 \le y \le 0.25\}$ ¿Son independientes *A* y *B*?

Solución



$AB = \{(x, y) \mid 0 \le x \le 0.5,\ \ 0 \le y \le 0.25\}$

$P(AB) = \frac{1}{8} = \left(\frac{1}{2}\right)\left(\frac{1}{4}\right) = P(A)P(B) \ \therefore\ \underline{A \perp B}$

5._ Si $A \perp B$ entonces ¿$A' \perp B'$?

Solución

$A \perp B$ entonces $P(AB) = P(A)P(B)$

Por las Leyes de D'Morgan se tiene que

$$\begin{aligned} P(A'B') &= P(A \cup B)' \\ &= 1 - P(A \cup B) = 1 - [P(A) + P(B) - P(AB)] = 1 - P(A) - P(B) + P(AB) \\ &= [1 - P(A)] - P(B) + P(A)P(B) = P(A') - P(B)[1 - P(A)] \\ &= P(A') - P(B)P(A') = P(A')[1 - P(B)] = P(A')P(B') \ \therefore\ \underline{A' \perp B'} \end{aligned}$$

6._ La probabilidad de que Tomás esté vivo dentro de 20 años es de 0.7 y la probabilidad de que Nancy esté viva dentro de 20 años es 0.9. Si se asume independencia entre ambos eventos ¿Cuál es la probabilidad de que ninguno de los dos esté vivo dentro de 20 años?

Solución

Sean los eventos *T*: Tomás viva dentro de 20 años y *N*: Nancy viva dentro de 20 años.

$P(T) = 0.7 \quad P(T') = 1 - P(T) = 0.3 \quad P(N) = 0.9 \quad P(N') = 1 - P(N) = 0.1$

Se sabe que $T \perp N$, por lo tanto los complementos también son independientes, y entonces

$$P(T'N') = P(T')P(N') = (0.3)(0.1) = \underline{0.03}$$

7._ Si $A \perp B$ y la probabilidad de que *A* o *B* ocurran es 0.6, mientras que la probabilidad de que *A* ocurra es 0.4. ¿Cuál es la probabilidad de que ocurra *B*?

Solución

Datos: $A \perp B$, $P(A \cup B) = 0.6$, $P(A) = 0.4$ $P(A') = 0.6$

Sabemos que $P(A \cup B) = P(A) + P(B) - P(AB)$

$$= P(A) + P(B)[1 - P(A)] = P(A) + P(B)P(A')$$

Entonces $P(B) = \dfrac{P(A \cup B) - P(A)}{P(A')} = \dfrac{0.6 - 0.4}{0.6} = \dfrac{0.2}{0.6} = \underline{0.33}$

8._ Supón que A, B y C son eventos mutuamente independientes y que $P(AB) \neq 0$. Muestra que $P(C \mid AB) = P(C)$

<u>Solución</u>

$$P(C \mid AB) = \frac{P(ABC)}{P(AB)} = \frac{P(AB)P(C)}{P(AB)} = P(C)$$

9._ Supón que se lanzan 3 monedas idénticas y perfectamente balanceadas una sola vez. Sea A_i el evento de que la *i*-ésima moneda cae en cara. Muestra que los eventos A_1, A_2, A_3 son mutuamente independientes.

<u>Solución</u>

El espacio muestral del experimento es

S ={xxx, xcc, cxc, ccx, xxc, xcx, cxx, ccc}

A_1 ={cxc, ccx, cxx, ccc} $\quad P(A_1) = \dfrac{1}{2}$

A_2 ={xcc, ccx, xcx, ccc} $\quad P(A_2) = \dfrac{1}{2}$

A_3 ={xcc, cxc, xxc, ccc} $\quad P(A_3) = \dfrac{1}{2}$

$A_1\ A_2$ ={ccx, ccc} $\quad P(A_1A_2) = \dfrac{1}{4}$

$A_1\ A_3$ ={cxc, ccc} $\quad P(A_1A_3) = \dfrac{1}{4}$

$A_2\ A_3$ ={xcc, ccc} $\quad P(A_2A_3) = \dfrac{1}{4}$

$A_1\ A_2\ A_3$ ={ccc} $\quad P(A_1A_2A_3) = \dfrac{1}{8}$

Se tiene que $P(A_iA_j) = P(A_i)P(A_j) \quad \forall \quad i \neq j = 1, 2, 3.$

y $P(A_1A_2A_3) = \prod_{i=1}^{3} P(A_i)$ entonces <u>los tres eventos son mutuamente independientes</u>.

Usando la definición de probabilidad condicional se tiene que

$$P(A \mid B)=\frac{P(AB)}{P(B)} \quad \therefore\; P(AB)=P(A \mid B)P(B) \quad \text{y también}$$

$$P(B \mid A)=\frac{P(AB)}{P(A)} \quad \therefore\; P(AB)=P(B \mid A)P(A)$$

esto se conoce como Teorema o Ley de la multiplicación de probabilidades y se puede leer como, *la probabilidad de la intersección de dos eventos es igual a la probabilidad condicional de uno de los eventos dado el otro por la probabilidad de la condición. O bien; la probabilidad de la intersección de dos eventos es igual a la probabilidad de uno de los eventos por la probabilidad condicional del otro evento dado el que ya tomamos.*

Teorema de la multiplicación $\boxed{P(AB)=P(A \mid B)P(B)=P(B \mid A)P(A)}$

Este teorema se puede generalizar de la manera siguiente

$$\boxed{P(A_1A_2\cap\cdots\cap A_n)=P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1A_2)\cdots P(A_n \mid A_1A_2\cap\cdots\cap A_{n-1})}$$

<u>**Demostración:**</u>

$$P(A_1A_2\cap\cdots\cap A_n)=P(A_1)\frac{P(A_1A_2)}{P(A_1)}\frac{P(A_1A_2A_3)}{P(A_1A_2)}\cdots\frac{P(A_1A_2\cap\cdots A_{n-1})}{P(A_1A_2\cap\cdots A_{n-2})}\frac{P(A_1A_2\cap\cdots\cap A_n)}{P(A_1A_2\cap\cdots\cap A_{n-1})} \qquad \blacklozenge$$

Ejemplo:

10._ La probabilidad de que le apliquen rayos X a una persona que llega a su dentista es de 0.6, la probabilidad de que una persona a la que se le han aplicado rayos X tenga también caries es de 0.3, y la probabilidad de que a alguna persona a la que se le han aplicado rayos X y a la que se le ha curado una caries, se le extraiga también un diente es de 0.1. ¿Cuál es la probabilidad de que a una persona que visite a su dentista se le apliquen rayos X, se le cure una caries y se le extraiga un diente?

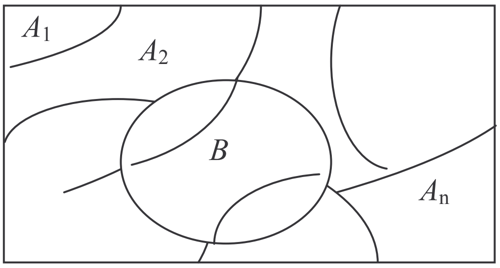
<u>**Solución**</u>

Sean los eventos *X*: se le aplican rayos X, *C*: tiene caries y *E*: se le extrae un diente.
Datos: $P(X)=0.6$
$P(C \mid X)=0.3$
$P(E \mid XC)=0.1$

Entonces $P(XCE)=P(X)P(C \mid X)P(E \mid XC)=0.6(0.3)(0.1)=\underline{0.018}$

1.6 TEOREMA DE BAYES.

Definición: Si $A_1, A_2, \ldots, A_n$ son eventos mutuamente excluyentes, es decir $A_i A_j = \varnothing \quad \forall\ i \neq j$, tales que $S = \bigcup_{i=1}^{n} A_i$ se dice que $\{A_1, A_2, \ldots, A_n\}$ es una partición de *S*.



Supón que tenemos una partición $\{A_1, A_2, \ldots, A_n\}$ de *S*, un evento posible B $(P(B) > 0)$ y que se conocen las probabilidades $P(B \mid A_i)$ y $P(A_i)$ para *i* = 1, 2, ... , *n*. ¿Cómo encontramos $P(B)$?

Observa que $B = BS = B\left[\bigcup_{i=1}^{n} A_i\right] = \bigcup_{i=1}^{n} BA_i$

Pero $BA_i \cap BA_j = \varnothing \quad \forall \quad i \neq j = 1, 2, ..., n$, por tratarse de una partición de *S*. Entonces:

$$P(B) = P\left[\bigcup_{i=1}^{n} BA_i\right] = \sum_{i=1}^{n} P(BA_i)$$

$$\boxed{P(B) = \sum_{i=1}^{n} P(B \mid A_i) P(A_i)}$$

Este resultado se conoce como **probabilidad total** de B.

Si ahora tenemos interés en $P(A_i \mid B)$ procedemos usando la definición de probabilidad condicional, la Ley de la multiplicación y la probabilidad total. Así;

$$P(A_i \mid B) = \frac{P(A_i B)}{P(B)}$$

$$\boxed{P(A_i \mid B) = \frac{P(B \mid A_i) P(A_i)}{\sum_{i=1}^{n} P(B \mid A_i) P(A_i)}}$$

Este resultado se conoce como **Teorema de Bayes**.

Ejemplos:

1._ Cuando los artículos llegan al final de una línea de producción, un supervisor escoge los que deben pasar por una inspección completa, 10% de todos los artículos son defectuosos, 60% de todos los artículos defectuosos y 20% de todos los artículos buenos pasan por una inspección completa. ¿Cuál es la probabilidad de que un artículo
a) haya pasado por la inspección completa?
b) sea defectuoso dado que pasó por una inspección completa?

Solución

Sean los eventos D: artículo defectuoso.
I: inspección completa.
Datos: $P(D) = 0.1$ $\quad P(D') = 0.9$
$P(I \mid D) = 0.6$
$P(I \mid D') = 0.2$

a) La probabilidad total de evento I está dada por:

$$\boxed{P(B) = \sum_{i=1}^{n} P(B \mid A_i)P(A_i)}$$

Observa que la partición en este problema se compone de dos eventos, que son el artículo es defectuoso o no defectuoso, entonces $n = 2$ y $A_1 = D$, $A_2 = D'$ y el evento de interés es que haya pasado por una inspección completa, es decir $B = I$. Sustituyendo en la ecuación para la probabilidad total del evento de interés tenemos que:

$$P(I) = \sum_{i=1}^{2} P(I \mid A_i)P(A_i) = P(I \mid D)P(D) + P(I \mid D')P(D') = 0.6(0.1) + 0.2(0.9) = \underline{0.24}$$

b) Usando el Teorema de Bayes

$$P(D \mid I) = \frac{P(I \mid D)P(D)}{P(I \mid D)P(D) + P(I \mid D')P(D')} = \frac{0.6(0.1)}{0.24} = \underline{0.25}$$

2._ En cierta región del país se sabe por experiencias pasadas que la probabilidad de elegir a un adulto de más de 40 años con cáncer es de 0.02. Si la probabilidad de que un doctor diagnostique en forma correcta que una persona tiene cáncer es de 0.78 y la probabilidad de que diagnostique en forma incorrecta que una persona tiene cáncer si no lo tiene es de 0.06.
a) ¿Cuál es la probabilidad de que diagnostique que una persona tiene cáncer?
b) Si el diagnóstico del doctor es que el paciente tiene cáncer. ¿Cuál es la probabilidad de que en efecto tenga la enfermedad?
c) Y ¿Cuál la probabilidad de que no la tenga?

Solución

Sean los eventos C: tiene cáncer y D: diagnostico de cáncer.

Datos: $P(C) = 0.02 \quad P(C') = 0.98$
$P(D \mid C) = 0.78$
$P(D \mid C') = 0.06$

a) Probabilidad total del evento D.

$$P(D) = P(D \mid C)P(C) + P(D \mid C')P(C') = 0.78(0.02) + 0.06(0.98) = \underline{0.0744}$$

b) $P(C \mid D) = \dfrac{P(D \mid C)P(C)}{P(D)} = \dfrac{0.78(0.02)}{0.0744} = \underline{0.2096}$

c) $P(C' \mid D) = \dfrac{P(D \mid C')P(C')}{P(D)} = \dfrac{0.06(0.98)}{0.0744} = \underline{0.79}$

También $P(C' \mid D) = 1 - P(C \mid D) =$ 1-0.2096 = $\underline{0.79}$

3._ Cinco urnas llevan los números 1, 2, 3, 4 y 5 respectivamente. La urna i contiene i bolas blancas y $(5-i)$ bolas negras, con i = 1,...,5. Se selecciona al azar una urna y después se sacan sin reposición dos bolas de dicha urna.
a) ¿Cuál es la probabilidad de que ambas bolas seleccionadas sean blancas?
b) Dado que ambas bolas seleccionadas son blancas ¿Cuál es la probabilidad de que se haya seleccionado la urna 3?

Solución

Sean los eventos:
u_i : urna i, con $i = 1, 2, \ldots, 5$.
b_1 : la primera bola es blanca
b_2 : la segunda bola es blanca.

1 blanca	2 blancas	3 blancas	4 blancas	5 blancas
4 negras	3 negras	2 negras	1 negra	0 negras
urna 1	urna 2	urna 3	urna 4	urna 5

a) Dado que la urna 1 sólo tiene una bola blanca las dos bolas pueden ser blancas sólo si provienen de cualquiera de las otras cuatro urnas, es decir, la probabilidad total del evento *dos bolas blancas* está dada por

$$b_1b_2 = b_1b_2u_2 + b_1b_2u_3 + b_1b_2u_4 + b_1b_2u_5$$

$$\therefore \quad P(b_1b_2) = \sum_{i=2}^{5} P(b_1b_2u_i) = \sum_{i=2}^{5} P(u_i)P(b_1|u_i)P(b_2|u_ib_1)$$

$$P(u_i) = \frac{1}{5}$$

$$P(b_1|\; u_2) = \frac{2}{5} \qquad P(b_2|\; u_2b_1) = \frac{1}{4}$$

$$P(b_1|\; u_3) = \frac{3}{5} \qquad P(b_2|\; u_3b_1) = \frac{2}{4}$$

$$P(b_1|\; u_4) = \frac{4}{5} \qquad P(b_2|\; u_4b_1) = \frac{3}{4}$$

$$P(b_1|\; u_5) = 1 \qquad P(b_2|\; u_3b_1) = 1$$

$$P(b_1b_2) = \frac{1}{5}\left(\frac{2}{5}\frac{1}{4} + \frac{3}{5}\frac{2}{4} + \frac{4}{5}\frac{3}{4} + 1\right) = \frac{1}{20}(2+6+12+20) = \frac{40}{5(20)} = \underline{\frac{2}{5}}$$

b) $P(u_3|b_1b_2) = \dfrac{P(b_1b_2|u_3)P(u_3)}{P(b_1b_2)} = \underline{\dfrac{3}{20}}$

2 VARIABLES ALEATORIAS DISCRETAS.

OBJETIVO: Caracterizar las distribuciones de probabilidad de las variables aleatorias discretas más comunes.

2.1 VARIABLES ALEATORIAS DISCRETAS.

Considera el experimento de lanzar una moneda 3 veces, la probabilidad de que caiga cara en un lanzamiento es p. Supón que por cada cara ganas 1 peso y por cada cruz pierdes 1 peso, si X es la cantidad total ganada al cabo de los tres lanzamientos. Para cada evento simple (E_i) del espacio muestral se tiene un valor de X y un valor para su probabilidad como se ilustra en el siguiente cuadro:

i	E_i	X	$P(E_i)$
1	ccc	3	p^3
2	ccx	1	$p^2(1-p)$
3	cxc	1	$p^2(1-p)$
4	xcc	1	$p^2(1-p)$
5	cxx	-1	$p(1-p)^2$
6	xcx	-1	$p(1-p)^2$
7	xxc	-1	$p(1-p)^2$
8	xxx	-3	$(1-p)^3$

Cada lanzamiento es independiente de los demás por lo que $P(ccc)=\prod_{i=1}^{3} P(c)=p^3$.

Además $\sum_{i=1}^{8} P(E_i)=p^3+3\left[p^2(1-p)\right]+3p(1-p)^2+(1-p)^3=(p+1-p)^3=1$.

Si A es el evento de que la ganancia es 1 peso, se tiene que $A=\{E_i \mid X(E_i)=1\}=\{ccx,cxc,xcc\}$
Entonces $P(A)=3p^2(1-p)$.

A menudo se hace referencia a este hecho diciendo que $X=1$ tiene probabilidad $3p^2(1-p)$ o bien que $P(X=1)=3p^2(1-p)$, de igual manera se dice que:
$P(X=-1)=3p(1-p)^2$, $P(X=3)=p^3$, $P(X=-3)=(1-p)^3$ y $P(X=2)=0$.

Definición: Sea S el espacio muestral asociado a un experimento. Una función X que asigna a cada uno de los elementos $E_i \in S$ un número real $X(E_i)$ se llama **variable aleatoria (v.a)** y se denota con una letra mayúscula *X, Y, Z*, etc la v.a es una "descripción" por ejemplo X puede ser el número de artículos defectuosos en una muestra aleatoria, o la ganancia al cabo del juego descrito anteriormente, etc.

La v.a toma valores numéricos, por ejemplo si la variable aleatoria X es el número de materias aprobadas por un alumno al cabo del primer semestre en la ESCOM los valores posibles de esta v.a, bajo el supuesto de que se cursan 6 materias en el primer semestre, son $x = 0, 1, 2, 3, 4, 5, 6$. Entonces la expresión "$X = x$" se lee "el número de materias aprobadas por un alumno al cabo del primer semestre en la ESCOM es igual a x". Recuerda que X es una descripción de algo que se cuenta (en el caso discreto) y x es un número.

Definición: Un valor x de la v.a es posible si tiene probabilidad de ocurrencia diferente de cero, es decir si $P(X = x) \neq 0$.

Definición: Si el número de valores posibles que puede tomar la variable aleatoria X es finito (los puedes contar y acabas) o infinito numerable (los puedes contar pero nunca acabas) se dice que X es una v.a **discreta.**

Ejemplo:
Un envío de 5 automóviles contiene dos de ellos con pequeñas fallas en la pintura. Si una agencia recibe en forma aleatoria tres de estos automóviles, obtén una lista de los elementos del espacio muestral S utilizando las letras B y D para bueno y defectuoso respectivamente; después asigna a cada punto muestral un valor x de la v.a X que representa el número de automóviles adquiridos por la agencia que tuvieron defectos de pintura.

Solución

Se tienen 2 automóviles defectuosos y 3 buenos de los cuales queremos seleccionar 3 al azar, no importa el orden por lo tanto la cardinalidad del espacio muestral está dada por:

$$\binom{5}{3} = \frac{5!}{(3!)(2!)} = \frac{5(4)(3!)}{(3!)(2!)} = \frac{20}{2} = 10$$

De estas 10 posibles muestras de tamaño 3 se tienen sólo tres posibilidades dado que hay 3 automóviles buenos puede ser que ninguno en la muestra sea defectuoso, puede ser que sólo uno sea defectuoso o que dos sean defectuosos, esto porque sólo hay 2 defectuosos en los automóviles disponibles.

¿Cuántas muestras de 3 automóviles se pueden sacar de los 5 que tenemos siendo los 3 buenos? $\binom{3}{3} = 1$

¿Cuántas muestras de 3 automóviles se pueden sacar de los 5 que tenemos siendo 2 buenos y 1 defectuoso?

$$\binom{3}{2}\binom{2}{1} = \frac{3!}{(2!)(1!)}(2) = 6$$

¿Cuántas muestras de 3 automóviles se pueden sacar de los 5 que tenemos siendo 1 bueno y 2 defectuosos?

$$\binom{3}{1}\binom{2}{2} = \frac{3!}{(1!)(2!)}(1) = 3$$

Listando las 10 posibles muestras y el valor de la v.a se tiene que:

Muestra i	E_i	$X = x$
1	$B_1B_2B_3$	0
2	$B_1B_2D_1$	1
3	$B_1B_2D_2$	1
4	$B_1B_3D_1$	1
5	$B_1B_3D_2$	1
6	$B_2B_3D_1$	1
7	$B_2B_3D_2$	1
8	$B_1D_1D_2$	2
9	$B_2D_1D_2$	2
10	$B_3D_1D_2$	2

El espacio muestral está dado por el conjunto siguiente:

$$S = \{B_1B_2B_3, B_1B_2D_1, B_1B_2D_2, B_1B_3D_1, B_1B_3D_2, B_2B_3D_1, B_2B_3D_2, B_1D_1D_2, B_2D_1D_2, B_3D_1D_2\}$$

y los valores posibles de la v.a son 0, 1, 2.

2.1.1 FUNCIÓN DE MASA DE PROBABILIDAD.

Definición: La función f definida sobre $\Re$ como $\boxed{f_X(x) = P(X = x)}$ se conoce como la función de masa de probabilidad o función de probabilidad f.d.p de la v.a X si cumple con las propiedades siguientes:

i) $f_X(x) \geq 0 \quad \forall \quad x \in \Re$.
ii) $\{x \mid f_X(x) \neq 0\}$ es un conjunto finito o infinito numerable.
iii) $\sum_x f_X(x) = 1$

Nota._ Si x es tal que $f_X(x) > 0$ entonces x es un valor posible de la v.a X.

Ejemplo:
En base al ejemplo de lanzar 3 veces la moneda de la sección anterior, se dijo que la probabilidad de una cara en un lanzamiento era p. Si $p = 0.4$ la v.a X tiene la función de masa de probabilidad siguiente:

$$f_X(-3) = P(X = -3) = (1-p)^3 = (1-0.4)^3 = 0.216.$$
$$f_X(-1) = P(X = -1) = 3p(1-p)^2 = 3(0.4)(1-0.4)^2 = 3(0.4)(0.6)^2 = 0.432.$$
$$f_X(1) = P(X = 1) = 3p^2(1-p) = 3(0.4)^2(1-0.4) = 3(0.4)^2(0.6) = 0.288.$$
$$f_X(3) = P(X = 3) = p^3 = (0.4)^3 = 0.064.$$

La función de masa de probabilidad de una variable aleatoria discreta se puede presentar como una tabla de valores de la manera siguiente:

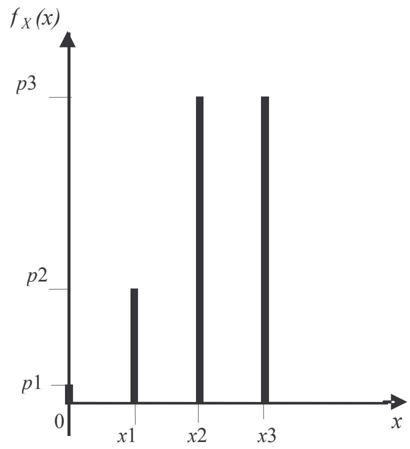
Función de probabilidad de *X*: ganancia al cabo de los tres lanzamientos.

$X = x$	$f_X(x) = P(X = x)$
-3	0.216
-1	0.432
1	0.288
3	0.064

Observa que $0.216 + 0.432 + 0.288 + 0.064 = 1$.

La f.d.p de una v.a además de representarse por una tabla se puede representar por una gráfica o por una fórmula.

Si la v.a X toma los valores posibles $x = x1, x2, x3$ y sus probabilidades asociados son $p1,\ p2,\ p3$ respectivamente entonces su f.d.p se puede representar con la gráfica siguiente:



Función de probabilidad para la v.a *X*.

También se puede decir que la v.a X tiene f.d.p dada por $f_X(x) = \binom{n}{x} p^x q^{n-x}$, con $x = 0, 1, 2, ..., n$.

O que la f.d.p de la v.a Y es de la forma $f_Y(y) = \dfrac{e^{-\lambda}\lambda^y}{y!}$ con $y = 0, 1, 2, ...$

Observa que la manera más compacta de expresar una f.d.p es mediante una fórmula pero a veces no es tan fácil construirla.

2.1.2 FUNCIÓN ACUMULATIVA.

Definición: Se define a la **función de probabilidad acumulada o acumulativa** de la v.a X en el número x, $F_X(x)$ con $-\infty < x < \infty$ como $\boxed{F_X(x) = P(X \le x) = \sum_{i \le x} P(X = i) = \sum_{i \le x} f_X(i)}$

Ejemplos:

1._ 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.

<u>**Solución**</u>

Primero necesitamos encontrar la acumulada en cada uno de los valores posibles de la v.a:

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

Observa que

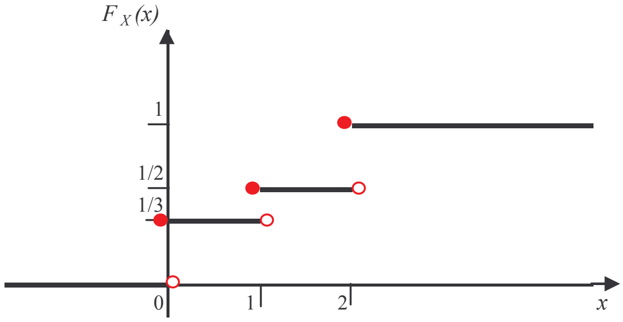
$$F_X(0.5) = P(X \le 0.5) = P(X \le 0) = P(X=0) = \frac{1}{3}.$$
$$F_X(1.9) = P(X \le 1.9) = P(X=0) + P(X=1) = \frac{1}{2}.$$
$$F_X(-1) = P(X \le -1) = 0$$

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

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

Nota: La función de probabilidad acumulada $F_X(x)$ de una v.a discreta siempre es una función por partes.

Su gráfica es de la forma escalonada:



Función de probabilidad acumulada para la v.a X

Las propiedades de la función de probabilidad acumulada son:

i) $0 \le F_X(x) \le 1 \quad , \quad -\infty < x < \infty$

Esto es porque la función de distribución acumulada es una suma de valores positivos y cuando se agotan todos los valores posibles de la v.a se tiene la probabilidad de *todo* que es igual a 1.

ii) $\lim_{x \to -\infty} F_X(x) = 0$

$\lim_{x \to \infty} F_X(x) = 1$

Esto es cierto ya que si aún no se alcanza el primer valor posible de la v.a no se ha acumulado nada de probabilidad pero sí ya se ha rebasado el último valor posible de la v.a. entonces ya se tiene toda la probabilidad posible que es uno.

iii) Si $x_1 \le x_2$ entonces $F_X(x_1) \le F_X(x_2)$

Si el valor de x_1 y el valor de x_2 están entre dos valores posibles consecutivos de la v.a entonces las acumuladas coinciden. Si entre x_1 y x_2 hay al menos un valor posible de la v.a. entonces la acumulada en x_1 es menor que la acumulada en x_2.

iv) Es continua por la derecha, es decir $F_X(x)$ existe y además $F_X(x) = \lim_{t \to x^+} F_X(t)$.

Cálculo de probabilidades en función de acumuladas.

Toda probabilidad se puede expresar en términos de la función de probabilidad acumulada, tal como se muestra en los siguientes casos:

<u>Caso 1</u> $P(X > x) = 1 - P(X \le x)$ es decir $\boxed{P(X > x) = 1 - F_X(x)}$

Observa que $S = (X > x) \cup (X \le x)$ sacando probabilidad $P(S) = P(X > x) + P(X \le x)$ despejando la probabilidad de interés $P(X > x) = 1 - P(X \le x) = 1 - F_X(x)$.

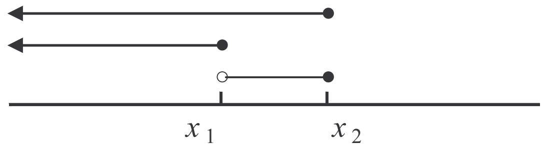
<u>Caso 2</u> Si $x_1 < x_2$ $\quad P(x_1 < X \le x_2) = P(X \le x_2) - P(X \le x_1)$ es decir

$$\boxed{P(x_1 < X \le x_2) = F_X(x_2) - F_X(x_1)}$$

$F_X(x_2) = P(X \le x_2)$

$F_X(x_1) = P(X \le x_1)$

$P(x_1 < X \le x_2)$



Caso 3 $\boxed{P(X < x) = \lim\limits_{t \to x^-} F_X(t)}$

Dado que las v.v.a.a discretas toman valores enteros esta probabilidad también se puede expresar como $P(X < x) = P(X \le x-1) = F_X(x-1)$.

Caso 4 $\boxed{P(X = x) = \lim\limits_{t \to x^+} F_X(t) - \lim\limits_{t \to x-} F_X(t)}$

Ya que $P(X = x) = P(X \le x) - P(X < x) = F_X(x) - P(X < x)$ y usando la propiedad iv) de la definición y el caso 3 llegamos al caso 4.
Otra manera de expresar esta probabilidad puntual es;
$P(X = x) = P(X \le x) - P(X \le x-1) = F_X(x) - F_X(x-1)$.

2._ Se sabe que al lanzar una moneda, a menudo sale cara (c) tres veces más que cruz (x). Esta moneda se lanza 3 veces. Sea *X* la v.a que representa el número de caras que aparecen:
a) Encuentra la función de masa de probabilidad de *X*.
b) Encuentra la función de probabilidad acumulada de *X*.
c) Grafica ambas funciones.

Solución

a) La variable aleatoria es *X*: número de caras en los tres lanzamientos. La ocurrencia de cara (c) es tres veces la de cruz (x), por lo tanto, si la ocurrencia de cruz está dada por a, la de cara es $3a$, y el número total de lanzamientos sería $4a$, con lo cual se tiene que $P(x) = \frac{a}{4a} = \frac{1}{4}$ y $P(c) = \frac{3a}{4a} = \frac{3}{4}$.

E_i	$X = x$	$P(E_i)$
ccc	3	$\left(\frac{3}{4}\right)^3$
ccx	2	$\left(\frac{3}{4}\right)^2\left(\frac{1}{4}\right)$
cxc	2	$\left(\frac{3}{4}\right)^2 \frac{1}{4}$
xcc	2	$\left(\frac{3}{4}\right)^2\left(\frac{1}{4}\right)$

cxx	1	$\left(\frac{3}{4}\right)\left(\frac{1}{4}\right)^2$
xcx	1	$\left(\frac{3}{4}\right)\left(\frac{1}{4}\right)^2$
xxc	1	$\left(\frac{3}{4}\right)\left(\frac{1}{4}\right)^2$
xxx	0	$\left(\frac{1}{4}\right)^3$

La f.d.p de la variable aleatoria se puede presentar en la tabla siguiente:

Función de probabilidad para el número de caras
en tres lanzamientos de una moneda.

$X = x$	$f_X(x) = P(X = x)$
0	$\frac{1}{64}$
1	$\frac{9}{64}$
2	$\frac{27}{64}$
3	$\frac{27}{64}$

Observa que $\frac{1}{64} + \frac{9}{64} + \frac{27}{64} + \frac{27}{64} = 1$.

b) La función de probabilidad acumulada se construye de la manera siguiente:

$$F_X(0) = P(X \leq 0) = P(X = 0) = \frac{1}{64}$$

$$F_X(1) = P(X \leq 1) = P(X = 0) + P(X = 1) = \frac{1}{64} + \frac{9}{64} = \frac{10}{64}$$

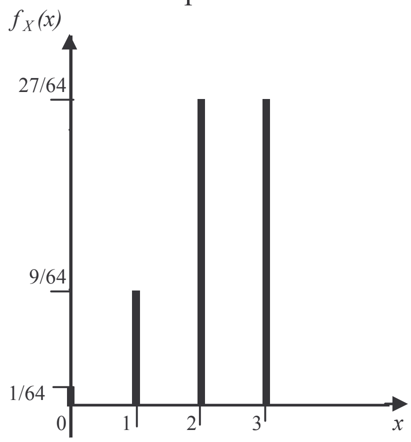
$$F_X(2) = P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2) = \frac{1}{64} + \frac{9}{64} + \frac{27}{64} = \frac{37}{64}$$

$$F_X(3) = P(X \leq 3) = P(X = 0) + \ldots + P(X = 3) = \frac{37}{64} + \frac{27}{64} = 1$$

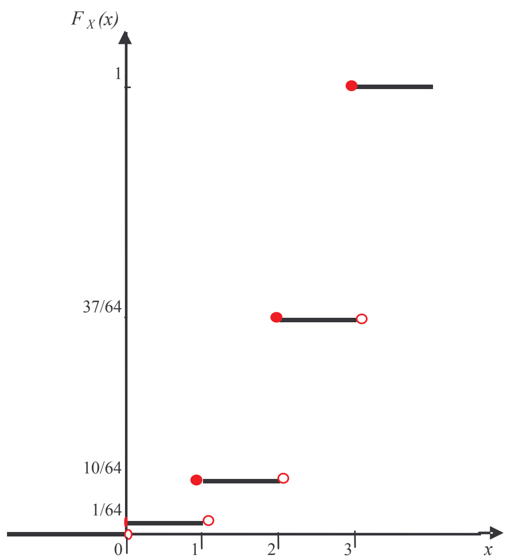
La función de probabilidad acumula de la v.a X es la función escalonada siguiente:

$$F_X(x)=\begin{cases}0 & \text{si} \quad x<0\\ \dfrac{1}{64} & \text{si} \quad 0\le x<1\\ \dfrac{10}{64} & \text{si} \quad 1\le x<2\\ \dfrac{37}{64} & \text{si} \quad 2\le x<3\\ 1 & \text{si} \quad x\ge 3\end{cases}$$

c) Las gráficas de la f.d.p y de la distribución de probabilidad acumulada son:



Función de probabilidad para la v.a X



Función de probabilidad acumulada para la v.a X.

Observa que de la gráfica de la probabilidad acumulada se tiene que:

$$P(X=1)=\frac{10}{64}-\frac{1}{64}=\frac{9}{64}$$

$$P(X=3)=\frac{64}{64}-\frac{37}{64}=\frac{27}{64}$$

$$P(X=-1)=P(X=1.4)=0$$

2.2 VALOR ESPERADO Y VARIANZA DE UNA VARIABLE ALEATORIA DISCRETA.

Definición: Si *X* es una v.a discreta con función de probabilidad $f_X(x)$, **la esperanza**, **valor esperado** o **media** de la v.a *X* está dada por

$$\boxed{E(X)=\sum_x xf_X(x)=\sum_x xP(X=x)}$$

Ejemplos:

1._ En base al ejemplo de la sección 2.1.1 donde se lanzaba una moneda 3 veces y por cada cara se ganaba 1 peso y por cada cruz se perdía 1 peso. ¿Cuál es el valor esperado de la v.a. *X*: ganancia al cabo de los tres lanzamientos?

__**Solución**__

Función de probabilidad de *X*: ganancia al cabo de los tres lanzamientos.

$X=x$	$f_X(x)=P(X=x)$
-3	0.216
-1	0.432
1	0.288
3	0.064

.

$$E(X)=\sum_x xf_X(x)=-3P(X=-3)+(-1)P(X=-1)+1P(X=1)+3P(X=3)$$

$$=-3(0.216)-1(0.432)+1(0.288)+3(0.064)=-0.6$$

La ganancia esperada en este juego es de -0.6 pesos. Lo cual indica una pérdida de dinero.

Observación: El valor esperado de la v.a no necesariamente es uno de los valores posibles de la v.a.

¿Cómo se interpreta el valor esperado de una variable aleatoria?

Imagina que el experimento se repite *n* veces, se tendría entonces *n* ganancias, si se saca el promedio de estas ganancias el número resultante sería muy parecido al valor esperado, es decir a –0.6 pesos. La similitud entre el promedio de las *n* ganancias y el valor esperado calculado aumenta a medida que *n* aumenta.

Entonces no se debe pensar que al hacer una vez el experimento se perderá la cantidad de 0.6 pesos, de hecho ya nos dimos cuenta que esto no ocurre nunca ya que este no es un valor posible de la v.a. en un solo experimento, se puede perder uno o tres pesos y se puede ganar uno o tres pesos. Pero participar muchas veces en el experimento, a la larga produce una pérdida de dinero.

2._ Un fabricante produce artículos de tal modo que el 10% es defectuoso y el 90% no lo es. Si se produce un artículo defectuoso el fabricante pierde \$1.00, mientras que un artículo sin defectos le produce una utilidad de \$5.00. Si X es la utilidad neta por artículo, encuentra su valor esperado.

Solución

X: utilidad neta por artículo.
$x = -1,\ 5.$
$P(Defectuoso) = 0.1$
$P(Bueno) = 0.9$

Los valores posibles de la v.a son –1 si el artículo es defectuoso y 5 si no lo es. La esperanza de la utilidad es:

$$E(X) = \sum_x x f_X(x) = -1P(X = -1) + 5P(X = 5)$$
$$= -1P(Defectuoso) + 5P(bueno) = -1(0.1) + 5(0.9) = 4.40$$

Nota 1._ Si la v.a X puede tomar un número infinito numerable de valores, entonces se dice que la esperanza de X existe si y sólo si $\sum_x |x| f_X(x) < \infty$, es decir la serie tiene que ser absolutamente convergente.

Nota 2._ Si X es una v.a discreta con función de probabilidad $f_X(x)$ y $g(x)$ es una función de valores reales de x, se tiene que el valor esperado de la función de la v.a está dado por la expresión siguiente:

$$\boxed{E[g(x)] = \sum_x g(x) f_X(x)}$$

PROPIEDADES DEL VALOR ESPERADO.

Sea X una v.a discreta con función de probabilidad $f_X(x)$

P1._ Si a y b son constantes y $Y = aX + b$, entonces $\boxed{E(Y) = aE(X) + b}$

Ya que; $E(Y) = \sum_x (ax+b) f_X(x) = \sum_x ax f_X(x) + \sum_x b f_X(x)$

$$= a\sum_x x f_X(x) + b\sum_x f_X(x) = aE(X) + b \quad \blacklozenge$$

Observación._ Si $a = 0$ la esperanza de la constante b es ella misma, $\boxed{E(a) = a}$

P2._ Si $X_1, X_2, \ldots, X_n$ son variables aleatorias con valores esperados $E(X_i)$, i = 1, 2, ..., n respectivamente, entonces la esperanza de la suma de las variables aleatorias es la suma de las esperanzas de las variables aleatorias.

$$\boxed{E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i)}$$

Definición: La **covarianza** entre las variables aleatorias *X* y *Y* con $E(X) = \mu_X$ y $E(Y) = \mu_Y$ está dada por:

$$\boxed{Cov(X,Y) = E[(X-\mu_X)(Y-\mu_Y)]}$$

Se obtiene una definición alterna con un poco de álgebra.

$$\begin{aligned} E[(X-\mu_X)(Y-\mu_Y)] &= E(XY - X\mu_Y - \mu_X Y + \mu_X\mu_Y) \\ &= E(XY) - E(X)\mu_Y - \mu_X E(Y) + \mu_X\mu_Y \\ &= E(XY) - \mu_X\mu_Y - \mu_X\mu_Y + \mu_X\mu_Y \end{aligned}$$

$$\boxed{Cov(X,Y) = E(XY) - \mu_X\mu_Y}$$

Observación._ En el capítulo 4 se verán distribuciones de probabilidad conjuntas y se entenderá la siguiente justificación. Si $X \perp Y$ entonces $Cov(X,Y) = 0$.

Para $X \perp Y$ se tiene que $f_{XY}(x,y) = f_X(x) f_Y(y)$ por lo que la esperanza del producto es el producto de las esperanzas.

$$E(XY) = \sum_x \sum_y xyf_{XY}(x,y) = \sum_x \sum_y xyf_X(x) f_Y(y)$$

$$E(XY) = \sum_x xf_X(x) \sum_y yf_Y(y) = E(X)E(Y) = \mu_X\mu_Y$$

Es decir, la esperanza de un producto de variables aleatorias independientes es el producto de las esperanzas de las variables aleatorias.

La covarianza entre dos variables aleatorias independientes es cero.

$$Cov(X,Y) = E(XY) - \mu_X\mu_Y = E(X)E(Y) - \mu_X\mu_Y = 0$$

VARIANZA DE UNA VARIABLE ALEATORIA DISCRETA.

Definición._ Si *X* es una v.a con media μ, la **varianza de** *X* se define como $\boxed{V(X) = E(X-\mu)^2}$

Observaciones:

1._ La varianza de la v.a es un número no negativo.

2._ Otra manera de expresar la varianza se obtiene al desarrollar el cuadrado de la expresión original y con un poco de álgebra:

$V(X) = E(X-\mu)^2 = E(X^2 - 2\mu X + \mu^2) = E(X^2) - 2\mu E(X) + \mu^2$ es decir $\boxed{V(X) = E(X^2) - \mu^2}$

3._ La varianza proporciona una medida de la variación o dispersión de una distribución alrededor de su media. Un valor pequeño de la varianza indica que la distribución de probabilidad está muy concentrada alrededor de la media y un valor grande indica que la distribución de probabilidad tiene gran dispersión alrededor de su valor central.

4._ La varianza se denota con la letra griega sigma cuadrada $V(X)=\sigma^2$, pero como la varianza nos lleva a hablar de unidades cuadradas y esto deja de tener sentido para el investigador, se acostumbra trabajar con su raíz cuadrada positiva, a la cual se le llama **desviación estándar** y se denota por sigma $\boxed{\sigma = \sqrt{V(X)} = \sqrt{\sigma^2}}$

PROPIEDADES DE LA VARIANZA.

P1._ La varianza de una constante es cero. Si a = Cte entonces $\boxed{V(a)=0}$

$$V(a)=E\left(a^2\right)-[E(a)]^2=a^2-a^2=0 \qquad \blacklozenge$$

P2._ La varianza del producto de una constante por una v.a. es el producto del cuadrado de la constante por la varianza de la v.a. $\boxed{V(aX)=a^2V(X)}$

$$V(aX)=E\left(a^2X^2\right)-[E(aX)]^2=a^2\left\{E\left(X^2\right)-[E(X)]^2\right\}=a^2V(X) \qquad \blacklozenge$$

P3._ Si X y Y son variables aleatorias cualesquiera $\boxed{V(X+Y)=V(X)+V(Y)+2Cov(X,Y)}$

$$\begin{aligned} V(X+Y) &= E[(X+Y)^2]-E^2(X+Y) \\ &= E\left(X^2+Y^2+2XY\right)-[E(X)+E(Y)]^2 \\ &= E\left(X^2\right)+E\left(Y^2\right)+2E(XY)-E^2(X)-E^2(Y)-2E(X)E(Y) \\ &= [E\left(X^2\right)-E^2(X)]+[E\left(Y^2\right)-E^2(Y)]+2[E(XY)-E(X)E(Y)] \\ &= V(X)+V(Y)+2Cov(X,Y) \qquad \blacklozenge \end{aligned}$$

Si X y Y son variables aleatorias independientes entonces la $Cov(X,Y)=0$ y por lo tanto la varianza de la suma es la suma de las varianzas.

$$\boxed{V(X+Y)=V(X)+V(Y)}$$

La varianza de la suma de tres variables aleatorias cualesquiera.

$$\begin{aligned} V(X+Y+Z) &= V[(X+Y)+Z]=V(X+Y)+V(Z)+2Cov[(X+Y),Z] \\ &= V(X)+V(Y)+V(Z)+2[Cov(X,Y)+Cov(X,Z)+Cov(Y,Z)] \end{aligned}$$

Ya que $Cov[(X+Y),Z]=E[(X+Y)Z]-E(X+Y)E(Z)$

$$\begin{aligned} &= E(XZ+YZ)-E(X)E(Z)-E(Y)E(Z) \\ &= [E(XZ)-E(X)E(Z)]+[E(YZ)-E(Y)E(Z)] \\ &= Cov(X,Z)+Cov(Y,Z) \end{aligned}$$

Si X , Y y Z son mutuamente independientes entonces $Cov(X,Y)=Cov(X,Z)=Cov(Y,Z)=0$ y por lo tanto $V(X+Y+Z)=V(X)+V(Y)+V(Z)$.

En general $V\left[\sum_{i=1}^{n} X_i\right]=\sum_{i=1}^{n} V(X_i)+2\sum_{i<j} Cov(X_i,X_j)$.

Observa también que la segunda sumatoria consta de $\frac{n(n-1)}{2}$ términos.

Si las n variables aleatorias son independientes la segunda sumatoria se hace cero y la varianza de la suma de variables aleatorias independientes es la suma de las varianzas de las variables aleatorias.

Generalizando, se tiene que si $X_1, X_2,..., X_n$ son variables aleatorias independientes entonces la varianza de su suma es la suma de las varianzas.

$$\boxed{V\left(\sum_{i=1}^{n} X_i\right)=\sum_{i=1}^{n} V(X_i)}$$

Ejemplo:

Si X es una v.a con media y varianza dadas por $E(X)=\mu$ y $V(X)=\sigma^2$. Demuestra que $E[X(X-1)]=\mu(\mu-1)+\sigma^2$.

<u>Solución</u>

$E[X(X-1)]=E(X^2-X)=E(X^2)-E(X)$

Pero $V(X)=E(X^2)-E^2(X)\ \therefore\ E(X^2)=V(X)+E^2(X)$

Entonces $E[X(X-1)]=\sigma^2+\mu^2-\mu=\sigma^2+\mu(\mu-1)$ ♦

2.3 FUNCIÓN GENERADORA DE MOMENTOS DE UNA VARIABLE ALEATORIA DISCRETA.

Definición: Si X es una v.a con $E(X)=\mu_X$ y k un entero positivo se definen dos tipos de momentos para la v.a:

1._ k-ésimo momento alrededor del origen definido como $\boxed{\mu_k^o=E(X^k)=\sum_x x^k f_X(x)}$

2._ k-ésimo momento alrededor de la media o k-ésimo momento central, definido por:

$$\boxed{\mu_k^c=E\left[(X-\mu_X)^k\right]=\sum_x (x-\mu_X)^k f_X(x)}$$

Observaciones:

i) La media de una v.a es el momento de orden 1 o primer momento alrededor del origen.

$$\mu_X=\mu_1^o$$

ii) La varianza de la v.a X es el segundo momento central de la misma.

$$\sigma_X^2=\mu_2^c$$

Definición: Si X es una v.a discreta, la función generadora de momentos f.g.m se define como:

$$\boxed{\psi_X(t) = E\left(e^{tX}\right) = \sum_x e^{tx} f_X(x)}$$

Observaciones:
1) Sacando la primera derivada de la f.g.m y evaluando en $t = 0$ se obtiene la media de la v.a.

$$\psi'_X(t=0) = \frac{d}{dt}\left[E\left(e^{tX}\right)\right]\Big|_{t=0} = E\left[\frac{d}{dt}\left(e^{tX}\right)\right]\Big|_{t=0} = E\left(Xe^{tX}\right)\Big|_{t=0} = E\left(Xe^{0(X)}\right) = E(X)$$

$$\boxed{\psi'_X(t=0) = E(X) = \mu_X}$$

Es más intuitiva la demostración usando el hecho de que la derivada de una suma es la suma de las derivadas, de la manera siguiente:

$$\psi'_X(t) = \frac{d}{dt}\left[E\left(e^{tX}\right)\right] = \frac{d}{dt}\sum_x e^{tx} f_X(x) = \sum_x \frac{d}{dt} e^{tx} f_X(x) = \sum_x e^{tx} x f_X(x)$$

Evaluando en cero se tiene:

$$\psi'_X(t=0) = \sum_x e^{0(x)} x f_X(x) = \sum_x x f_X(x) = E(X) = \mu_X$$

2) La n-ésima derivada de la f.g.m evaluada en cero es el n-ésimo momento.

La misión en la vida de la f.g.m es generar momentos y ahora ya sabemos cómo lo hace.

PROPIEDADES DE LA f.g.m.

P1._ Si X es una v.a con f.g.m $\psi_X(t)$ y $Y = aX + b$ es otra v.a, con a y b constantes, entonces

$$\boxed{\psi_Y(t) = e^{bt}\psi_X(at)}$$

Ya que $\psi_Y(t) = E\left(e^{tY}\right) = E\left(e^{t(aX+b)}\right) = E\left(e^{atX}e^{bt}\right) = e^{bt}E\left(e^{atX}\right) = e^{bt}\psi_X(at)$

P2._ Si $X_1, X_2, ..., X_n$ son variables aleatorias independientes cada una con f.g.m $\psi_{X_i}(t)$, $i = 1, 2, ... , n$ y $Y = \sum_{i=1}^{n} X_i$, entonces la f.g.m de la v.a. Y, está dada por $\boxed{\psi_Y(t) = \prod_{i=1}^{n}\psi_{X_i}(t)}$

Ya que $\psi_Y(t) = E\left(e^{tY}\right) = E\left(e^{t\sum_{i=1}^{n} X_i}\right) = E\left(\prod_{i=1}^{n} e^{tX_i}\right) = \prod_{i=1}^{n} E\left(e^{tX_i}\right) = \prod_{i=1}^{n}\psi_{X_i}(t)$

Ejemplos:

1._ Supón que X es una v.a con $E(X) = 1$, $E\left(X^2\right) = 2$ y $E\left(X^3\right) = 5$. Determina el valor del tercer momento central de X.

Solución

$\mu_3^c = E\left[(X-\mu)^3\right] = E\left(X^3 - 3X^2\mu + 3X\mu^2 - \mu^3\right) = E\left(X^3\right) - 3\mu E\left(X^2\right) + 3\mu^3 - \mu^3$

$\mu_3^c = 5 - 3(2) + 2 = \underline{1}$

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

Solución

La media de la v.a es el primer momento, entonces necesitamos la primera derivada de la f.g.m y evaluar en cero:

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

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

$$V(X) = E\left(X^2\right) - \mu_X^2 = 1 - \left(\frac{1}{2}\right)^2 = \frac{3}{4}$$

3._ 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

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

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

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

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

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

$$\psi''_Y(0) = ce^{-c}\left[e^{c\psi_X(0)}\psi''_X(0) + \mu^2 e^{c\psi_X(0)}c\right]$$

$$= ce^{-c}\left(e^c\psi''_X(0) + c\mu^2 e^c\right)$$

$$= c\psi''_X(0) + c^2\mu^2$$

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 $V(Y) = E\left(Y^2\right) - E^2(Y) = c\sigma^2 + c\mu^2 + c^2\mu^2 - c^2\mu^2 = c\left(\sigma^2 + \mu^2\right)$

4._ Si $\psi_X(t) = \frac{e^t}{6} + \frac{2e^{2t}}{6} + \frac{3e^{3t}}{6}$. Encuentra a) la esperanza de la v.a, b) su varianza y c) Su función de probabilidad.

Solución

a) $\psi'_X(t=0) = \frac{e^t}{6} + \frac{4e^{2t}}{6} + \frac{9e^{3t}}{6}\Big|_{t=0} = \frac{1+4+9}{6} = \frac{14}{6} \;\therefore\; \mu_X = \frac{7}{3}$

b) $\psi''_X(t=0) = \frac{e^t}{6} + \frac{8e^{2t}}{6} + \frac{27e^{3t}}{6}\Big|_{t=0} = \frac{1+8+27}{6} = \frac{36}{6} \;\therefore\; E(X^2) = 6$

$V(X) = E(X^2) - E^2(X) = 6 - (7/3)^2 = 6 - \frac{49}{9} = \frac{5}{9}$

c) $\psi_X(t) = E(e^{tX}) = \sum_x e^{tx} P(X=x) = \frac{1}{6}e^t + \frac{2}{6}e^{2t} + \frac{3}{6}e^{3t}$

La función de probabilidad de esta v.a está dada por la tabla siguiente:

$X = x$	$f_X(x) = P(X=x)$
1	$\frac{1}{6}$
2	$\frac{2}{6}$
3	$\frac{3}{6}$

Observa que $\frac{1}{6} + \frac{2}{6} + \frac{3}{6} = 1$.

5._ 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: 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.

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

La f.d.p puede presentarse mediante la tabla siguiente:

$X = x$	$f_X(x) = P(X = x)$
2	$\frac{1}{10}$
3	$\frac{2}{10}$
4	$\frac{3}{10}$
5	$\frac{4}{10}$

O mediante la fórmula $f_X(x) = \frac{x-1}{10}$ 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 $E(X) = \sum_{x=2}^{5} xf_X(x) = 0.2 + 0.6 + 1.2 + 2 = \underline{4}$

Para calcular la varianza necesitamos 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 $F_X(x) = P(X \leq x)$, primero encontramos las acumuladas en cada uno de los valores posibles de la v.a y luego construimos la función por partes.

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

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

$F_X(4) = P(X \leq 4) = P(X = 2) + P(X = 3) + P(X = 4) = 0.6$

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

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

2.4 DISTRIBUCIÓN DE BERNOULLI Y BINOMIAL.

Muchos experimentos tienen sólo dos posibles resultados por ejemplo el lanzamiento de una moneda puede resultar en cara o cruz, el género de un bebe al nacer puede ser femenino o masculino, una prueba médica puede ser diagnosticada como positiva o negativa, al evaluar un trabajo terminal puede catalogarse como bueno o como malo, etc. Este tipo de experimentos se llaman **ensayos de Bernoulli,** el resultado que es de interés para el investigador se conoce como **éxito** y el otro resultado se conoce como **fracaso**, con probabilidades asociadas $P(\acute{e}xito) = p$ y $P(fracaso) = 1 - p = q$.

DISTRIBUCIÓN DE BERNOULLI.

La variable aleatoria *X:* resultado en un ensayo Bernoulli, con valores posibles x = 0, 1. Cero para el fracaso y uno para el éxito. Tiene una distribución de Bernoulli con parámetro *p,* dada por la expresión

$$f_X(x) = p^x q^{1-x} \quad x = 0,\ 1$$

Su valor esperado es $E(X) = 0 f_X(0) + 1 f_X(1) = 0(q) + 1(p)$, es decir $\boxed{E(X) = p}$

El segundo momento de la v.a Bernoulli es: $E(X^2) = 0^2 q + 1^2 p = E(X) = p$.
Dado que:

$V(X) = E(X^2) - E^2(X) = p - p^2 = p(1-p)$ se tiene que $\boxed{V(X) = pq}$

La **f.g.m**:

$$\psi_X(t) = E(e^{tX}) = e^{t(0)} f_X(0) + e^{t(1)} f_X(1) = P(X = 0) + e^t P(X = 1) = q + e^t p$$

$$\boxed{\psi_X(t) = q + pe^t}$$

No hay más que decir de esta distribución que es la más sencilla, pero muy útil para las distribuciones de variables aleatorias discretas que veremos en este capítulo.

DISTRIBUCIÓN BINOMIAL.

Definición: Un experimento es **binomial** si cumple con las dos condiciones siguientes:
i) Consiste de n ensayos Bernoulli independientes.
ii) La probabilidad de éxito p en cada ensayo es constante.
Se dice que la variable aleatoria X: número de éxitos en un experimento binomial, con valores posibles $x = 0, 1, 2, ..., n$, tiene una distribución de probabilidad binomial con parámetros n y p, lo cual se denota como $X \sim b(x; n, p)$, si su f.d.p está dada por

$$b(x;n,p) = P(X = x) = \binom{n}{x} p^x q^{n-x} \quad x = 0, 1, \dots, n$$

Observa que si $X_1, X_2, ..., X_n$ son n variables aleatorias Bernoulli, entonces la v.a binomial X puede escribirse como la suma de estas, es decir $X = X_1 + X_2 + \cdots + X_n$.
Recuerda que la f.g.m de una suma de variables aleatorias independientes es el producto de las funciones generadoras de momentos individuales.
Por lo tanto, la f.g.m para X es $\psi_X(t) = \prod_{i=1}^{n} \psi_{X_i}(t)$ o bien, $\boxed{\psi_X(t) = \left(q + pe^t\right)^n}$
La esperanza de la v.a binomial es la primera derivada de su f.g.m evaluada en cero:

$$\psi'_X(t=0) = n\left(q + pe^t\right)^{n-1} pe^t \,|_{t=0} = n\left(q + pe^0\right)^{n-1} pe^0 = np(q+p)^{n-1} \quad \therefore \quad \boxed{E(X) = np}$$

Una manera alternativa de sacar la esperanza de la binomial es recordando que la v.a binomial es una suma de n Bernoullis independientes cada una con media p, y que la esperanza de una suma de variables aleatorias independientes es la suma de las esperanzas,

$$E(X) = E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i) = np$$

Para encontrar **la varianza** se necesita el segundo momento de la v.a por lo tanto se calcula la segunda derivada de la f.g.m evaluada en cero:

$$\psi''_X(t=0) = np\left[\left(q + pe^t\right)^{n-1} e^t + e^t(n-1)\left(q + pe^t\right)^{n-2} pe^t\right]\Bigg|_{t=0}$$

$$= np\left[\left(q + pe^0\right)^{n-1} e^0 + e^0(n-1)\left(q + pe^0\right)^{n-2} pe^0\right]$$
$$= np[1 + (n-1)p] = np(1 + np - p) = np(q + np) = npq + n^2p^2$$

$$\therefore \; V(X) = E\left(X^2\right) - E^2(X) = npq + n^2p^2 - n^2p^2$$

La varianza de la v.a binomial es $\boxed{V(X) = npq}$

La varianza de la v.a Binomial también se obtiene con el hecho de que la varianza de una suma de variables aleatorias independientes es la suma de las varianzas de las variables aleatorias que se están sumando y recordando que la varianza de una Bernoulli es pq.

$$V(X) = V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V(X_i) = npq$$

La función de probabilidad acumulada.

$F_X(x) = P(X \leq x) = \sum_{i=0}^{x} \binom{n}{i} p^i q^{n-i}$. Los valores ya están tabulados en las llamadas tablas binomiales que se encuentran en los anexos de los libros de probabilidad.

Teorema._ Si $X_1, X_2, ..., X_k$ son k variables aleatorias independientes y $X_i \sim b(x_i; n_i, p)$ $i = 1, 2, ..., k$. Entonces la v.a $X = X_1 + X_2 + \cdots + X_k$ tiene una distribución binomial con parámetros $n = \sum_{i=1}^{k} n_i$ y p.

Demostración
La f.g.m para X es

$$\psi_X(t) = \prod_{i=1}^{k} \psi_{X_i}(t) = (q + pe^t)^{n_1} (q + pe^t)^{n_2} \cdots (q + pe^t)^{n_k} = (q + pe^t)^{\sum_{i=1}^{k} n_i}$$

$$\boxed{\psi_X(t) = (q + pe^t)^n \quad \text{con} \quad n = \sum_{i=1}^{k} n_i}$$

Observa que es la f.g.m para una binomial con parámetros $n = \sum_{i=1}^{k} n_i$ y probabilidad de éxito igual a p.

Ejemplos:

1._ Una moneda con probabilidad de cara 0.6 se lanza nueve veces. ¿Cuál es la probabilidad de obtener un número par de caras?

Solución

X: Número de caras.
$X \sim b(x; n = 9, p = 0.6)$
A: Número par de caras.

$P(A) = P(X = 0) + P(X = 2) + P(X = 4) + P(X = 6) + P(X = 8)$

$$= \binom{9}{0}(0.6)^0(0.4)^9 + \binom{9}{2}(0.6)^2(0.4)^7 + \binom{9}{4}(0.6)^4(0.4)^5 + \binom{9}{6}(0.6)^6(0.4)^3 + \binom{9}{8}(0.6)^8(0.4)$$

$= 0.0002621 + 0.0212336 + 0.1672151 + 0.2508226 + 0.0604661 = 0.499 \approx 0.5$

2._ Tres hombres A, B y C disparan a un blanco. Supón que A dispara 3 veces y la probabilidad de que dé en el blanco en un disparo concreto es $\frac{1}{8}$, que B dispara 5 veces y la probabilidad de que dé en el blanco es $\frac{1}{4}$ y C dispara sólo 2 veces con probabilidad de dar en el blanco de $\frac{1}{2}$.

a) ¿Cuál es el número esperado de disparos que darán en el banco?
b) ¿Cuál es la varianza del número de disparos que darán en el blanco?

Solución

X_1: Número de disparos de A que dan en el blanco.

X_2: Número de disparos de B que dan en el blanco.

X_3: Número de disparos de C que dan en el blanco.

X: Número total de disparos que dan en el blanco.

$n_1 = 3$ $\qquad n_2 = 5$ $\qquad n_3 = 2$

$p_1 = \frac{1}{8}$ $\qquad p_2 = \frac{1}{4}$ $\qquad p_3 = \frac{1}{2}$

$X_i \sim b(x_i; n_i, p_i) \quad i = 1,\ 2,\ 3.$

a) El número total de disparos que dan en el blanco es la suma del número de disparos del señor A que dan en el blanco, el número de disparos del señor B que dan en el blanco y el número de disparos del señor C que dan en el blanco.

$$X = X_1 + X_2 + X_3$$

$$E(X) = E(X_1 + X_2 + X_3) = \sum_{i=1}^{3} E(X_i) = \sum_{i=1}^{3} n_i p_i$$

$$E(X) = 3\left(\frac{1}{8}\right) + 5\left(\frac{1}{4}\right) + 2\left(\frac{1}{2}\right) = \frac{3+10+8}{8} = \frac{21}{8} = 2.625$$

b) Los tiros son independientes unos de otros, por lo tanto

$$V(X) = \sum_{i=1}^{3} V(X_i) = \sum_{i=1}^{3} n_i p_i q_i = 3\left(\frac{1}{8}\right)\left(\frac{7}{8}\right) + 5\left(\frac{1}{4}\right)\left(\frac{3}{4}\right) + 2\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)$$

$$= \frac{21}{64} + \frac{15}{16} + \frac{1}{2} = \frac{21+60+32}{64} = \frac{113}{64} = 1.766$$

3._ La probabilidad de que un paciente se recupere de una rara enfermedad en la sangre es de 0.4. Si se sabe que 15 personas han contraído esta enfermedad ¿Cuál es la probabilidad de que:

a) Por lo menos 10 de ellos sobrevivan?
b) Sobrevivan de 3 a 8 personas?
c) Se salven exactamente 5?

Solución

X: Número de sobrevivientes.

$p = 0.4$, $q = 0.6$, $n = 15$
$X \sim b(x; n = 15, p = 0.4)$

Este problema lo resolvemos utilizando las tablas binomiales.
a) $P(X \geq 10) = 1 - P(X < 10) = 1 - P(X \leq 9) = 1 - F_X(9) = 1 - 0.9662 = 0.0338$
b) $P(3 \leq X \leq 8) = F_X(8) - F_X(2) = 0.905 - 0.0271 = 0.8779$
c) $P(X = 5) = F_X(5) - F_X(4) = 0.4032 - 0.2173 = 0.1859$

4._ La probabilidad de que una cierta clase de componentes sobreviva a una prueba de choque dada es $\frac{3}{4}$. Determina la probabilidad de que resistan exactamente 2 de los siguientes 4 componentes que se van a probar.

Solución

X: Número de componentes que sobreviven a la prueba.
$p = \frac{3}{4}$, $q = \frac{1}{4}$, $n = 4$
$X \sim b\left(x; n = 4, p = \frac{3}{4}\right)$

$$P(X = 2) = \binom{4}{2}\left(\frac{3}{4}\right)^2\left(\frac{1}{4}\right)^2 = \left(\frac{4!}{2!2!}\right)\left(\frac{9}{16}\right)\left(\frac{1}{16}\right) = \frac{54}{16^2} = \frac{27}{128} = 0.21$$

5._ Un fabricante de cera para pisos desarrolla dos productos nuevos, A y B que desea someter a la evaluación de las amas de casa para determinar cuál es mejor. Las dos ceras A y B se aplican en los pisos de 15 casas. Se supone que en realidad no hay diferencia en calidad entre las dos marcas. ¿Cuál es la probabilidad de que
a) 10 o más amas de casa vayan a preferir la marca A?
b) 10 o más amas de casa prefieran A o B?

Solución

Sean las variables aleatorias:
X: Número de amas de casa que prefieren A.
Y: Número de amas de casa que prefieren B.

Observa que se cumple lo siguiente:

$X = x$	$Y = y$	$X = x$	$Y = y$
0	15	8	7
1	14	9	6
2	13	10	5
3	12	11	4
4	11	12	3
5	10	13	2
6	9	14	1
7	8	15	0

a) $X \sim b(x; n=15, p=0.5)$

$P(X \geq 10) = 1 - P(X \leq 9) = 1 - F_X(9) = 1 - 0.849 = 0.151$

b) $Y \sim b(y; n=15, p=0.5)$

P(10 o más amas de casa prefieran B) = $P(Y \geq 10) = P(X \leq 5) = F_X(5) = 0.151$

Entonces $P(X \geq 10 \text{ o } X \leq 5) = P(X \geq 10) + P(X \leq 5) = 0.151 + 0.151 = 0.302$

2.5 Distribución Geométrica.

La distribución geométrica está relacionada con una secuencia de ensayos Bernoulli con misma probabilidad de éxito pero el número de ensayos no es fijo. La variable aleatoria de interés *X*, es el número de ensayos necesarios hasta obtener el primer éxito. Los valores posibles que toma la variable aleatoria son x = 1, 2, 3, ... Se dice que *X* tiene una distribución geométrica con parámetro *p* (probabilidad de éxito) lo cual se escribe como $X \sim G(x; p)$ si su f.d.p es de la forma:

$$\boxed{G(x; p) = P(X = x) = q^{x-1} p} \text{ con } x = 1, 2, ...$$

La f.g.m. para esta distribución es:

$$\psi_X(t) = E\left(e^{tX}\right) = \sum_{i=1}^{\infty} e^{tx} pq^{x-1} = p\left(e^t + e^{2t}q + e^{3t}q^2 + e^{4t}q^3 + \cdots\right) = pe^t\left(1 + e^t q + e^{2t}q^2 + e^{3t}q^3 + \cdots\right)$$

$$= pe^t \sum_{i=0}^{\infty} e^{it} q^i = pe^t \sum_{i=0}^{\infty} \left(e^t q\right)^i = pe^t \left[\frac{1}{1 - qe^t}\right]$$

$$\therefore \boxed{\psi_X(t) = \frac{pe^t}{1 - qe^t}}$$

Nota._ Utilizamos el resultado de la serie geométrica $\sum_{i=0}^{\infty} r^i = \frac{1}{1-r}$

Media y varianza.

$$E(X) = \psi'_X(t=0) = \left.\frac{\left(1 - qe^t\right)pe^t - pe^t\left(-qe^t\right)}{\left(1 - qe^t\right)^2}\right|_{t=0} = \left.\frac{pe^t - pqe^{2t} + pqe^{2t}}{\left(1 - qe^t\right)^2}\right|_{t=0} = \frac{pe^t}{\left(1 - qe^t\right)^2} = \frac{p}{p^2}$$

La media de la v.a Geométrica es: $\boxed{E(X) = \frac{1}{p}}$

$$\psi''_X(t=0) = \left.\frac{\left(1 - qe^t\right)^2 pe^t + pe^t 2\left(1 - qe^t\right)qe^t}{\left(1 - qe^t\right)^4}\right|_{t=0} = \frac{p^3 + 2p^2 q}{p^4} = \frac{p^2(p + 2q)}{p^4} = \frac{p + 2q}{p^2}$$

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

La varianza de la v.a Geométrica es: $\boxed{V(X)=\frac{q}{p^2}}$

Función de probabilidad acumulada.

$$F_X(x)=P(X\le x)=\sum_{i=1}^{x}pq^{i-1}=p\sum_{i=1}^{x}q^{i-1}=p\left[1+q+q^2+q^3+\cdots+q^{x-1}\right]=p\sum_{i=0}^{x-1}q^i=p\left[\frac{1-q^x}{1-q}\right]$$

La función de probabilidad acumulada de la v.a geométrica es $\boxed{F_X(x)=1-q^x}$

Nota._ Utilizamos el resultado de la serie geométrica $\sum_{i=0}^{x}r^i=\frac{1-r^{x+1}}{1-r}$

Ejemplos:

1._ Se va a realizar cierto experimento hasta que se obtenga un resultado exitoso. Los ensayos son independientes y el costo de efectuar el experimento es $25,000 dólares, sin embargo, si se produce una falla, cuesta $5,000 dólares iniciar el siguiente ensayo. Al experimentador le gustaría determinar el costo esperado del proyecto. Supón que $p = 0.25$.

Solución

La v.a es

X: Número de ensayos hasta obtener un experimento exitoso.

$X \sim G(x; p = 0.25)$

$C(X)$: Costo de experimento en función de la v.a.

Necesitamos construir el costo del proyecto como una función del número de ensayos requeridos hasta el primer éxito.

$$C(X)=25{,}000X+5{,}000(X-1)=30{,}000X-5{,}000$$

$$E[C(X)]=30{,}000E(X)-5{,}000=30{,}000\left(\frac{1}{p}\right)-5{,}000=30{,}000\left(\frac{100}{25}\right)-5{,}000=\$115{,}000.$$

2._ Una compañía aeroespacial ha construido cinco mísiles. La probabilidad de un disparo exitoso en cualquier prueba es 0.95. Supón lanzamientos independientes ¿Cuál es la probabilidad de que la primer falla ocurra en el quinto disparo?

Solución

El éxito es el evento de interés y en este problema nos interesa que el disparo sea no exitoso entonces:

$p = P(\text{éxito}) = P(\text{disparo no exitoso}) = 0.05$

$q = P(\text{fracaso}) = P(\text{disparo exitoso}) = 0.95$

X: Número de disparos hasta que ocurre la primera falla.

$X \sim G(x; p = 0.05)$

$P(X = 5) = (0.95)^4 (0.05) = 0.0407$

3._ La compañía A planea visitar clientes potenciales hasta que se realice una venta considerable. Cada presentación de venta cuesta 1,000 dólares y cuesta 4,000 dólares viajar para visitar al siguiente cliente y realizar una nueva presentación.
a) ¿Cuál es el costo esperado de la realización de una venta si la probabilidad de hacer una venta después de cualquier presentación es 0.10?
b) Si la ganancia esperada en cada venta es 15,000 dólares, ¿Deben efectuarse los viajes?
c) Si el presupuesto para publicidad es sólo 100, 000 dólares ¿Cuál es la probabilidad de que esta suma sea gastada sin que se logre ningún pedido?

Solución

X: Número de visitas hasta realizar la primera venta importante.

$X \sim G(x; p = 0.1)$

Sabemos que $E(X) = \frac{1}{p} = \frac{1}{0.1} = 10$

$C(X)$: Costo de la realización de la primera venta en función de la v.a.

a) El costo en función del número de visitas es $C(X) = 1000X + 4000(X - 1) = 5000X - 4000$
Por lo tanto el costo esperado es $E[C(X)] = 5000E(X) - 4000 = 5000(10) - 4000 = \$46,000$.

b) El costo esperado para realizar la primera venta es de \$46,000 > \$15,000 que es la ganancia esperada en esa venta, por lo tanto no deben efectuarse los viajes.

c) El costo para realizar la primera venta está dado por

$$C(X) = 5,000X - 4,000$$

Entonces
$$\begin{aligned} P[C(X) > 100,000] &= P(5,000X - 4,000 > 100,000) \\ &= P(5,000X > 104,000) \\ &= P\left(X > \frac{104}{5}\right) = P(X > 20.8) = 1 - P(X \le 20) \\ &= 1 - F_X(20) = 1 - \left(1 - q^{20}\right) = (0.9)^{20} = 0.1216 \end{aligned}$$

4._ Dado que se ha lanzado una moneda normal 10 veces, y se han obtenido cero caras. ¿Cuál es la probabilidad de que se tenga que lanzar al menos dos veces más para obtener la primera cara?

Solución

X : Número de lanzamientos para obtener la primera cara.

$X \sim G(x; p = 0.5)$

Entonces

$$P(X \ge 12 | X > 10) = \frac{P(X \ge 12,\ X > 10)}{P(X > 10)} = \frac{P(X \ge 12)}{P(X > 10)} = \frac{1 - F_X(11)}{1 - F_X(10)} = \frac{q^{11}}{q^{10}} = q$$

Propiedad._ Si X tiene una distribución geométrica, para dos enteros positivos cualesquiera s y t, se tiene que

$$\boxed{P(X \geq s+t \mid X > s) = P(X \geq t)}$$

Esta propiedad se conoce como la propiedad de la falta de memorio o pérdida de la memoria.

<u>Demostración:</u>

Observa que: $P(X \geq t) = 1 - P(X \leq t-1)$

$$= 1 - F_X(t-1)$$

$$= 1 - \left(1 - q^{t-1}\right) = q^{t-1}$$

Y por otro lado

$$P(X \geq s+t \mid X > s) = \frac{P(X > s, X \geq s+t)}{P(X > s)} = \frac{P(X \geq s+t)}{1 - P(X \leq s)} = \frac{1 - P(X \leq s+t-1)}{1 - F_X(s)}$$

$$= \frac{1 - F_X(s+t-1)}{1 - F_X(s)} = \frac{1 - \left(1 - q^{s+t-1}\right)}{1 - \left(1 - q^{s}\right)} = \frac{q^{s+t-1}}{q^{s}} = q^{t-1} = P(X \geq t)$$

Así por ejemplo el problema de las 10 cruces pudo resolverse usando esta propiedad de la pérdida de la memoria;

$$P(X \geq 12 \mid X > 10) = P(X \geq 2) = 1 - F_X(1) = 1 - (1-q) = \frac{1}{2}$$

5._ Sea Y una v.a geométrica con probabilidad de éxito igual a p, demuestra que:
a) para un entero positivo a , $P(Y > a) = q^a$
b) para los enteros positivos a y b $P(Y > a+b \mid Y > a) = P(Y > b)$
<u>Solución</u>

a) $P(Y > a) = 1 - P(Y \leq a) = 1 - F_Y(a) = 1 - \left(1 - q^a\right) = q^a$.

b) $P(Y > a+b \mid Y > a) = \dfrac{P(Y > a+b)}{P(Y > a)} = \dfrac{q^{a+b}}{q^a} = q^b = P(Y > b)$.

6._ Un contador público ha encontrado que 9 de 10 auditorias de compañías contienen errores importantes. Si el contador revisa la contabilidad de una serie de compañías.
a) ¿Cuál es la probabilidad de que la primera contabilidad con errores importantes se encuentre en la tercera contabilidad revisada?
b) ¿Cuál es la probabilidad de que la primera contabilidad con errores importantes se encuentre después de revisar la tercera?
c) Encuentra la media y la desviación estándar del número de contabilidades que hay que revisar para obtener la primera con errores importantes?

<u>Solución</u>

X: Número de contabilidades revisadas para obtener el primer error importante.

$$X \sim G\left(x; p = \frac{9}{10}\right)$$

a) $P(X=3) = \left(\frac{1}{10}\right)^2\left(\frac{9}{10}\right) = \frac{9}{1000} = 0.009$

b) $P(X>3) = 1 - F_X(3) = q^3 = \left(\frac{1}{10}\right)^3 = 0.001$

c) $E(X) = \frac{1}{p} = \frac{10}{9} = 1.1$

$V(X) = \frac{q}{p^2} = \frac{1}{10}\left(\frac{100}{81}\right) = \frac{10}{81} = 0.123$

Sacando la raíz positiva de la varianza obtenemos la desviación estándar, es decir $\sigma = \sqrt{V(X)} = 0.35$.

2.6 Distribución Hipergeométrica y Distribución de Poisson.

Supón ahora que tenemos una población con *r* objetos, de los cuales r_1 son del tipo A y r_2 del tipo B. Entonces el número de objetos en la población es $r = r_1 + r_2$.
Se toma una muestra aleatoria sin reemplazo de tamaño *n*, menor o igual a *r*, siendo de interés la v.a
X: número de objetos del tipo A en la muestra.

Esta v.a se conoce como **hipergeométrica.** Observa que el valor de la v.a no puede ser mayor que *n* ni mayor que r_1 es decir $x \le \min(r_1, n)$.

Ejemplos:

i) $n = 10$ muestra de tamaño 10.
$r_1 = 5$ en la población hay 5 objetos del tipo A.
El número de objetos del tipo A en la muestra a lo más es 5 que es precisamente r_1.

ii) $n = 6$ muestra de tamaño 6.
$r_1 = 7$ en la población hay 7 objetos del tipo A.
El número de objetos del tipo A en la muestra a lo más es 6 que es precisamente el tamaño de la muestra *n*.

De igual manera el número de objetos del tipo B en la muestra, $n - x$, no puede exceder al número total de objetos del tipo B en la población, es decir, $n - x \le r_2 \;\therefore\; n - r_2 \le x$ y tampoco puede ser un número negativo por lo tanto $\max(0, n - r_2) \le x \le \min(r_1, n)$.

Ejemplos:

i) $n = 6$ muestra de tamaño 6.

$r_1 = 7$ en la población hay 7 objetos del tipo A.
$r_2 = 5$ y 5 objetos del tipo B, entonces el número total de elementos en la población es $r = 12$.
La variable aleatoria de interés toma valores entre 1 y 6 inclusive, ya que $\max(0,6-5) \le x \le \min(7,6)$.

ii) $n = 6$ muestra de tamaño 6.
$r_1 = 5$ en la población hay 5 objetos del tipo A.
$r_2 = 10$ elementos del tipo B, entonces el número total de elementos en la población es $r = 15$.
La variable aleatoria de interés toma valores entre 0 y 5 inclusive, ya que $\max(0,6-10) \le x \le \min(5,6)$.

Si tomamos una muestra aleatoria de tamaño 6, el mínimo número de objetos de la clase A es cero (hay 10 objetos de la clase B, entonces los 6 de la muestra pueden ser de este tipo) y el máximo número posible es 5 ya que sólo hay 5 del tipo A en la población.

Función de probabilidad.
La expresión $X \sim H(x; r_1, r_2, n)$ indica que la variable aleatoria X tiene una distribución hipergeométrica con parámetros r_1, r_2 y *n,* tal distribución es de la forma:

$$f_X(x) = \frac{\binom{r_1}{x}\binom{r-r_1}{n-x}}{\binom{r}{n}} \qquad \max(0, n-r_2) \le x \le \min(r_1, n)$$

Esperanza y varianza.
Sean las variables aleatorias Bernoulli $X_i = \begin{cases} 1 & \text{si } X_i \text{ es del tipo A} \\ 0 & \text{si } X_i \text{ es del tipo B} \end{cases}$ para $i = 1, 2, \ldots, n$.

Para las cuales $E(X_i) = p = \frac{r_1}{r}$.

Observa que la v.a. hipergeométrica es la suma de estas variables aleatorias Bernoulli, $X = \sum_{i=1}^{n} X_i$, usando las propiedades del valor esperado se tiene que:

$$E(X) = E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i) = nE(X_i)$$

Y así se tiene que la esperanza de la v.a hipergeométrica es: $\boxed{E(X) = n\frac{r_1}{r}}$

La muestra se toma sin reemplazo, por lo tanto las variables aleatorias X_i no son independientes. Para encontrar la varianza de la v.a hipergeométrica necesitamos la varianza de cada X_i y la covarianza $Cov(X_i, X_j)\ \forall\ i < j$.

La varianza de cada Bernoulli es $V(X_i) = pq = \left(\frac{r_1}{r}\right)\left(1 - \frac{r_1}{r}\right)$ (1)

Y la covarianza entre X_i y X_j está dada por:
$Cov(X_i, X_j) = E(X_i X_j) - E(X_i)E(X_j) \quad \forall\ i < j$

Los valores posibles que toma el producto son 0 y 1 de la manera siguiente:

$$\left.\begin{array}{l} X_i = 0 \text{ y } X_j = 1 \\ X_i = 1 \text{ y } X_j = 0 \\ X_i = 0 \text{ y } X_j = 0 \end{array}\right\} X_i X_j = 0$$

$X_i = 1$ y $X_j = 1 \quad X_i X_j = 1$

Por lo tanto

$$E(X_i X_j) = 1P(X_i = 1, X_j = 1) = P(X_i = 1)P(X_j = 1 \mid X_i = 1) = \frac{r_1}{r}\frac{r_1 - 1}{r - 1}$$

Entonces $Cov(X_i, X_j) = \frac{r_1}{r}\frac{r_1 - 1}{r - 1} - \left(\frac{r_1}{r}\right)^2 = \frac{r_1}{r}\left(\frac{r_1 - 1}{r - 1} - \frac{r_1}{r}\right) = \frac{r_1}{r}\left(\frac{r_1 r - r - r_1 r + r_1}{r(r-1)}\right) = \frac{r_1}{r}\left(\frac{r_1 - r}{r(r-1)}\right)$ (2)

Usando las ecuaciones (1) y (2) en la expresión de la varianza:

$$V(X) = V\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} V(X_i) + 2\sum_{i<j} Cov(X_i, X_j)$$

$$= n\frac{r_1}{r}\left(1 - \frac{r_1}{r}\right) + 2\frac{n(n-1)}{2}\frac{r_1}{r}\left(\frac{r_1 - r}{r(r-1)}\right)$$

$$= n\frac{r_1}{r}\left(1 - \frac{r_1}{r}\right) + n(n-1)\frac{r_1}{r}\left(\frac{r_1 - r}{r(r-1)}\right)$$

$$= n\frac{r_1}{r}\left(\frac{r - r_1}{r}\right) - n\left(\frac{r_1}{r}\right)\left(\frac{r - r_1}{r}\right)\left(\frac{n-1}{r-1}\right)$$

$$= n\frac{r_1}{r}\left(\frac{r - r_1}{r}\right)\left(1 - \frac{n-1}{r-1}\right)$$

Finalmente tenemos que la varianza de la v.a hipergeometrica está dada por:

$$\boxed{V(X) = n\left(\frac{r_1}{r}\right)\left(1 - \frac{r_1}{r}\right)\left(1 - \frac{n-1}{r-1}\right)}$$

Nota._ Supón que se tiene una población de tamaño r con r_1 objetos de la clase A y $r - r_1$ de la clase B, y se toma una muestra aleatoria (m.a) de tamaño n. La v.a Y: Número de objetos de la clase A en la muestra puede tener dos distribuciones de probabilidad diferentes, dependiendo de la manera en que se toma la m.a:

i) Si la muestra se toma con reemplazo, la probabilidad de éxito (el objeto es de la clase A) es constante igual a $\frac{r_1}{r}$, hay independencia y por lo tanto $Y \sim b(y;n,p)$ con $E(X) = np = n\frac{r_1}{r}$ y

$V(Y) = npq = n\frac{r_1}{r}\left(1-\frac{r_1}{r}\right)$

ii) Si la muestra se toma sin reemplazo entonces $Y \sim H(y;r_1,r_2,n)$ con $E(Y) = np = n\frac{r_1}{r}$ y

$$V(Y) = n\left(\frac{r_1}{r}\right)\left(1-\frac{r_1}{r}\right)\left(1-\frac{n-1}{r-1}\right)$$

El valor esperado es el mismo pero si el muestreo es sin reemplazo la varianza es menor ya que:

$$n\left(\frac{r_1}{r}\right)\left(1-\frac{r_1}{r}\right) > n\left(\frac{r_1}{r}\right)\left(1-\frac{r_1}{r}\right)\left(1-\frac{n-1}{r-1}\right)$$

pero $1 > 1-\frac{n-1}{r-1}$.

Si *r* es grande comparado con *n*, entonces $\frac{n}{r} \to 0$ y la varianza de la hipergeométrica se parece mucho a la varianza de la binomial, es decir para *n* fijo y *r* grande (poblaciones muy grandes) hay poca diferencia entre muestreo con y sin reemplazo (binomial e hipergeométrica).

Función de probabilidad acumulada.

Existen en escuelas del IPN tablas de la distribución acumulada realizadas por los profesores, pero en los libros difícilmente se encontraran debido a la diversidad de casos, en general esta probabilidad se puede encontrar usando la definición de acumulada y la f.d.p de la Hipergeométrica como sigue:

$$F_X(x) = P(X \le x) = \sum_{i=\max(0,n-r_2)}^{x} \frac{\binom{r_1}{i}\binom{r-r_1}{n-i}}{\binom{r}{i}}$$

Ejemplos:

1._ Supón que una urna contiene cinco bolas rojas y diez azules. Si se seleccionan al azar sin reemplazo siete bolas.
a) ¿Cuál es la probabilidad de obtener al menos tres bolas rojas?
b) Si $\overline{X}$ representa la proporción de bolas rojas en la muestra ¿Cuáles son la media y la varianza de $\overline{X}$.
<u>**Solución**</u>

$r_1 = 5$ bolas rojas.
$r_2 = 10$ bolas azules.
$n = 7$ tamaño de la muestra.
$r = 15$ total de bolas en la urna.
X: Número de bolas rojas en la muestra.
$X \sim H(x; r_1 = 5, r_2 = 10, n = 7)$

$$P(X \geq 3) = P(X = 3) + P(X = 4) + P(X = 5)$$

$$= \frac{\binom{5}{3}\binom{10}{4} + \binom{5}{4}\binom{10}{3} + \binom{5}{5}\binom{10}{2}}{\binom{15}{7}} = \frac{2100 + 600 + 45}{6435} = \frac{2745}{6435} = \frac{61}{143} = 0.4265$$

b) Observa que $\overline{X} = \frac{X}{n}$ por lo tanto $E(\overline{X}) = E\left(\frac{X}{n}\right) = \frac{1}{n}E(X)$ y $V(\overline{X}) = V\left(\frac{X}{n}\right) = \frac{1}{n^2}V(X)$

$$E(\overline{X}) = \frac{1}{7}\left(\frac{7(5)}{15}\right) = \frac{1}{3}$$

$$V(\overline{X}) = \left(\frac{1}{7}\right)^2\left(\frac{7(5)}{15}\right)\left(1 - \frac{5}{15}\right)\left(1 - \frac{6}{14}\right) = \frac{10(8)}{7(3)(15)(14)} = \frac{8}{441} = 0.018$$

2._ En un almacén se tienen 10 impresoras de las cuales 4 son defectuosas. Una compañía selecciona 5 de las máquinas al azar.
a) ¿Cuál es la probabilidad de que las 5 máquinas sean no defectuosas?
b) La compañía repara las impresoras defectuosas a un costo de 50 dólares cada una. Encuentra la media y la varianza del costo total de reparación.

Solución

$r_1 = 6$ máquinas buenas.
$r_2 = 4$ máquinas defectuosas.
$n = 5$ tamaño de la muestra.
$r = 10$ total de máquinas.

X: Número de máquinas buenas en la muestra.

$X \sim H(x; r_1 = 6, r_2 = 4, n = 5)$

a) $P(X = 5) = \dfrac{\binom{6}{5}\binom{4}{0}}{\binom{10}{5}} = \dfrac{6}{252} = \dfrac{1}{42} = 0.024$

b) Ahora sea la v.a de interés Y: Número de máquinas defectuosas en la muestra.

$$Y \sim H(y; r_1 = 4, r_2 = 6, n = 5)$$

El costo de reparación está dado como $C(Y) = 50Y$.
Por lo tanto:

$$E[C(Y)] = 50E(Y) = 50(5)\frac{4}{10} = 100$$

$$V[C(Y)] = (50)^2 V(Y) = (50)^2 (5)\left(\frac{4}{10}\right)\left(\frac{6}{10}\right)\left(1 - \frac{4}{9}\right) = 1{,}666.66$$

3._ Una corporación muestrea sin reemplazo $n = 3$ empresas para adquirir ciertos suministros. La muestra se selecciona de un conjunto de 6 empresas, de las cuales cuatro son locales y dos no lo son. Sea Y el número de empresas foráneas entre las tres escogidas. Encuentra las siguientes probabilidades:

a) $P(Y = 1)$

b) $P(Y \geq 1)$

c) $P(Y \leq 1)$

<u>Solución</u>

$r_1 = 2$ empresas foráneas.

$r_2 = 4$ $r_2 = 4$ empresas locales.

$n = 3$ tamaño de la muestra.

$r = 6$ total de empresas.

Y: Número de empresas foráneas.

$Y \sim H(y; r_1 = 2, r_2 = 4, n = 3)$

a) $P(Y = 1) = \dfrac{\binom{2}{1}\binom{4}{2}}{\binom{6}{3}} = \dfrac{3}{5} = 0.6$

b) $P(Y \geq 1) = 1 - P(Y = 0) = 1 - \dfrac{\binom{2}{0}\binom{4}{3}}{\binom{6}{3}} = 1 - \dfrac{1}{5} = 0.8$

c) $P(Y \leq 1) = P(Y = 0) + P(Y = 1) = \dfrac{1}{5} + \dfrac{3}{5} = \dfrac{4}{5} = 0.8$

DISTRIBUCIÓN DE POISSON.

La distribución de Poisson a menudo servirá como una distribución de probabilidad apropiada para variables aleatorias tales como el número de llamadas telefónicas que entran a un conmutador en un periodo de tiempo dado, el número de partículas emitidas por una fuente en un periodo de 15 minutos, o el número de defectos en una longitud específica de una tela fabricada. Cada una de estas variables aleatorias representa el número total de ocurrencias de un fenómeno durante un período de tiempo fijo o en una región fija del espacio. Esta v.a tiene una distribución de probabilidad conocida como Poisson con parámetro λ, y su f.d.p está dada por:

$$P(x;\lambda)=\frac{e^{-\lambda}\lambda^{x}}{x!} \quad \text{si } x=0,1,2,\ldots$$

Función generadora de momentos.

$$\psi_X(t)=E\left(e^{tX}\right)=\sum_{x=0}^{\infty}e^{tx}\frac{e^{-\lambda}\lambda^{x}}{x!}=e^{-\lambda}\sum_{x=0}^{\infty}\frac{e^{tx}\lambda^{x}}{x!}=e^{-\lambda}\sum_{x=0}^{\infty}\frac{\left(e^{t}\lambda\right)^{x}}{x!}=e^{-\lambda}e^{e^{t}\lambda} \quad \text{Es decir } \boxed{\psi_X(t)=e^{\lambda(e^{t}-1)}}$$

Nota._ Utilizando el resultado de la serie $e^{\lambda}=\sum_{x=0}^{\infty}\frac{\lambda^{x}}{x!}$

Media y varianza.

$$\mu=\psi'_X(t=0)=e^{\lambda(e^{t}-1)}\lambda e^{t}\Big|_{t=0}=e^{\lambda(1-1)}\lambda e^{0}$$

La media de la v.a Poisson es $\boxed{\mu=\lambda}$ es decir, el parámetro de la distribución es el número promedio de ocurrencias en la unidad de tiempo, área, espacio, etc.

Para enconrar la varianza necesitamos el segundo momento:

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

Sustituyendo en la expresión para la varianza $V(X)=E\left(X^{2}\right)-E^{2}(X)=\lambda+\lambda^{2}-\lambda^{2}$ por lo tanto $\boxed{V(X)=\sigma_X^{2}=\lambda}$

Observación._ Esta distribución tiene la característica de que su media y su varianza son iguales.

Función de probabilidad acumulada.

$F_X(x)=P(X\le x)=\sum_{k=0}^{x}\frac{e^{-\lambda}\lambda^{k}}{k!}$ Existen tablas para estos valores en los anexos de los libros de probabilidad.

Teorema._ Si las variables aleatorias $X_1,X_2,\ldots,X_k$ son independientes y $X_i\sim P(x_i;\lambda_i)$ para $i=1,\ 2,$ ... , k. Entonces la v.a $X=\sum_{i=1}^{k}X_i$ tiene una distribución de Poisson con parámetro $\lambda=\sum_{i=1}^{k}\lambda_i$.

Demostración:

La f.g.m para X_i es $\psi_{X_i}(t) = e^{\lambda_i(e^t-1)}$, para $i = 1,2,...,k$. Debido a que estas variables aleatorias son independientes, tenemos que la f.g.m de la suma de ellas está dada por el producto de las f.g.m. de las variables aleatorias individuales:

$$\psi_X(t) = \prod_{i=1}^{k} \psi_{X_i}(t) = \prod_{i=1}^{k} e^{\lambda_i(e^t-1)} = e^{(e^t-1)\sum_{i=1}^{k}\lambda_i}$$ que se puede escribir como

$\psi_X(t) = e^{\lambda(e^t-1)}$ con parámetro $\lambda = \sum_{i=1}^{k} \lambda_i$.

Ejemplos:

1._ Supón que en un fin de semana concreto el número de accidentes en un cierto cruce tiene una distribución de Poisson con media 0.7. ¿Cuál es la probabilidad de que haya al menos tres accidentes en el cruce durante el fin de semana?

Solución

X : Número de accidentes en un cruce un fin de semana.

$X \sim P(x; \lambda = 0.7)$

$$P(X \geq 3) = 1 - F_X(2) = 1 - e^{-0.7}\left(\frac{(0.7)^0}{0!} + \frac{(0.7)^1}{1!} + \frac{(0.7)^2}{2!}\right) = 1 - 0.966 = 0.034$$

2._ Supón que el número de defectos en un rollo de tela fabricado con un cierto proceso tiene una distribución de Poisson con media 0.4. Si se inspecciona una muestra aleatoria de cinco rollos de tela. ¿Cuál es la probabilidad de que el número total de defectos en los cinco rollos sea al menos 6?

Solución

X_i: Número de defectos en el rollo i.

$X_i \sim P(x_i; \lambda_i = 0.4)$

X: Número de defectos en los 5 rollos.

$X \sim P(x; \lambda = 5(0.4) = 2)$

Ya que $X = X_1 + X_2 + X_3 + X_4 + X_5$

$$P(X \geq 6) = 1 - F_X(5) = 1 - e^{-2}\left(\frac{(2)^0}{0!} + \frac{(2)^1}{1!} + \frac{(2)^2}{2!} + \frac{(2)^3}{3!} + \frac{(2)^4}{4!} + \frac{(2)^5}{5!}\right) = 1 - 0.983 = 0.017$$

3._ 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.

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

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

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

4._ Supón que una fuente radioactiva emite partículas y que el número de las que se emiten durante un período de una hora tiene una distribución de Poisson con parámetro λ. Considera que el instrumento para contar esas emisiones en ocasiones falla al anotar una partícula emitida. Supón específicamente que cualquier partícula emitida tiene una probabilidad p de ser anotada.
a) Si Y está definida como el número de partículas anotadas. ¿Encuentra una expresión para la distribución de probabilidades de Y?
b) Calcula $P(Y = 0)$ si $\lambda = 4$ y $p = 0.9$.

Solución

Sean las variables aleatorias:

X: Número de partículas emitidas en una hora.
Con valores posibles $x = 0,\ 1, 2, ...$
$X \sim P(x; \lambda)$
Y: Número de partículas anotadas.
$y = 0,\ 1, 2, ...$

Para entender el problema observa el siguiente cuadro, si se anotan 3 partículas puede ser que se hayan emitido 3, 4, 5 o más partículas.

Ejemplos de valores posibles de las variables de interés.

X : Partículas emitidas	Y : Partículas anotadas
0	0
1	0, 1
2	0, 1, 2
3	0, 1, 2, **3**
4	0, 1, 2, **3**, 4
5	0, 1, 2, **3**, 4, 5
6	0, 1, 2, **3**, 4, 5, 6
$\vdots$	$\vdots$

Si el número de partículas emitido es fijo, digamos x, entonces la distribución de Y es una binomial dada por

$$P(Y=y \mid X=x)=\binom{x}{y} p^{y} q^{x-y}$$

Debido a que no se pueden emitir 3 y 4 partículas al mismo tiempo tenemos eventos disjuntos así que

$$(Y=y)=(Y=y, X \geq y)=(X=y, Y=y) \cup(X=y+1, Y=y) \cup(X=y+2, Y=y) \cup \cdots$$

Entonces la probabilidad de anotar exactamente y partículas está dada por:

$P(Y=y)=\sum_{x=y}^{\infty} P(X=x, Y=y)$ Usando la ley de la multiplicación se tiene que

$$=\sum_{x=y}^{\infty} P(X=x) P(Y=y \mid X=x)$$

$$=\sum_{x=y}^{\infty} \frac{e^{-\lambda} \lambda^{x}}{x!}\binom{x}{y} p^{y} q^{x-y}$$

$$=e^{-\lambda}\left(\frac{p}{q}\right)^{y} \sum_{x=y}^{\infty} \frac{\lambda^{x}}{x!} \frac{x!}{y!(x-y)!} q^{x}$$

$$=e^{-\lambda}\left(\frac{p}{q}\right)^{y} \frac{1}{y!} \sum_{x=y}^{\infty} \frac{(\lambda q)^{x}}{(x-y)!}$$

$$=e^{-\lambda}\left(\frac{p}{q}\right)^{y} \frac{1}{y!} \sum_{x=y}^{\infty} \frac{(\lambda q)^{(x-y)+y}}{(x-y)!}$$

$$=\frac{e^{-\lambda}}{y!}\left(\frac{\lambda p q}{q}\right)^{y} \sum_{x=y}^{\infty} \frac{(\lambda q)^{x-y}}{(x-y)!}$$

$$=\frac{e^{-\lambda}}{y!}(\lambda p)^{y} \sum_{u=0}^{\infty} \frac{(\lambda q)^{u}}{u!}=\frac{(\lambda p)^{y}}{y!} e^{\lambda(q-1)}$$

Finalmente la f.d.p de la v.a de interés se expresa como:

$\boxed{P(Y=y)=\frac{(\lambda p)^{y} e^{-\lambda p}}{y!}}$ con valores posibles $y=0,\ 1,\ 2,\ \ldots$

b) si $\lambda=4$ y $p=0.9$.

$P(Y=0)=e^{-4(0.9)}=e^{-3.6}=0.027$

5._ Sea Y una v.a de Poisson con media λ. Encuentra $E[Y(Y-1)]$ y usa este resultado para demostrar que $V(Y)=\lambda$.

Solución

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

Observa que $y=(y-2)+2$, entonces

$$E[Y(Y-1)]=e^{-\lambda}\sum_{y=2}^{\infty}\frac{\lambda^{(y-2)+2}}{(y-2)!}=e^{-\lambda}\lambda^2\sum_{y=2}^{\infty}\frac{\lambda^{y-2}}{(y-2)!}=e^{-\lambda}\lambda^2\sum_{z=0}^{\infty}\frac{\lambda^z}{z!}=\lambda^2$$

con $z=y-2$, se tiene que cuando $y=2$ el valor de z es 0.

Y además $E[Y(Y-1)]=E(Y^2)-E(Y)=\lambda^2$

$$\therefore\ E\left(Y^2\right)=\lambda^2+\lambda \quad \text{y} \quad V(Y)=E\left(Y^2\right)-E^2(Y)=\lambda^2+\lambda-\lambda^2=\lambda$$

2.7 Aproximación entre las distribuciones Binomial y Poisson.

Supón que $X\sim b(x;n,p)$ con función de probabilidad dada por

$$P(X=x)=\binom{n}{x}p^x q^{n-x}=\frac{n!}{x!(n-x)!}p^x q^{n-x}=\frac{n(n-1)(n-2)\cdots(n-x+1)(n-x)!}{(n-x)!x!}p^x q^{n-x}$$

Observa que en el numerador del primer factor de esta ecuación hay exactamente x términos. Si hacemos $\lambda=np$ se tiene que $p=\frac{\lambda}{n}$, sustituyendo en la ecuación anterior tenemos que:

$$P(X=x)=\frac{\lambda^x}{x!}\frac{(n-x+1)}{n}\cdots\frac{(n-1)}{n}\frac{n}{n}\left[1-\frac{\lambda}{n}\right]^{n-x}$$

$$=\frac{\lambda^x}{x!}\frac{(n-x+1)}{n}\cdots\frac{(n-1)}{n}\frac{n}{n}\left[1-\frac{\lambda}{n}\right]^{n}\left[1-\frac{\lambda}{n}\right]^{-x}$$

Tomando el límite cuando n tiende a infinito:

$$\lim_{n\to\infty}P(X=x)=\frac{\lambda^x}{x!}\lim_{n\to\infty}\frac{(n-x+1)}{n}\cdots\lim_{n\to\infty}\frac{(n-1)}{n}\lim_{n\to\infty}\frac{n}{n}\lim_{n\to\infty}\left[1-\frac{\lambda}{n}\right]^{n}\lim_{n\to\infty}\left[1-\frac{\lambda}{n}\right]^{-x}$$

$$=\frac{\lambda^x}{x!}(1)\cdots(1)\lim_{n\to\infty}\left[1-\frac{\lambda}{n}\right]^{n}\lim_{n\to\infty}\left[1-\frac{\lambda}{n}\right]^{-x}$$

Recuerda que $(1+x)^k=1+kx+\frac{k(k-1)}{2!}x^2+\cdots+\frac{k(k-1)\cdots(k-(n-1))}{n!}x^n+\cdots$

Entonces $\left(1-\frac{\lambda}{n}\right)^n = 1+n\left(-\frac{\lambda}{n}\right)+\frac{n(n-1)}{2!}\frac{\lambda^2}{n^2}+\cdots+\frac{n(n-1)\cdots(n-(m-1))}{m!}\frac{\lambda^m}{n^m}+\cdots$

$$\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n = 1+(-\lambda)+\frac{(-\lambda)^2}{2!}+\cdots+\frac{(-\lambda)^m}{m!}+\cdots=\sum_{i=0}^{\infty}\frac{(-\lambda)^i}{i!}=e^{-\lambda}$$

Por lo tanto $\lim_{n\to\infty} P(X=x)=e^{-\lambda}\frac{\lambda^x}{x!}$ que es precisamente la función de probabilidad de Poisson con parámetro $\lambda = np$.

Cuando n es grande y p es chica, el valor de la distribución binomial se puede aproximar con el valor de la distribución de Poisson con media $\lambda = np$.

Ejemplos:

1._ Supón que en una gran población la proporción de personas que tienen una cierta enfermedad es 0.01. Se desea encontrar la probabilidad de que en un grupo aleatorio de 200 personas al menos cuatro tengan la enfermedad.

Solución

La v.a de interés es X: Número de personas con la enfermedad.
$X \sim b(x; n=200, p=0.01)$
Como n es grande, podemos aproximar con una distribución de Poisson con parámetro $\lambda = np = 200(0.01)=2$. Entonces la probabilidad buscada está dada por:

$$P(X \geq 4) = 1 - F_X(3) = 1 - 0.8571 = 0.1429$$

2._ Supón que la proporción de personas daltónicas en cierta población es 0.005. ¿Cuál es la probabilidad de que no haya más de una persona daltónica en un grupo de 600 personas seleccionadas aleatoriamente?

Solución

X: Número de personas daltónicas.

$X \sim b(x; n=600, p=0.005)$

Aproximando con una Poisson,
$X \approx P(x; \lambda = np = 600(0.005)=3)$
$P(X \leq 1) = F_X(1) = 0.1992$

3._ La probabilidad de trillizos en nacimientos humanos es aproximadamente 0.001. ¿Cuál es la probabilidad de que haya exactamente un conjunto de trillizos entre 700 nacimientos en un gran hospital?

Solución

X: Número de nacimientos de trillizos.

$X \sim b(x; n = 700, p = 0.001)$

Que se puede aproximar con una Poisson dada por
$X \approx P(x; \lambda = np = 700(0.001) = 0.7)$

$P(X = 1) = 0.7e^{-0.7} = 0.3476$.

4._ 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 $P(\hat{p} \le 0.03)$, 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

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

Entonces

$$P(\hat{p} \le 0.03) = P\left(\frac{X}{n} \le 0.03\right) = P(X \le n(0.03)) = P(X \le 100(0.03)) = P(X \le 3) = F_X(3)$$

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

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

$$P(X \le 3) = \sum_{i=0}^{3} \frac{e^{-\lambda}\lambda^x}{x!} = \sum_{i=0}^{3} \frac{e^{-1}1^x}{x!} = e^{-1}\sum_{i=0}^{3} \frac{1}{x!} = e^{-1}\left(1 + 1 + \frac{1}{2!} + \frac{1}{3!}\right) = e^{-1}\left(\frac{6+6+3+1}{6}\right)$$

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

5._ 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, lo cual se denota como

$X \sim bin(x; n = 30, p = 0.2)$.

Si se aproxima con una Poisson $\lambda = np = 30(0.2) = 6$

$$P(X \leq 3) = \sum_{x=0}^{3} e^{-\lambda} \frac{\lambda^x}{x!} = e^{-\lambda}\left[\frac{\lambda^0}{0!} + \frac{\lambda^1}{1!} + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!}\right]$$

$$= e^{-6}\left[\frac{6^0}{0!} + \frac{6^1}{1!} + \frac{6^2}{2!} + \frac{6^3}{3!}\right] = e^{-6}\left[1 + 6 + 18 + 36\right] = 61e^{-6} = 0.151$$

6._ 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

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

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

$$P(6 \leq X \leq 8) = \sum_{x=6}^{8} \frac{e^{-\lambda}\lambda^x}{x!} = e^{-10}\left(\frac{10^6}{6!} + \frac{10^7}{7!} + \frac{10^8}{8!}\right) = 0.2657$$

3 VARIABLES ALEATORIAS CONTINUAS.

OBJETIVO: Caracterizar las distribuciones de probabilidad de las variables aleatorias continuas más comunes.

3.1 VARIABLES ALEATORIAS CONTINUAS.

Si se tuviera interés en ***medir*** la estatura de una persona, el peso de un artículo, la vida útil de algún aparato electrónico, etc. no se podría descartar ningún número real como posible resultado, una variable de este tipo se llama **continua** y debido a que el número de valores posibles que puede tomar la variable es infinito, la probabilidad de que tome un valor x en particular es cero. Lo cual se escribe como $P(X=x)=0$, $-\infty < x < \infty$.

Observa que:

i) $P(a < X \le b) = P(a < X < b) + P(X = b) = P(a < X < b)$.
ii) $P(a \le X < b) = P(X = a) + P(a < X < b) = P(a < X < b)$.
iii) $P(a \le X \le b) = P(X = a) + P(a < X < b) + P(X = b) = P(a < X < b)$.

Es decir; $P(a < X \le b) = P(a \le X < b) = P(a \le X \le b) = P(a < X < b)$.

3.1.1 FUNCIÓN DE DENSIDAD DE PROBABILIDAD.

Definición: Una **función de densidad continua** (para la v.a continua) es una función no negativa que cumple dos propiedades:

i) $\int_{-\infty}^{\infty} f_X(x)dx = 1$

ii) $P(a < X < b) = \int_a^b f_X(x)dx$

3.1.2 FUNCIÓN DE DISTRIBUCIÓN DE PROBABILIDAD.

Definición: La **función de distribución o función de probabilidad acumulada** para una v.a continua, X con función de densidad $f_X(x)$ está dada por

$$F_X(x) = P(X \le x) = \int_{-\infty}^{x} f_X(t)dt, \qquad -\infty < x < \infty$$

Observación:

i) $f_X(x) = F'_X(x), \quad -\infty < x < \infty$ en todos los puntos donde $f_X(x)$ es continua.

ii) $f_X(x) = 0$ en otro caso.

Ejemplo 1: Una v.a continua X que puede asumir valores entre 1 y 3 tiene una función de densidad de probabilidad dada por $f_X(x) = \frac{1}{2}$.

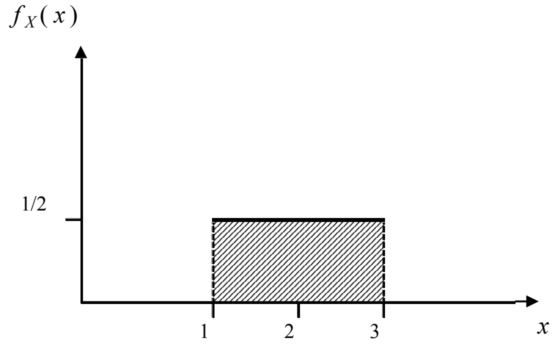
a) Muestra que el área bajo la curva es 1.

b) Encuentra $P(2 < X < 2.5)$.

c) Encuentra la $P(X \leq 1.6)$.

d) Encuentra $F_X(x)$ y úsala para encontrar $P(2 < X < 2.5)$.

Solución



Función de densidad para la v.a X

Observa que esta v.a sólo toma valores de 1 a 3, en los demás puntos la función es cero. El área total bajo la curva debe ser 1, por definición de f.d.p, debido a que la representación gráfica de esta función es una figura geométrica conocida, también podemos encontrar el área bajo la curva con la fórmula del área para el rectángulo, con base igual a 2 y altura igual a $\frac{1}{2}$.

a) $\int_1^3 f_X(x)dx = \int_1^3 \frac{1}{2}dx = \frac{1}{2}x\Big|_1^3 = \frac{1}{2}(3-1) = 1$

b) $P(2 < X < 2.5) = \int_2^{2.5} f_X(x)dx = \int_2^{2.5} \frac{1}{2}dx = \frac{1}{2}x\Big|_2^{2.5} = \frac{2.5-2}{2} = \frac{1}{4}$

c) $P(X \leq 1.6) = \int_1^{1.6} f_X(x)dx = \int_1^{1.6} \frac{1}{2}dx = \frac{1}{2}x\Big|_1^{1.6} = \frac{1.6-1}{2} = 0.3$

d) $F_X(x) = P(X \leq x) = \int_1^{x} f_X(t)dt = \int_1^{x} \frac{1}{2}dt = \frac{1}{2}t\Big|_1^{x} = \frac{x-1}{2}$ para $1 \leq x \leq 3$

$$F_X(x) = \begin{cases} 0 & x < 1 \\ \dfrac{x-1}{2} & 1 \leq x \leq 3 \\ 1 & x > 3 \end{cases}$$

Usando la acumulada se tiene que: $P(2 < X < 2.5) = F_X(2.5) - F_X(2) = \frac{2.5-1}{2} - \frac{2-1}{2} = \frac{0.5}{2} = \frac{1}{4}$

Ejemplo 2: Sea Y una v.a continua con función de densidad de probabilidad

$$f_Y(y) = \begin{cases} 3y^2 & 0 \le y \le 1 \\ 0 & \text{o.c} \end{cases}$$

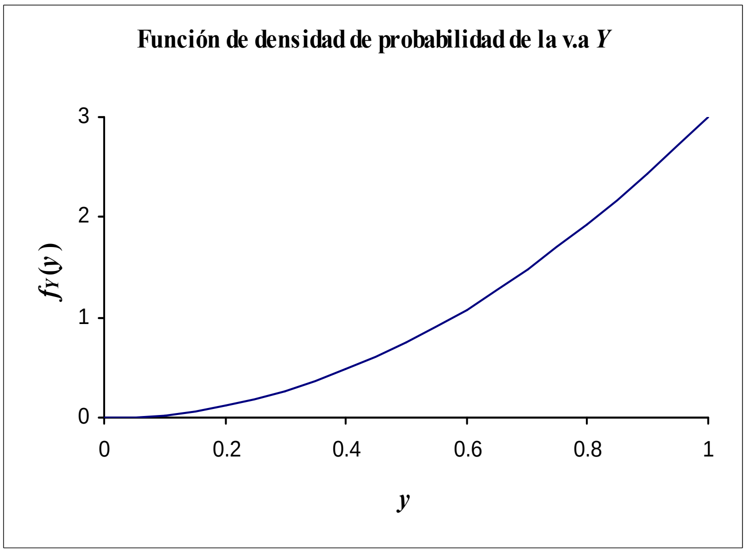
Encuentra la función de probabilidad acumulada y grafica ambas funciones.

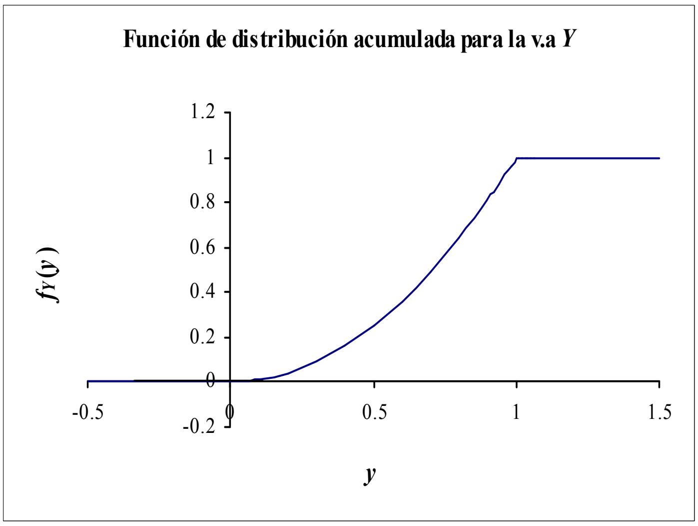
<u>Solución</u>

$$F_Y(y) = \int_0^y f_Y(t)dt = \int_0^y 3t^2 dt = \left.\frac{3t^3}{3}\right|_0^y = y^3$$

Entonces

$$F_Y(y) = \begin{cases} 0 & y < 0 \\ y^3 & 0 \le y \le 1 \\ 1 & y > 1 \end{cases}$$





3.2 Valor esperado y varianza de una variable aleatoria continua.

Las definiciones para media, varianza y f.g.m de una v.a continua son similares al caso discreto, sólo se cambian las sumatorias por integrales.

Media: $\mu_X = E(X) = \int_{-\infty}^{\infty} x f_X(x)dx$

Varianza: $\sigma_X^2 = V(X) = E\left[(X-\mu_X)^2\right] = \int_{-\infty}^{\infty} (x-\mu_X)^2 f_X(x)dx$

Otra manera de expresar la varianza de la v.a X es: $\sigma_X^2 = E\left(X^2\right) - \mu_X^2 = \int_{-\infty}^{\infty} x^2 f_X(x)dx - \mu_X^2$

Función generadora de momentos: $\psi_X(t) = E\left(e^{tX}\right) = \int_{-\infty}^{\infty} e^{tx} f_X(x)dx$

Ejemplo 1: Si la v.a X tiene una f.d. expresada por la siguiente función:

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

a) Demuestra que es una función de densidad de probabilidad.
b) Encuentra la media y la varianza de la v.a X.
c) Encuentra la función de distribución acumulada.

Solución

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

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

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

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

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

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

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

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

Entonces la función de distribución acumulada queda como

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

Ejemplo 2: Si la función de probabilidad de una v.a. continua X es

$$f_X(x)=\begin{cases}ce^{-2x} & x>0\\ 0 & \text{o.c}\end{cases}$$

a) Determina el valor de c.
b) Encuentra $P(1<X<2)$.
Solución

a) Se sabe que $\int_{-\infty}^{\infty} f_X(x)dx=1$

Entonces $\int_0^{\infty} ce^{-2x}dx=c\left(-\frac{1}{2}\right)e^{-2x}\Big|_0^{\infty}=\frac{c}{2}=1 \quad \therefore \quad c=2$

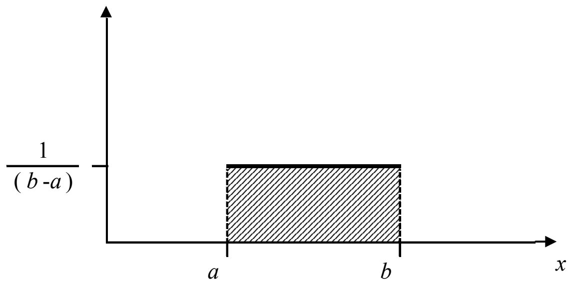
b) $P(1<X<2)=\int_1^2 2e^{-2x}dx=-\int_1^2 e^{-2x}(-2)dx=-e^{-2x}\Big|_1^2=-e^{-4}+e^{-2}=0.117$

DISTRIBUCIÓN UNIFORME.

La más sencilla de las distribuciones continuas es la uniforme, la cual recibe su nombre debido a que es constante en todo el intervalo donde existe. Se dice que la v.a continua X, tiene una distribución uniforme si su **f.d.p** está dada por la expresión:

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

Dado que el área bajo la función es la probabilidad total, podemos encontrar esta área con la fórmula conocida para un rectángulo, con base igual a $(b-a)$ y altura $\left(\frac{1}{b-a}\right)$, con lo cual se obtiene la unidad. La gráfica de esta distribución es la siguiente:



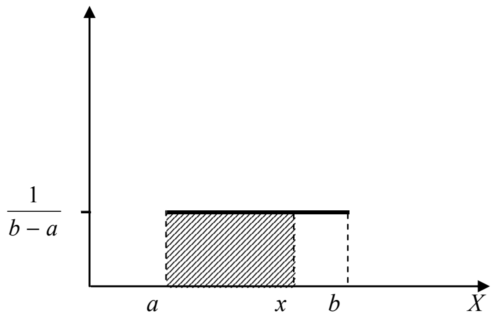
Distribución Uniforme en el intervalo (a , b)

La notación para esta distribución es $X \sim U(a,b)$ y se lee "la v.a X se distribuye uniformemente en el intervalo (a,b).

La **función de probabilidad acumulada** para esta distribución está dada por:

$$F_X(x) = \int_a^x \frac{dt}{b-a} = \left.\frac{t}{b-a}\right|_a^x = \frac{x-a}{b-a} \quad \text{si } a < x < b.$$

Observa que es la fórmula para encontrar el área del rectángulo delimitado por las rectas $x = a$, $x = x$, $y = 0$ y $y = \frac{1}{b-a}$.



Función de distribución acumulada en x para la distribución uniforme en el intervalo (a, b).

Valor esperado, varianza y f.g.m.

$$\mu_X = E(X) = \int_a^b x f_X(x)dx = \int_a^b x\left(\frac{1}{b-a}\right)dx = \left.\frac{x^2}{2(b-a)}\right|_a^b = \frac{b^2-a^2}{2(b-a)} = \frac{(b+a)(b-a)}{2(b-a)}$$

$$\boxed{\mu_X = \frac{a+b}{2}}$$

Para el cálculo de la varianza necesitamos el segundo momento:

$$E\left(X^2\right)=\int_a^b \frac{x^2dx}{b-a}=\left.\frac{x^3}{3(b-a)}\right|_a^b=\frac{b^3-a^3}{3(b-a)}=\frac{\left(b^2+ab+a^2\right)(b-a)}{3(b-a)}=\frac{a^2+ab+b^2}{3}$$

Por lo tanto la varianza es:

$$\sigma^2=V(X)=E\left(X^2\right)-E^2(X)=\frac{a^2+ab+b^2}{3}-\left(\frac{a+b}{2}\right)^2$$

$$=\frac{4a^2+4ab+4b^2-3a^2-6ab-3b^2}{12}=\frac{a^2+b^2-2ab}{12}$$

$$\boxed{\sigma^2=\frac{(a-b)^2}{12}}$$

La **f.g.m** es de la forma:

$$\psi_X(t)=E\left(e^{tX}\right)=\int_a^b \frac{e^{tx}dx}{b-a}=\left.\frac{e^{tx}}{t(b-a)}\right|_a^b=\frac{e^{bt}-e^{at}}{t(b-a)} \qquad \boxed{\psi_X(t)=\frac{e^{bt}-e^{at}}{t(b-a)}}$$

Ejemplo 1: Si X se distribuye uniformemente, es simétrica respecto al origen y tiene varianza 1. ¿Cuánto vale a y cuánto b?

__Solución__

Por ser simétrica respecto al origen se tiene que si $b=k$, entonces $a=-k$. Además la información que se tiene es que: $V(X)=\frac{(a-b)^2}{12}=\frac{(-k-k)^2}{12}=1 \;\therefore\; 12=4k^2$

es decir; $k=\sqrt{3}$ y por lo tanto $a=-\sqrt{3}$ y $b=\sqrt{3}$.

Ejemplo 2: Si $X\sim U(a=0,b=4)$. ¿Cuál es la probabilidad de que las raíces de $y^2+4xy+x+1=0$ sean reales?

__Solución__

Las raíces de la ecuación cuadrática $y^2+(4x)y+(x+1)=0$, se sacan usando la fórmula general con $a=1$, $b=4x$ y $c=x+1$.

$$x_{1,2}=\frac{-b\pm\sqrt{b^2-4ac}}{2a}=\frac{-4x\pm\sqrt{16x^2-4(x+1)}}{2}$$

Para que las raíces sean reales se necesita que el radicando sea mayor o igual a cero, es decir:

$$16x^2-4(x+1)\ge 0$$

Dividiendo entre 4, se tiene que $4x^2-x-1\ge 0$

Esta última nuevamente es una ecuación de segundo grado pero ahora en x, con raíces dadas por;

$$x_{1,2} = \frac{1 \pm \sqrt{1-4(4)(-1)}}{2(4)} = \frac{1 \pm \sqrt{17}}{8} = \begin{cases} 0.64 \\ -0.39 \end{cases}$$

Observa que la función es positiva en el intervalo (0.64, 4). La probabilidad de que las raíces sean reales coincide con la probabilidad de que X esté en ese intervalo.
Dado que $X \sim U(0,4)$ procedemos como sigue:

$$P(0.64 < X < 4) = \int_{0.64}^{4} \frac{dx}{4} = \left.\frac{x}{4}\right|_{0.64}^{4} = \frac{4-0.64}{4} = 0.84$$

3.3.1 Distribución Exponencial.

Se utiliza a menudo para representar la distribución del tiempo que transcurre antes de la ocurrencia de un suceso. Por ejemplo; período que un componente electrónico funciona sin fallar, también conocido como vida útil del componente, período requerido para atender a un cliente en un banco, período entre las llegadas de dos clientes sucesivos a un servicio, etc.
La v.a X tiene una distribución exponencial con parámetro $\beta > 0$, y se escribe $X \sim \exp(x;\beta)$ si su **f.d.p** es de la forma:

$$f_X(x) = \begin{cases} \beta e^{-\beta x} & x > 0 \\ 0 & x \le 0 \end{cases}$$

La **función de distribución acumulada** de la v.a es;

$$F_X(x) = \int_0^x \beta e^{-\beta t}\,dt = \int_0^x e^{-\beta t}\beta dt = -\left.e^{-\beta t}\right|_0^x = -e^{-\beta x} + e^{-0} = 1 - e^{-\beta x}$$

Para encontrar media y varianza de esta distribución, primero encontramos la **f.g.m**,

$$\psi_X(t) = E\left(e^{tX}\right) = \int_0^\infty e^{tx}\beta e^{-\beta x}\,dx = \beta\int_0^\infty e^{(t-\beta)x}\,dx = \left.\frac{\beta}{t-\beta}e^{-(\beta-t)x}\right|_0^\infty$$

$$\boxed{\psi_X(t) = \frac{\beta}{\beta - t}}$$

$$\mu_X = E(X) = \psi'_X(t=0) = \frac{d}{dt}\left[\beta(\beta-t)^{-1}\right]\Big|_{t=0} = \beta(\beta-t)^{-2}\Big|_{t=0} = \frac{\beta}{\beta^2}$$

$$\boxed{\mu_X = \frac{1}{\beta}}$$

$$E\left(X^2\right) = \psi''_X(t=0) = \frac{d}{dt}\left[\beta(\beta-t)^{-2}\right]\Big|_{t=0} = 2\beta(\beta-t)^{-3}\Big|_{t=0} = \frac{2\beta}{\beta^3} = \frac{2}{\beta^2}$$

$$\sigma^2 = V(X) = E\left(X^2\right) - E^2(X) = \frac{2}{\beta^2} - \frac{1}{\beta^2} \quad \therefore \quad \boxed{\sigma^2 = \frac{1}{\beta^2}}$$

Esta distribución también tiene la propiedad de falta de memoria, es decir para los enteros positivos t y h se cumple que:

$$P(X \geq t+h \mid X \geq t) = \frac{P(X \geq t+h)}{P(X \geq t)} = \frac{1-F_X(t+h)}{1-F_X(t)} = \frac{1-(1-e^{-t\beta-h\beta})}{1-(1-e^{-t\beta})} = \frac{e^{-t\beta}e^{-h\beta}}{e^{-t\beta}} = e^{-t\beta}$$

Y por otro lado $P(X \geq h) = 1 - F_X(h) = 1 - \left(1 - e^{-h\beta}\right) = e^{-h\beta}$

Entonces $\boxed{P(X \geq t+h \mid X \geq t) = P(X \geq h)}$

Teorema: Si las variables aleatorias $X_1, X_2, \ldots, X_n$ son variables aleatorias independientes con distribución exponencial con parámetro β. Entonces la v.a $Y = \min(X_1, X_2, \ldots, X_n)$ también se distribuye como una exponencial pero con parámetro $n\beta$.

<u>Demostración</u>

Observa que si el tiempo mínimo de espera de los n que tenemos es mayor que t entonces todos los demás también son mayores que t, es decir:

$$P(Y > t) = P(X_1 > t, \ldots, X_n > t) = P(X_1 > t) \cdots P(X_n > t) = e^{-t\beta} \cdots e^{-t\beta} = e^{-nt\beta}$$

Por lo tanto $Y \sim \exp(y; n\beta)$. ♦

Ejemplo 1: El motor y el tren de transmisión de un automóvil nuevo están garantizados por un año. Las vidas medias de estos componentes se estiman en tres años, y el tiempo transcurrido hasta la falla tiene una distribución exponencial. La ganancia en un auto nuevo es de $1,000. Incluyendo los costos de refacciones y de mano de obra, la agencia debe pagar $250 para reparar la falla. ¿Cuál es la utilidad esperada por automóvil?

<u>Solución</u>

La vida media del motor y del tren de transmisión es de 3 años, es decir $E(X) = \frac{1}{\beta} = 3$, despejando se tiene que el parámetro de la distribución es $\frac{1}{3}$.

X : tiempo transcurrido hasta la falla.

$X \sim \exp\left(x; \beta = \frac{1}{3}\right)$

$U(X)$: utilidad en función del tiempo transcurrido hasta la falla.

$$U(X) = \begin{cases} 1000 & X > 1 \text{ año} \\ 1000 - 250 = 750 & X \leq 1 \text{ año} \end{cases}$$

La utilidad esperada está dada por:

$$\begin{aligned}E[U(X)] &= 1000P\,[U(X)=1000]+750P\,[U(X)=750]\\ &= 1000P(X>1)+750P(X\le 1) = 1000[1-F_X(1)]+750F_X(1)\\ &= 1000-F_X(1)[1000-750] = 1000-250F_X(1)\\ &= 1000-250\left(1-e^{-\frac{1}{3}}\right) = 750+250e^{-\frac{1}{3}} = 929.13\end{aligned}$$

Ejemplo 2: ¿Hay una densidad exponencial que cumple la siguiente condición? $P(X\le 2)=\frac{2P(X\le 3)}{3}$. Si es así encuentra el valor del parámetro de la distribución.

Solución

$$P(X\le 2)=F_X(2)=\frac{2F_X(3)}{3}=\frac{2P(X\le 3)}{3}$$
$$1-e^{-2\beta}=\frac{2}{3}\left(1-e^{-3\beta}\right)$$
$$3\left(1-e^{-2\beta}\right)=2\left(1-e^{-3\beta}\right)$$
$$3-3e^{-2\beta}=2-2e^{-3\beta}$$
$$3-2=3e^{-2\beta}-2e^{-3\beta}$$
$$1=3e^{-2\beta}-2e^{-3\beta}$$

Lo cual ocurre sólo si $\beta=0$, pero una de las condiciones de la definición de la f.d.p exponencial es que el parámetro debe ser estrictamente mayor que cero, es decir no existe una distribución exponencial que cumpla la condición pedida.

Ejemplo 3: Se sabe que un componente electrónico tiene una vida útil representada por una densidad exponencial con tasa de falla de 10^{-5} fallas por hora. ¿Cuál es la fracción de componentes que fallan antes de su vida media?

Solución

La fracción de componentes de interés coincide con la probabilidad de que un componente falle antes de su vida esperada, por lo tanto procedemos como sigue:

X : vida útil del componente.
$X\sim \exp\left(x;10^{-5}\right)$

La tasa de falla es $\beta=10^{-5}$ por lo tanto $E(X)=\frac{1}{\beta}=10^{5}$.

Entonces $P\left(X<\frac{1}{\beta}\right)=P\left(X<10^{5}\right)=1-e^{-10^{-5}(10^{5})}=1-e^{-1}=0.632$

Ejemplo 4: Supón que se quiere decidir entre dos procesos de manufactura para fabricar cierto componente. Con el proceso *A*, producir un componente cuesta *C* dólares, con el proceso *B*, producir un componente cuesta kC dólares, con $k > 1$. El tiempo de vida de los componentes se distribuye exponencialmente con una tasa de $\frac{1}{200}$ fallas por hora para el proceso *A*. Los componentes producidos con el proceso *B* tienen una vida útil distribuida exponencialmente con una tasa de $\frac{1}{300}$ fallas por hora. Debido a una cláusula de garantía, si un componente dura menos de 400 hrs, el fabricante debe pagar una multa de M dólares. ¿Cuál proceso escogerías?

<u>**Solución**</u>

X : Vida útil del componente.
C_i : Costo bajo el proceso *i*, con $i = A, B$.

$$X_A \sim \exp\left(x_A; \beta_A = \frac{1}{200}\right)$$
$$X_B \sim \exp\left(x_B; \beta_B = \frac{1}{300}\right)$$

$$C_A = \begin{cases} C & X \geq 400 \text{ hrs.} \\ C + M & X < 400 \text{ hrs.} \end{cases}$$

$$C_B = \begin{cases} kC & X \geq 400 \text{ hrs.} \\ kC + M & X < 400 \text{ hrs.} \end{cases}$$

$$\begin{aligned} E(C_A) &= CP(X \geq 400) + (C+M)P(X < 400) \\ &= C(1 - F_X(400)) + (C+M)F_X(400) \\ &= C - CF_X(400) + (C+M)F_X(400) \\ &= C + [C + M - C]F_X(400) = C + MF_X(400) \end{aligned}$$
$$E(C_A) = C + M\left(1 - e^{-\frac{400}{200}}\right) = C + M\left(1 - e^{-2}\right)$$

Para el proceso *B* se tiene que:

$$\begin{aligned} E(C_B) &= kCP(X \geq 400) + (kC+M)P(X < 400) \\ &= kC(1 - F_X(400)) + (kC+M)F_X(400) \\ &= kC + MF_X(400) = kC + M\left(1 - e^{-\frac{400}{300}}\right) = kC + M\left(1 - e^{-1.33}\right) \end{aligned}$$

M es una multa fija que pagan ambos procesos en caso de que la vida útil del componente sea menor a 400 hrs. Si el costo esperado de producción bajo el proceso *A* es menor que el costo esperado bajo el proceso *B*, se prefiere el proceso *A*, lo cual ocurre si:

$$C + M\left(1 - e^{-2}\right) < kC + M\left(1 - e^{-1.33}\right)$$
$$C + M\left(1 - e^{-2}\right) - M\left(1 - e^{-1.33}\right) < kC$$
$$\frac{1}{C}\left[C + M\left(1 - e^{-2} - 1 + e^{-1.33}\right)\right] < k$$
$$1 + \frac{M}{C}\left(e^{-1.33} - e^{-2}\right) < k$$
$$1 + 0.129\frac{M}{C} < k$$

Ejemplo 5: Supón que las variables aleatorias $X_1, X_2, \ldots, X_k$ son independientes y que $X_i \sim \exp(x_i;\beta_i)$ con $i = 1,...,k$. Si $Y = \min(X_1, X_2, \ldots, X_k)$ ¿Cómo se distribuye *Y*?

Solución

$X_i \sim \exp(x_i;\beta_i)$, entonces para $t > 0$

$$\begin{aligned} P(Y > t) &= P(X_1 > t, X_2 > t, \ldots, X_k > t) \\ &= P(X_1 > t)P(X_2 > t)\cdots P(X_k > t) \\ &= \left(1 - F_{X_1}(t)\right)\left(1 - F_{X_2}(t)\right)\cdots\left(1 - F_{X_k}(t)\right) = e^{-t\beta_1}e^{-t\beta_2}\cdots e^{-t\beta_k} = e^{-[\beta_1+\beta_2+\cdots+\beta_k]t} \end{aligned}$$

Por lo tanto $Y \sim \exp\left(y;\beta = \sum_{i=1}^{k}\beta_i\right)$.

Ejemplo 6: Supón que cierto sistema contiene tres componentes que funcionan independientemente unos de otros y que están conectados en serie de forma que el sistema falla tan pronto como uno de los componentes falla. Si el tiempo de vida de cada uno de los componentes medido en horas se distribuye exponencialmente con parámetros 0.001, 0.003 y 0.006 respectivamente. Determina la probabilidad de que el sistema no falle antes de 100 horas.

Solución

X_i: tiempo de vida del *i*-ésimo componente.
$X_1 \sim \exp(x_1;0.001)$
$X_2 \sim \exp(x_2;0.003)$
$X_3 \sim \exp(x_3;0.006)$
X : tiempo de vida del sistema.

Debido a la construcción del sistema, su tiempo de vida es $X = \min(X_1, X_2, X_3)$. Entonces $X \sim \exp(x;0.01)$ ya que $0.001 + 0.003 + 0.006 = 0.01$.
Por lo tanto $P(X > 100) = 1 - F_X(100) = 1 - \left(1 - e^{-100(0.01)}\right) = e^{-1} = 0.3678$

Ejemplo 7: Si un sistema electrónico contiene *n* componentes similares que funcionan independientemente unos de otros y están conectados en serie (el sistema falla tan pronto como uno de los componentes falla). Si el tiempo de vida de cada componente medido en horas, tiene una distribución exponencial con media μ. Encuentra la media y la varianza del tiempo de espera hasta que falle el sistema.

Solución

X_i : tiempo de vida del *i*-ésimo componente.

$X_i \sim \exp\left(x_i; \frac{1}{\mu}\right)$ con $i = 1,\ldots,n$.

X : tiempo de vida del sistema.

$X = \min(X_1, \ldots, X_n)$

$X \sim \exp\left(x; \frac{n}{\mu}\right)$

Entonces; $E(X) = \frac{1}{\beta} = \frac{\mu}{n}$ y $V(X) = \frac{1}{\beta^2} = \frac{\mu^2}{n^2}$

3.3.2 DISTRIBUCIÓN GAMMA.

Para esta distribución necesitamos recordar la función gamma por lo que es indispensable comenzar con la siguiente:

Definición: La función gamma para $\alpha > 0$ se define como $\boxed{\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx}$

Nota: Para que la integral converja se requiere $x > 0$.

La función gamma tiene las siguientes propiedades:

1) Si $\alpha > 1$ entonces $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$

2) Si $n \in Z^+$ entonces $\Gamma(n) = (n-1)!$

Por este motivo también se conoce a esta función como función factorial generalizada.

Demostración

1) $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$

Integrando por partes con $u = x^{\alpha-1}$, $du = (\alpha-1)x^{\alpha-2}dx$, $dv = e^{-x}dx$ y $v = -e^{-x}$

$$\Gamma(\alpha) = -x^{\alpha-1}e^{-x}\Big|_0^\infty + \int_0^\infty e^{-x}(\alpha-1)x^{\alpha-2}dx = (\alpha-1)\int_0^\infty x^{(\alpha-1)-1}e^{-x}dx = (\alpha-1)\Gamma(\alpha-1) \quad \blacklozenge$$

2) Aplicando repetidas veces la propiedad uno se tiene que:

$$\Gamma(n) = (n-1)\Gamma(n-1) = (n-1)(n-2)\Gamma(n-2) = \cdots = (n-1)(n-2)\cdots\Gamma(1)$$

Pero $\Gamma(1) = \int_0^\infty x^{1-1}e^{-x}dx = \int_0^\infty e^{-x}dx = -e^{-x}\Big|_0^\infty = 1$ Entonces $\Gamma(n) = (n-1)!$ $\blacklozenge$

Los intervalos de tiempo entre dos fallas en un proceso de producción, los intervalos de tiempo entre dos llegadas a la caja en un banco, en general los tiempos entre dos ocurrencias de un evento son variables aleatorias continúas que se pueden modelar con la función de densidad tipo Gamma. Se dice que la v.a X tiene una distribución de probabilidad Gamma con parámetros $\alpha > 0$ y $\beta > 0$ y se escribe $X \sim Gamma(x;\alpha,\beta)$ o equivalentemente $X \sim \Gamma(x;\alpha,\beta)$ si su **f.d.p** está definida como:

$$f_X(x) = \begin{cases} \dfrac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} & x > 0 \\ 0 & \text{o.c} \end{cases}$$

La f.g.m para esta distribución está dada por:

$$\psi_X(t) = E\left(e^{tX}\right) = \int_0^\infty e^{tx} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} dx = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty x^{\alpha-1} e^{-(\beta-t)x} dx$$

Tomando $u = (\beta - t)x$ se tiene que $x = \dfrac{u}{\beta - t}$ y que $du = (\beta - t)dx$ entonces $dx = \dfrac{du}{\beta - t}$

Sustituyendo en la integral:

$$\psi_X(t) = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty \left(\frac{u}{\beta - t}\right)^{\alpha-1} e^{-u} \frac{du}{\beta - t} = \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{1}{(\beta-t)^{\alpha-1}} \frac{1}{(\beta-t)} \int_0^\infty u^{\alpha-1} e^{-u} du = \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha)}{(\beta-t)^\alpha}$$

$$\boxed{\psi_X(t) = \left(\frac{\beta}{\beta - t}\right)^\alpha}$$

Media y varianza.

$$\mu_X = E(X) = \psi'_X(t=0) = \alpha\left(\frac{\beta}{\beta - t}\right)^{\alpha-1} \beta(\beta - t)^{-2}\Bigg|_{t=0} \quad \therefore \boxed{\mu_X = \frac{\alpha}{\beta}}$$

El segundo momento de la variable es:

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

Entonces $\sigma^2 = V(X) = E\left(X^2\right) - E^2(X) = \dfrac{\alpha(\alpha+1)}{\beta^2} - \dfrac{\alpha^2}{\beta^2} = \dfrac{\alpha^2 + \alpha - \alpha^2}{\beta^2}$

$$\boxed{\sigma^2 = \frac{\alpha}{\beta^2}}$$

Ejercicio: Si α es un entero positivo, encuentra la probabilidad de que la v.a $X \sim Gamma(x;\alpha,\beta)$ sea mayor que el número x.

Solución

$$P(X > x) = \int_x^\infty \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} dx = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_x^\infty x^{\alpha-1} e^{-\beta x} dx$$

Integrando por partes con $u = x^{\alpha-1}$, $du = (\alpha-1)x^{\alpha-2}dx$, $dv = e^{-\beta x}dx$ y $v = \dfrac{-e^{-\beta x}}{\beta}$

Y tomando $A = \dfrac{\beta^\alpha}{\Gamma(\alpha)}$, tenemos que:

$$P(X > x) = A\left[-\frac{x^{\alpha-1}e^{-\beta x}}{\beta}\Big|_x^\infty + \int_x^\infty \frac{e^{-\beta x}}{\beta}(\alpha-1)x^{\alpha-2}dx \right]$$

$$= A\left[\frac{x^{\alpha-1}e^{-\beta x}}{\beta} + \frac{\alpha-1}{\beta}\int_x^\infty x^{\alpha-2}e^{-\beta x}dx \right]$$

Integrando nuevamente por partes con:

$$u = x^{\alpha-2}, \qquad du = (\alpha-2)x^{\alpha-3}dx$$

$$dv = e^{-\beta x}dx, \qquad v = \frac{-e^{-\beta x}}{\beta}$$

$$uv - \int vdu = \frac{x^{\alpha-2}e^{-\beta x}}{\beta} + \frac{\alpha-2}{\beta}\int_x^\infty x^{\alpha-3}e^{-\beta x}dx$$

$$= A\left[\frac{x^{\alpha-1}e^{-\beta x}}{\beta} + \frac{(\alpha-1)x^{\alpha-2}e^{-\beta x}}{\beta^2} + \frac{(\alpha-1)(\alpha-2)}{\beta^2}\int_x^\infty x^{\alpha-3}e^{-\beta x}dx \right]$$

$$= A\left\{ \frac{x^{\alpha-1}e^{-\beta x}}{\beta} + \frac{(\alpha-1)x^{\alpha-2}e^{-\beta x}}{\beta^2} + \frac{(\alpha-1)(\alpha-2)}{\beta^2}\left[\frac{x^{\alpha-3}e^{-\beta x}}{\beta} + \frac{\alpha-3}{\beta}\int_x^\infty x^{\alpha-4}e^{-\beta x}dx \right]\right\}$$

$$= A\left\{ \frac{x^{\alpha-1}e^{-\beta x}}{\beta} + \frac{(\alpha-1)x^{\alpha-2}e^{-\beta x}}{\beta^2} + \frac{(\alpha-1)(\alpha-2)x^{\alpha-3}e^{-\beta x}}{\beta^3} + \frac{(\alpha-1)(\alpha-2)(\alpha-3)}{\beta^3}\int_x^\infty x^{\alpha-4}e^{-\beta x}dx \right\}$$

$$= A\left\{ \frac{x^{\alpha-1}e^{-\beta x}}{\beta} + \frac{(\alpha-1)x^{\alpha-2}e^{-\beta x}}{\beta^2} + \cdots + \frac{(\alpha-1)(\alpha-2)\cdots(\alpha-(\alpha-1))}{\beta^{\alpha-1}}\int_x^\infty x^{\alpha-\alpha}e^{-\beta x}dx \right\}$$

Pero $\displaystyle\int_x^\infty e^{-\beta x}dx = \frac{-e^{-\beta x}}{\beta}\Big|_x^\infty = \frac{e^{-\beta x}}{\beta}$

Entonces:

$$= \frac{\beta^{\alpha}}{\Gamma(\alpha)}\left[\frac{x^{\alpha-1}e^{-\beta x}}{\beta} + \frac{(\alpha-1)x^{\alpha-2}e^{-\beta x}}{\beta^2} + \frac{(\alpha-1)(\alpha-2)x^{\alpha-3}e^{-\beta x}}{\beta^3} + \cdots + \frac{(\alpha-1)(\alpha-2)\cdots(1)}{\beta^{\alpha-1}}\frac{e^{-\beta x}}{\beta}\right]$$

Sustituyendo el valor de *A*, llegamos a:

$$P(X > x) = \frac{e^{-\beta x}(\beta x)^{\alpha-1}}{(\alpha-1)!} + \frac{e^{-\beta x}(\beta x)^{\alpha-2}}{(\alpha-2)!} + \frac{e^{-\beta x}(\beta x)^{\alpha-3}}{(\alpha-3)!} + \cdots + e^{-\beta x}$$

$$\boxed{P(X > x) = \sum_{k=0}^{\alpha-1}\frac{e^{-\beta x}(\beta x)^k}{k!}} \quad (*)$$

Recuerda que la f.d.p de la Poisson es de la forma $f_X(x) = \frac{e^{-\lambda}\lambda^x}{x!}$

Entonces, (*) es precisamente la probabilidad acumulada de la Poisson con parámetro $\lambda = \beta x$, evaluada en el número $(\alpha - 1)$. Que escribimos como $P(X > x) = F_X^P(\alpha - 1)$.

Finalmente se tiene que la función de distribución acumulada para la Gamma, está dada por:

$$\boxed{F_X(x) = P(X \le x) = 1 - P(X > x) = 1 - F_X^P(\alpha - 1)}$$

Observa que tomando $\alpha = 1$ en la distribución Gamma, $\Gamma(x;\alpha,\beta) = \frac{\beta^{\alpha}}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x} = \beta e^{-\beta x} = \exp(x;\beta)$ se obtiene la f.d.p exponencial con parámetro β.

Ejemplo 1: Si las variables aleatorias $X_1, X_2, \ldots, X_\alpha$ tienen una distribución exponencial con parámetro β y son independientes unas de otras. ¿Cómo se distribuye la v.a $X = \sum_{i=1}^{\alpha} X_i$?

Solución

Como las variables aleatorias son independientes, se usa la propiedad de que la f.g.m de una suma de variables aleatorias independientes es el producto de las f.g.m de cada una de las variables aleatorias que se están sumando;

$\psi_X(t) = \prod_{i=1}^{\alpha}\psi_{X_i}(t) = \prod_{i=1}^{\alpha}\frac{\beta}{\beta - t} = \left(\frac{\beta}{\beta - t}\right)^{\alpha}$ es la f.g.m para una Gamma con parámetros α y β.

Entonces, *suma de α variables aleatorias, independientes distribuidas exponencialmente con parámetro β es una* $\Gamma(x;\alpha,\beta)$.

Ejemplo 2: Si las variables aleatorias $X_1, X_2, \ldots, X_k$ son independientes y $X_i \sim \Gamma(x_i; \alpha_i, \beta)$ para $i = 1,\ldots,k$. Encuentra la distribución de la v.a $X = \sum_{i=1}^{k} X_i$.

Solución

$$\psi_{X_i}(t) = \left(\frac{\beta}{\beta - t}\right)^{\alpha_i}$$

$$\psi_X(t) = \prod_{i=1}^{k} \psi_{X_i}(t) = \left(\frac{\beta}{\beta - t}\right)^{\sum_{i=1}^{k} \alpha_i}$$

Entonces $X \sim \Gamma\left(x; \alpha = \sum_{i=1}^{k} \alpha_i, \beta\right)$.

Ejemplo 3: Un sistema redundante opera de la siguiente manera. Cuando la unidad 1 falla el tablero de decisiones pone la unidad 2 hasta que falla y entonces activa la unidad 3. El interruptor de decisión se supone perfecto, por lo que la vida del sistema puede representarse como la suma de las vidas de los subsistemas. Si las vidas de los subsistemas son independientes entre sí y cada subsistema tiene una vida X_j, $j = 1, 2, 3$, con densidad $g(x_j) = \dfrac{e^{-\frac{x_j}{100}}}{100}$, $x_j \geq 0$. Encuentra la función de confiabilidad $R(x)$ del sistema, donde $R(x) = P(\text{El sistema opere al menos } x \text{ horas}) = P(X \geq x) = 1 - F_X(x)$.

Solución

X_i: Vida del subsistema *i* con $i = 1, 2, 3$.
X : Vida del sistema.
$X = X_1 + X_2 + X_3$.

Si $X_i \sim \exp\left(x_i; \dfrac{1}{100}\right)$, entonces $X \sim \Gamma\left(x; \alpha = 3, \beta = \dfrac{1}{100}\right)$ y por lo tanto la confiabilidad del sistema es:

$$R(x) = 1 - \left(1 - \sum_{k=0}^{\alpha-1} \frac{e^{-\beta x} (\beta x)^k}{k!}\right) = \sum_{k=0}^{2} \frac{e^{-0.01x} (0.01x)^k}{k!} = e^{-0.01x}\left(1 + 0.01x + 0.00005x^2\right)$$

Ejemplo 4: Una caja de caramelos contiene 24 barras. El tiempo entre pedidos por barra se distribuye exponencialmente con media 10 minutos. Supón que se vende el primer caramelo en cuanto se abre la caja. ¿Cuál es la probabilidad de que una caja abierta a las 8:00 A.M. se haya terminado al medio día?

<u>**Solución**</u>

X_i : Tiempo entre pedidos $i = 1,\ 2,\ldots,23$.
X : tiempo en que se acaba la caja.

$X = \sum_{i=1}^{23} X_i$, pero $X_i \sim \exp(x_i;6)$ ya que 1 hora tiene 60 minutos.
Hay independencia entre pedidos, y de las 8:00 a las 12:00 hay 4 horas. Se desea encontrar la $P(X \leq 4)$.
Sabemos que $X \sim \Gamma(x;\alpha = 23, \beta = 6)$, y que la distribución acumulada de probabilidad de la distribución Gamma queda en términos de la distribución acumulada de una Poisson con parámetro $\lambda = \beta x = 6(4) = 24$, así que:

$$F_X(4) = 1 - \sum_{k=0}^{22} \frac{e^{-24}(24)^k}{k!} = 1 - F_X^P(22) = 1 - 0.3917 = 0.6083$$

Ejemplo 5: El tiempo de reabastecimiento de cierto producto cumple con la distribución Gamma con media 40 y varianza 400. Determina la probabilidad de que un pedido se envíe dentro de los 20 días posteriores a su solicitud.

<u>**Solución**</u>

X : Tiempo de reabastecimiento del producto en días.
$X \sim \Gamma(x;\alpha = 4, \beta = 0.1)$

$$\mu = \frac{\alpha}{\beta} = 40 \text{ y } \sigma^2 = \frac{\alpha}{\beta^2} = \left(\frac{\alpha}{\beta}\right)\frac{1}{\beta} = \frac{40}{\beta} = 400$$
Entonces $\alpha = 4$ y $\beta = 0.1$

$P(X < 20) = F_X(20) = 1 - F_X^P(3)$ con $\lambda = \beta x = 0.1(20) = 2$
$P(X < 20) = 0.1433$

3.4 DISTRIBUCIÓN NORMAL.

Es una de las distribuciones más importantes en estadística, es simétrica respecto a su media y casi el 98% de la distribución se encuentra entre $(\mu - 2\sigma, \mu + 2\sigma)$, en gran parte de las aplicaciones se supone normalidad en las variables.

Se dice que la v.a X tiene una distribución normal o Gaussiana con parámetros μ y σ^2 y se escribe $X \sim N(x;\mu,\sigma^2)$ si su **f.d.p** es de la forma:

$$f_X(x) = \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}} \quad \text{con} -\infty < x < \infty,\ -\infty < \mu < \infty,\ \sigma^2 > 0$$

Ejercicio: Demuestra que $\int_{-\infty}^{\infty} N(x;\mu,\sigma^2)dx = 1$

__Solución__

Sea $\mathrm{I} = \int_{-\infty}^{\infty} \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}} dx$

Si probamos que $\mathrm{I}^2 = 1$, se tendría que $\mathrm{I} = 1$ y habremos terminado.

Haciendo el cambio de variable $u = \frac{x-\mu}{\sigma}$ $\therefore$ $du = \frac{dx}{\sigma}$, sustituyendo en I nos queda:

$$\mathrm{I} = \frac{\sigma}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{u^2}{2}} du \quad \text{entonces} \quad \mathrm{I}^2 = \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{u^2}{2}} du\right)\left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{v^2}{2}} dv\right)$$

$$\mathrm{I}^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{1}{2}(u^2+v^2)} dudv$$

Trabajando con coordenadas polares,

Si $u = rCos\theta$ y $v = rSen\theta$

$u^2 + v^2 = r^2Cos^2\theta + r^2Sen^2\theta = r^2\left(Cos^2\theta + Sen^2\theta\right) = r^2$

$$\begin{vmatrix} \partial_r u & \partial_\theta u \\ \partial_r v & \partial_\theta v \end{vmatrix} = \begin{vmatrix} Cos\theta dr & -rSen\theta d\theta \\ Sen\theta dr & rCos\theta d\theta \end{vmatrix} = rCos^2\theta drd\theta + rSen^2\theta drd\theta = rdrd\theta\left(Cos^2\theta + Sen^2\theta\right) = rdrd\theta$$

$$\mathrm{I}^2 = \frac{1}{2\pi} \int_0^{2\pi}\int_0^{\infty} e^{-\frac{r^2}{2}} rdrd\theta$$

Tomando $u = -\frac{r^2}{2}$ y $du = -rdr$

$$I^2 = \frac{1}{2\pi}\int_0^{2\pi}\left[-e^{-\frac{r^2}{2}}\Bigg|_0^{\infty}\right]d\theta = \frac{1}{2\pi}\int_0^{2\pi} d\theta = \frac{1}{2\pi}\theta\Big|_0^{2\pi} = 1$$

Observa que $\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ es una normal con media 0 y varianza 1. Esta normal recibe el nombre de normal estándar y se denota como $Z \sim N(0,1)$.

3.4.1 ESTANDARIZACIÓN Y CÁLCULO DE PROBABILIDADES UTILIZANDO LAS TABLAS.

La v.a Normal estándar se representa con la letra zeta y se escribe $Z \sim N(z;0,1)$ o simplemente $Z \sim N(0,1)$, los valores de la distribución acumulada de la normal estándar se encuentran tabulados en unas tablas conocidas como tablas normales.

Función generadora de momentos.

$$\psi_X(x) = E\left(e^{tX}\right) = \int_{-\infty}^{\infty} e^{tx}\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}dx$$

Tomando $u = \frac{x-\mu}{\sigma}$ y $du = \frac{dx}{\sigma}$

$$\psi_X(t) = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{t(\mu+\sigma u)}e^{-\frac{1}{2}u^2}\sigma du = \frac{e^{t\mu}}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(u^2-2t\sigma u\right)}du$$

Completando el trinomio cuadrado perfecto en la exponencial.

$$\psi_X(t) = \frac{e^{t\mu}}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(u^2-2t\sigma u+t^2\sigma^2-t^2\sigma^2\right)}du = \frac{e^{t\mu}e^{\frac{1}{2}t^2\sigma^2}}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{1}{2}(u-t\sigma)^2}du$$

$z = u - t\sigma$ y $dz = du$

$$\psi_X(t) = \exp\left(t\mu + \tfrac{1}{2}t^2\sigma^2\right)\left[\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp\left(-\tfrac{1}{2}z^2\right)dz\right]$$

Observa que el término entre corchetes es una normal estándar por lo que la integral tiene el valor de 1 por ser una f.d.p. así que la f.g.m de la normal es: $\boxed{\psi_X(t) = \exp\left(t\mu + \tfrac{1}{2}t^2\sigma^2\right)}$

Media y varianza.

$$E(X)=\psi'_X(t=0)=e^{t\mu+\frac{1}{2}t^2\sigma^2}\left[\mu+\frac{\sigma^2 2t}{2}\right]\Bigg|_{t=0}=e^0(\mu+0)=\mu$$

$$E\left(X^2\right)=\psi''_X(t=0)=e^{t\mu+\frac{1}{2}t^2\sigma^2}\sigma^2+\left[\mu+\sigma^2 t\right]\,e^{t\mu+\frac{1}{2}t^2\sigma^2}\left(\mu+\sigma^2 t\right)\Bigg|_{t=0}=e^0\sigma^2+\mu e^0\mu=\sigma^2+\mu^2$$

$$V(X)=E\left(X^2\right)-E^2(X)=\sigma^2+\mu^2-\mu^2=\sigma^2$$

Función de distribución acumulada.

$$F_X(x)=\int_{-\infty}^{x}\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}dx$$

Esta integral se tiene que hacer con métodos de integración numérica, pero se tabularon los valores para la normal estándar es decir para la normal con media cero y varianza 1. La distribución acumulada en z se denota por

$\phi(z)=F_Z(z)=\int_{-\infty}^{z}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{z^2}{2}\right)dz$. Las tablas se conocen como tablas normales. Para consultarlas, primero se estandariza la distribución normal con media μ y σ^2, $N\left(x;\mu,\sigma^2\right)$ mediante la fórmula, $\boxed{z=\frac{x-\mu}{\sigma}}$.

Los valores originales de la variable se recuperan con la fórmula $x=z\sigma+\mu$.

Ejemplo 1: La resistencia al rompimiento en Newton de una tela sintética denotada por X se distribuye normal con media de 800 y varianza 144. El comprador de la tela requiere que esta tenga una resistencia de por lo menos 772N. Se selecciona al azar y se prueba una muestra de tela ¿Cuál es la probabilidad de que el comprador se lleve la tela?

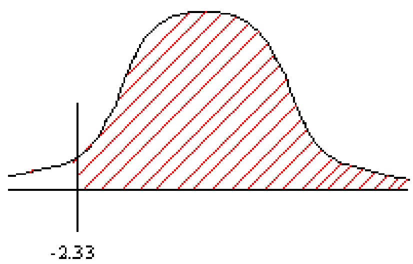
<u>Solución</u>

X : Resistencia de la tela al rompimiento.
$X\sim N\left(x;\mu=800,\sigma^2=144\right)$

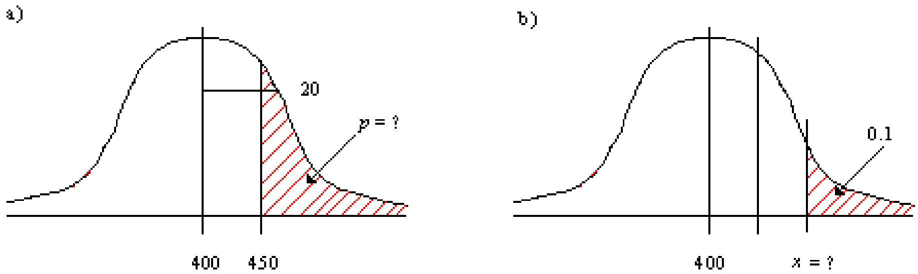
Primero se estandarizan ambos lados de la desigualdad para poder buscar en tablas.

$$P(X\geq 772)=P\left(\frac{X-\mu}{\sigma}\geq\frac{772-\mu}{\sigma}\right)=P\left(Z\geq\frac{772-800}{12}\right)=P(Z\geq -2.33)$$

$$=1-\phi(-2.33)=1-0.0099=0.9901$$

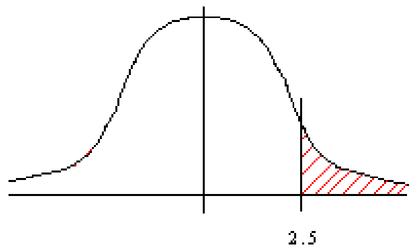


Ejemplo 2: Se observó durante un largo período que la cantidad semanal gastada en el mantenimiento y en las reparaciones en cierta fábrica tiene aproximadamente una distribución normal con una media de \$400 y una desviación estándar de \$20. Si el presupuesto para la próxima semana es de \$450.
a) ¿Cuál es la probabilidad de que los costos reales sean mayores que la cantidad presupuestada?
b) ¿De cuánto tendría que ser el presupuesto para reparaciones semanales y mantenimiento para que la cantidad presupuestada solamente sea rebasada con una probabilidad de 0.1?



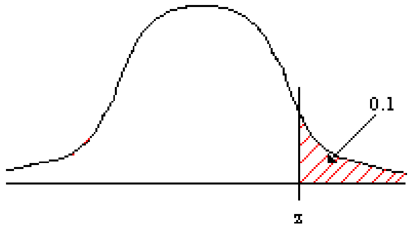
<u>Solución</u>

X : Cantidad semanal gastada.
$X \sim N\left(x; \mu = 400, \sigma^2 = 400\right)$



a) $P(X > 450) = P\left(Z \geq \dfrac{450 - 400}{20}\right) = P(Z \geq 2.5) = 1 - \phi(2.5) = 1 - 0.9938 = 0.0062$

b) Estoy buscando un valor x para la v.a tal que $P(X > x) = 0.1$. Al estandarizar la v.a hay un valor z para la normal estándar que corresponde a este valor de x. Es decir; $P(Z \geq z) = 0.1$.



Buscamos en las tablas el valor de z que corresponde a la probabilidad 0.1 y despejamos el valor de la v.a de interés de la fórmula de estandarización,

$$z = \frac{x-\mu}{\sigma} \quad \therefore \quad x = \mu + z\sigma$$

El valor de tablas es $z = 1.28$ por lo tanto $x = 400 + 1.28(20) = 425.6$.

Teorema: Si X tiene una distribución normal con media μ y varianza σ^2 y si $Y = aX + b$ donde a y b son constantes con a diferente de cero, entonces Y tiene una distribución normal con media $a\mu + b$ y varianza $a^2\sigma^2$.

Demostración

$$X \sim N(x;\mu,\sigma^2)$$
$$Y = aX + b$$
$$\psi_Y(t) = E\left(e^{tY}\right) = E\left(e^{t(aX+b)}\right) = e^{tb}E\left(e^{atX}\right) = e^{tb}\psi_X(at) = e^{tb}e^{at\mu+\frac{1}{2}a^2t^2\sigma^2} = \exp\left[(b+a\mu)t + \tfrac{1}{2}t^2\left(a^2\sigma^2\right)\right]$$

La f.g.m para la v.a Y corresponde a una normal con $\mu_Y = a\mu + b$ y $\sigma_Y^2 = a^2\sigma^2$.

Teorema: Si las variables aleatorias $X_1, X_2, ..., X_k$ son independientes y $X_i \sim N(x_i;\mu_i,\sigma_i^2)$ entonces $X_1 + \cdots + X_k$ tiene una distribución normal con media $\mu = \mu_1 + \cdots + \mu_k$ y varianza $\sigma^2 = \sigma_1^2 + \cdots + \sigma_k^2$.

Demostración

$$\psi_{X_i}(t) = e^{t\mu_i + \frac{1}{2}t^2\sigma_i^2}$$
$$X = \sum_{i=1}^{k} X_i$$
$$\psi_X(t) = \prod_{i=1}^{k}\psi_{X_i}(t) = \prod_{i=1}^{k} e^{t\mu_i+\frac{1}{2}t^2\sigma_i^2} = \exp\left(t\sum_{i=1}^{k}\mu_i + \tfrac{1}{2}t^2\sum_{i=1}^{k}\sigma_i^2\right)$$

Que es precisamente la f.g.m para una normal con media $\mu_X = \sum_{i=1}^{k}\mu_i$ y varianza $\sigma_X^2 = \sum_{i=1}^{k}\sigma_i^2$.

Corolario 1: Si las variables aleatorias $X_1, X_2, ..., X_k$ son independientes con $X_i \sim N(x_i; \mu_i, \sigma_i^2)$ y si $a_1, a_2, ..., a_k$ y b son constantes para las que al menos uno de los valores $a_1, a_2, ..., a_k$ es distinto de cero, entonces la variable $X = a_1 X_1 + \cdots + a_k X_k + b$ tiene una distribución normal con media $\mu = a_1\mu_1 + \cdots + a_k\mu_k + b$ y varianza $\sigma^2 = a_1^2\sigma_1^2 + \cdots + a_k^2\sigma_k^2$.

Corolario 2: Supón que las variables aleatorias $X_1, X_2, ..., X_k$ constituyen una muestra aleatoria de una distribución normal con media μ y varianza σ^2. Sea la variable aleatoria $\overline{X}$ conocida como la media muestral. Entonces $\overline{X} \sim N\left(\overline{x}; \mu, \frac{\sigma^2}{n}\right)$.

<u>**Demostración**</u>

$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ Por el corolario anterior se tiene que $\overline{X} \sim N\left(\overline{x}; \mu_{\overline{X}} = \frac{1}{n} n\mu = \mu, \sigma_{\overline{X}}^2 = \sum_{i=1}^{n} \frac{1}{n^2}\sigma^2 = \frac{\sigma^2}{n}\right)$

Ejemplo 1: Un eje con diámetro exterior distribuido normalmente con media 1.2 y varianza 0.0016 se inserta en un cojinete de manguito que tiene un diámetro interior distribuido normalmente con media 1.25 y varianza 0.0009 ¿Cuál es la probabilidad de que el eje no quepa en el cojinete?

<u>**Solución**</u>

X : Diámetro exterior del eje.
Y : Diámetro interior del cojinete.

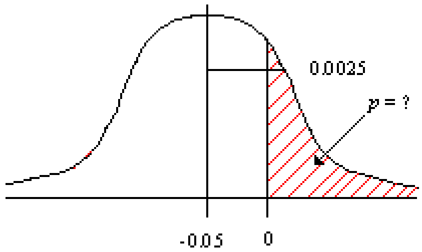
$X \sim N(x; \mu_X = 1.2, \sigma_X^2 = 0.0016)$
$Y \sim N(y; \mu_Y = 1.25, \sigma_Y^2 = 0.0009)$

Si el eje es más grande que el cojinete, entonces no cabe, por lo tanto la probabilidad de que el eje no quepa en el cojinete coincide con $P(X > Y)$, aplicando la teoría sobre suma de normales y estandarizando se tiene que:

$P(X > Y) = P(X - Y > 0) = P(W > 0)$

Con $W = X - Y$, entonces $W \sim N(w; \mu_W = 1.2 - 1.25 = -0.05, \sigma_W^2 = 0.0016 + 0.0009 = 0.0025)$.

Por lo tanto $P(W > 0) = P\left(Z > \frac{0 - (-0.05)}{\sqrt{0.0025}}\right) = P(Z > 1) = 1 - 0.84134 = 0.15866$



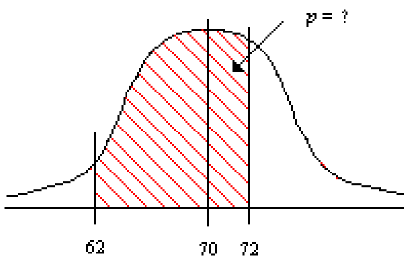
Ejemplo 2: La dureza de Rockwell de una aleación particular se distribuye normalmente con media de 70 y desviación estándar de 4.
a) Si un espécimen se acepta sólo si su dureza está entre 62 y 72. ¿Cuál es la probabilidad de que un espécimen elegido al azar tenga una dureza aceptable?
b) Si el intervalo de dureza aceptable es $(70-c, 70+c)$ ¿Para qué valor de c el 95% de los especimenes tendrían una dureza aceptable?
c) En el caso de que el intervalo aceptable sea el indicado en a) y la dureza de cada uno de 9 especimenes seleccionados al azar se determine en forma independiente ¿Cuál es el número esperado de especimenes aceptables de entre los 9?

<u>Solución</u>

X : Dureza de la aleación.
$X \sim N(x; \mu = 70, \sigma^2 = 16)$

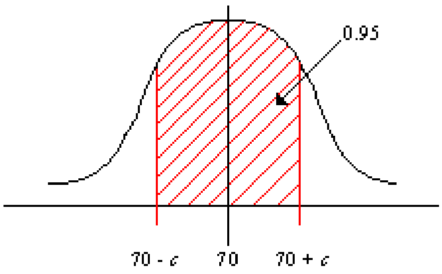
a) $P(62 < X < 72) = P\left(\frac{62-70}{4} < \frac{X-\mu}{\sigma} < \frac{72-70}{4}\right) = P(-2 < Z < 0.5)$

$$= \phi(0.5) - \phi(-2) = 0.69146 - 0.02275 = 0.66871$$



b) $P(70-c < X < 70+c) = P\left(\frac{70-c-70}{4} < \frac{X-\mu}{\sigma} < \frac{70+c-70}{4}\right) = P\left(\frac{-c}{4} < Z < \frac{c}{4}\right) = 0.95$

Debido a la simetría de la distribución y del intervalo se busca en tablas el valor de z que acumula una probabilidad de 0.975 el cual es $z = 1.96$ entonces, $1.96 = \frac{c}{4}$, por lo tanto $c = 4(1.96) = 7.84$



c) Sea Y : el número de especimenes aceptables.

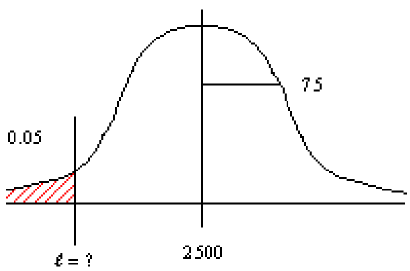
$\therefore\ Y \sim \text{bin}(y;\ n=9,\ p=0.66871)$

$E(Y) = np = 9(0.66871) = 6$

Ejemplo 3: Se sabe que cierta bombilla eléctrica tiene una salida que se distribuye normalmente con media de 2500 pie-candela y desviación estándar de 75 pie-candela. Determine un límite de especificación inferior tal que sólo 5% de las bombillas fabricadas sean defectuosas.

Solución

X : Salida de la bombilla.
$X \sim N\left(x; \mu = 2500, \sigma^2 = 75^2\right)$



$$P(X < l) = P\left(\frac{X-\mu}{\sigma} < \frac{l-2500}{75}\right) = P(Z < z) = 0.05$$

Buscando en tablas el valor de z que corresponde a la probabilidad acumulada de 0.05, se tiene que $z = \dfrac{l-2500}{75} = -1.65$ por lo tanto el límite inferior de especificación debe ser:
$l = 75z + 2500 = 75(-1.65) + 2500 = 2{,}376.25$ pie-candela.

3.4.2 Teorema Central del límite.

Si las variables aleatorias $X_1, X_2, ..., X_n$ constituyen una muestra aleatoria de tamaño *n* de una distribución con media μ y varianza σ^2 entonces para cualquier número fijo *x*.

$$\lim_{n\to\infty} P\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \le x\right) = \phi(x) \quad \text{con} \quad \phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

Interpretación: Si se selecciona una muestra aleatoria suficientemente grande de cualquier distribución con media μ y varianza σ^2 independientemente de si ésta distribución es discreta o continua, entonces la distribución de la variable aleatoria $\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}$ es aproximadamente normal estándar. Es decir $\overline{X}$ tiene una distribución aproximadamente normal con media μ y varianza $\left(\dfrac{\sigma}{\sqrt{n}}\right)^2$. Equivalentemente $\sum_{i=1}^{n} X_i$ se distribuye aproximadamente normal con media $n\mu$ y varianza $n\sigma^2$.

Ya que $E(\overline{X}) = \mu = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n} E\left(\sum_{i=1}^{n} X_i\right) \;\therefore\; E\left(\sum_{i=1}^{n} X_i\right) = n\mu$

$$V(\overline{X}) = \frac{\sigma^2}{n} = V\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2} V\left(\sum_{i=1}^{n} X_i\right) \;\therefore\; V\left(\sum_{i=1}^{n} X_i\right) = \frac{n^2\sigma^2}{n} = n\sigma^2$$

Ejemplo 1: Supón que se lanza una moneda 900 veces. ¿Cuál es la probabilidad de obtener más de 495 caras?

Solución

Sean las variables aleatorias

$$X_i = \begin{cases} 1 & \text{si cae cara en el } i\text{-ésimo tiro.} \\ 0 & \text{si cae cruz en el } i\text{-ésimo tiro.} \end{cases} \quad \text{con } i = 1, ..., 900.$$

Con $E(X_i) = p_i = \frac{1}{2}$ y $V(X_i) = p_i q_i = \frac{1}{4}$.

Sea $X = \sum_{i=1}^{900} X_i$ el número total de caras en los $n = 900$ lanzamientos. Entonces, por el Teorema Central del límite (TCL) se tiene que $X \sim N(x; \mu_X = n\mu, \sigma_X^2 = n\sigma^2)$, por lo tanto:

$$P(X > 495) = P\left(\frac{X - n\mu}{\sqrt{n\sigma^2}} > \frac{495 - 900(0.5)}{\sqrt{900(0.25)}}\right) = P(Z > 3) = 0.0013$$

Ejemplo 2: Supón que la distribución del número de defectos en un determinado rollo de tela es una distribución de Poisson con media 5 y que para una muestra aleatoria de 125 rollos se cuenta el número de defectos en cada rollo. Determina la probabilidad de que el número promedio de defectos por rollo en la muestra sea menor que 5.5.

Solución

X_i : Número de defectos en el i-ésimo rollo.

$\mu_i = 5$ y $\sigma_i^2 = 5$

$\overline{X}$: Número promedio de defectos por rollo en la muestra.

Por el TCL $\overline{X} \sim N\left(\overline{x}; \mu_{\overline{x}} = \mu = 5,\ \sigma_{\overline{x}}^2 = \frac{\sigma^2}{n} = \frac{5}{125} = 0.04\right)$.

Entonces: $P(\overline{X} < 5.5) = P\left(\frac{\overline{X} - \mu_{\overline{X}}}{\sigma_{\overline{X}}} < \frac{5.5 - 5}{\sqrt{0.04}}\right) = P(Z < 2.5) = 0.9938$

Ejemplo 3: Se empacan 250 piezas pequeñas en una caja. Los pesos de las piezas son variables aleatorias independientes con media de 0.5 libras y desviación estándar de 0.1 libras. Se cargan 20 cajas en una tarima. ¿Cuál es la probabilidad de que las piezas en la tarima excedan 2,510 libras de peso despreciando tanto el peso de la tarima como el de la caja?

Solución

X_i : Peso de la $i-$ésima pieza.

$250(20) = 5000$ piezas, entonces $n = 5000$.

$X_i \sim f_{X_i}\left(x_i; \mu_{X_i} = 0.5, \sigma_{X_i}^2 = (0.1)^2\right)$

Sea $X = \sum_{i=1}^{n} X_i$ el peso total de las piezas. Por el TCL se tiene que la distribución de esta v.a está dada por $X \sim N\left(\mu_X = 5000(0.5), \sigma^2 = 5000(0.1)^2\right)$.

$$P(X > 2510) = P\left(\frac{X - \mu_X}{\sigma_X} > \frac{2510 - 2500}{7.07}\right) = P(Z > 1.41) = 1 - \phi(1.41) = 1 - 0.9207 = 0.079$$

3.4.3 APROXIMACIÓN NORMAL A LA BINOMIAL.

Si repetimos n ensayos Bernoulli de manera independiente, se tiene que $E(X_i) = p$ y $V(X_i) = pq$ con

$$X_i = \begin{cases} 1 & \text{si éxito en el } i\text{-ésimo ensayo.} \\ 0 & \text{si fracaso en el } i\text{-ésimo ensayo.} \end{cases}$$

el TCL garantiza que para n grande $\bar{X} \sim N\left(\bar{x}; \mu_{\bar{x}} = \mu = p, \sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} = \frac{pq}{n}\right)$

y que $X = \sum_{i=1}^{n} X_i \sim N\left(x; \mu_X = n\mu = np, \sigma_X = n\sigma^2 = npq\right)$

Si n es grande, y no podemos usar tablas binomiales para encontrar probabilidades de este tipo, se aproxima con una normal.

La distribución binomial es discreta y la normal es continua, ¿Qué se hace si se tiene interés en una probabilidad puntual de tipo $P(X = x)$? Para resolver el problema se usa la llamada *corrección por continuidad* que consiste en dar un margen de media unidad alrededor del valor puntual, por ejemplo $P(X = x) \approx P(x - 0.5 \le X \le x + 0.5)$.

$P(X \ge x) \approx P(X \ge x - 0.5)$

$P(X \le x) \approx P(X \le x + 0.5)$

$P(a \le X \le b) \approx P(a - 0.5 < X < b + 0.5)$

Ejemplo 1: En el muestreo de un proceso que produce 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.
$p = 0.2 \quad q = 0.8 \quad n = 100$

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

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

Con corrección $P(X \le 15) = P(X \le 15.5) = P\left(Z \le \frac{15.5-20}{4}\right) = P(Z \le -1.125) = 0.1303$

Si comparamos estas probabilidades con la probabilidad correspondiente a la distribución exacta, es decir binomial, podemos observar que se obtiene una mejor aproximación al utilizar la corrección por continuidad.

Ver el cuadro siguiente:

Comparación de la $P(X \leq 15)$ según la distribución utilizada.

Binomial	Normal	Normal con corrección
0.12850551	0.10564977	0.13029452
Error	0.02285574	0.00178900

Como era de esperarse, es mejor la aproximación cuando se usa la corrección por continuidad observa que el error es más pequeño en este caso.

Ejemplo 2: 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 :El $i-$ésimo votante está a favor de la emisión.

Se trata de una población Bernoulli con media $\mu = p = 0.2$ y varianza $\sigma^2 = pq = 0.2(0.8) = 0.16$.

$\overline{X}$: fracción de votantes en la muestra a favor de la emisión.

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

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

$$= P\left(\frac{-0.06}{0.05} \leq \frac{\overline{X} - 0.2}{\sqrt{0.0025}} \leq \frac{0.06}{0.05}\right)$$

$$= P(-1.2 \leq Z \leq 1.2) = 0.7698$$

Ejemplo 3: 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 : El $i-$ésimo pasajero llega.
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 $\sigma^2 = pq = 0.05(0.95) = 0.0475$.

X : Número de personas que llegan.

$$X = \sum_{i=1}^{160} X_i$$

$p = 0.95 \quad q = 0.05 \quad n = 160$

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

Sin corrección: $P(X \leq 155) = P\left(\frac{X-\mu_X}{\sigma_X} \leq \frac{155-152}{\sqrt{7.6}}\right) = P(Z \leq 1.088) = 0.8617$

Con corrección: $P(X \leq 155.5) = P\left(\frac{X-\mu_X}{\sigma_X} \leq \frac{155.5-152}{\sqrt{7.6}}\right) = P(Z \leq 1.27) = 0.8979$

Binomial: $P(X \leq 155) = 0.9061$

Ejemplo 4: 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 pozo A.
$p_B = 0.2$ proporción de peces en el pozo B.
$\overline{X}_A$: proporción muestral de peces en el pozo A.
$\overline{X}_B$: proporción muestral de peces en el pozo B.
X : diferencia entre las proporciones muestrales.

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

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

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

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

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

$$= P(-1.4142 \le Z \le 1.4142) = 0.8414$$

3.4.4 Teorema de Chebyshev.

Es de gran ayuda cuando no conocemos la f.d.p exacta de la v.a, pero conocemos los valores para su media y su varianza. Por ser desconocida la f.d.p no podemos encontrar probabilidades exactas de cualquier evento, pero este teorema pone una cota para la probabilidad de que la v.a se encuentre dentro de k veces su desviación estándar respecto a la media.

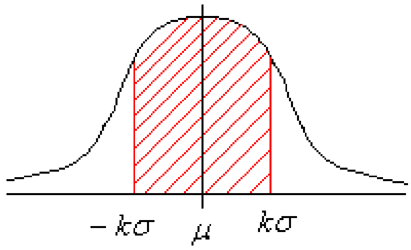
Teorema._ Sea X una v.a (discreta o continua) con f.d.p desconocida con media μ y varianza σ^2, para $k > 0$ se cumple la siguiente desigualdad:

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

restando μ en todas partes dentro del paréntesis se tiene que $P(-k\sigma < X - \mu < k\sigma) \ge 1 - \frac{1}{k^2}$

es decir $P(|X-\mu| < k\sigma) \ge 1 - \frac{1}{k^2}$ o equivalentemente $P(|X-\mu| \ge k\sigma) \le \frac{1}{k^2}$.

Observación: No conocemos la f.d.p de la v.a y por lo tanto el valor de la probabilidad exacta de que el valor de la v.a está entre $\pm k$ desviaciones estándar de la media es desconocido, pero el Teorema de Chebyshev nos dice que no puede ser más chico que $1 - \frac{1}{k^2}$.



Ejemplo 1: 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.
$X \sim f_X\left(x; \mu = 14, \quad \sigma^2 = 2\right)$

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

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

$$\mu \pm k\sigma = 14 \pm 2\sqrt{2}$$

El intervalo es [11, 17] días.

Ejemplo 2: El servicio postal requiere, en promedio, 2 días para entregar una carta en una Ciudad. La varianza se estima como 0.4 $(\text{días})^2$. 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 una carta.
$\mu = 2$ días.
$\sigma^2 = 0.4$ días^2

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

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

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

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

3.5.1 DISTRIBUCION MUESTRAL JI-CUADRADA.

Cuando se toma una muestra aleatoria de tamaño n de una población de tamaño N con f.d.p $f_X(x)$, cada una de las variables aleatorias $X_1, X_2, \ldots, X_n$ que conforman la muestra son independientes y con la misma distribución. Cualquier función de estas variables sigue siendo una variable aleatoria que debe tener una f.d.p propia, que se puede deducir de la distribución conjunta de las variables aleatorias $X_1, X_2, \ldots, X_n$, que conforman la muestra aleatoria, por esta razón se denomina a esta f.d.p como distribución muestral.
Estas distribuciones muestrales son indispensables en las técnicas estadísticas más usadas en la práctica diaria tales como los intervalos de confianza, prueba de hipótesis, prueba de independencia entre variables, análisis de varianza, entre otras.

Definición: Un **estadístico** es cualquier función de una muestra aleatoria $X_1, X_2, \ldots, X_n$.

La media muestral $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} X_i$ y la varianza muestral $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2$ son ejemplos de estadísticos.

Definición: Los **grados de libertad** $g.l$ para cualquier estadístico es el número de datos que pueden variar libremente al calcular ese estadístico.

Por ejemplo, la media muestral no tiene restricción alguna por esta razón los n datos pueden variar libremente, es decir $\overline{x}$ tiene $n\ g.l.$, pero la suma de las desviaciones de los datos a su media debe ser cero, es decir $\sum_{i=1}^{n}\left(x_i - \overline{x}\right) = 0$, por esta razón uno de los datos esta sujeto al valor de la suma de los $(n-1)$ restantes y con esto la varianza muestral tiene $(n-1)\ g.l$.

DISTRIBUCIÓN MUESTRAL JI-CUADRADA.

Si X es una v.a con distribución gamma con parámetros $\beta = \frac{1}{2}$ y $\alpha = \frac{n}{2}$ con n entero positivo, entonces se dice que X tiene una distribución llamada ji-cuadrada con n grados de libertad y se denota como $X \sim \chi_n^2$.

Su f.d.p es de la forma $f_X(x) = \frac{x^{\frac{n}{2}-1} e^{-\frac{x}{2}}}{2^{\frac{n}{2}}\Gamma(n/2)}$, $x > 0$. La media y la varianza de la distribución son $E(X) = n$ y $V(X) = 2n$ respectivamente.

En esta distribución n se conoce como el parámetro de forma. Los valores de la distribución de probabilidad acumulada se encuentran tabulados en tablas.

Observa que si la ji-cuadrada tiene dos grados de libertad, se convierte en una exponencial con parámetro $\frac{1}{2}$.

Por lo tanto la distribución gamma con parámetros $\alpha = 1$ y $\beta = \frac{1}{2}$, la distribución χ_2^2 y la distribución exponencial con parámetro $\beta = \frac{1}{2}$ son exactamente la misma distribución.

PARA LEER LAS TABLAS.

La notación $\chi^2_{\alpha,g.l}$ representa el valor de la v.a ji-cuadrada con $g.l$ grados de libertad que deja a la derecha una cola (probabilidad) de α.
Por ejemplo $\chi^2_{0.025,10} = 20.483$ indica que la probabilidad de que la v.a ji-cuadrada con 10 $g.l$ sea mayor que 20.483 es de 0.025, es decir $P\left(\chi^2_{10g.l} > 20.483\right) = 0.025$.

Generalmente las tablas de la distribución ji-cuadrada tienen como encabezado de las columnas los valores de α (probabilidad en la cola derecha de la distribución) y como encabezado de las filas los grados de libertad, lo que tenemos que hacer es buscar el número que se localiza en la intersección de la columna correspondiente a la probabilidad de interés y la fila correspondiente a los $g.l$ de interés y ese es el valor de la ji-cuadrada que estamos buscando.
En el ejemplo anterior buscamos la intersección de la columna encabezada por 0.025 con la fila encabezada por 10 y ahí se localiza el número 20.483.

Tabla de distribución ji-cuadrada.

Probabilidad de cola derecha

$g.l$	**0.025**
10	**20.483**

Ejercicio: Verificar en las tablas los siguientes valores:

a) $\chi^2_{0.975,10} = 3.247$
b) $\chi^2_{0.01,25} = 44.314$
c) $\chi^2_{0.975,35} = 20.569$
d) $\chi^2_{0.995,29} = 13.1211$
e) $\chi^2_{0.95,60} = 43.1879$
f) $\chi^2_{0.01,17} = 33.4087$
g) $\chi^2_{0.025,50} = 71.4202$

Con Excel:

También podemos encontrar estos valores en Excel, con la instrucción "= *prueba.chi.inv(*α*,g.l)*" donde α es la probabilidad que queda en la cola derecha de la distribución ji-cuadrada y $g.l$ son los grados de libertad.
Por ejemplo; la instrucción $=$ *prueba.chi.inv(*$0.975,10$*)* reporta como resultado el valor 3.247 lo cual significa que $P(\chi_{10}^2 > 3.247) = 0.975$.

3.5.2 DISTRIBUCION MUESTRAL t de Student.

Estrechamente relacionada con muestras aleatorias de una distribución normal. Si Y y Z son dos variables aleatorias independientes tales que $Y \sim N(y; \mu = 0, \sigma^2 = 1)$ y $Z \sim \chi_n^2$, entonces la v.a $t = \frac{Y}{\sqrt{\frac{Z}{n}}}$ tiene una distribución que se llama t de Student (en honor a su autor quien firmó sus investigaciones de esta distribución con el seudónimo de Student) con n grados de libertad, con f.d.p dada por la expresión.

$$t \sim t_n(t;n) = \frac{\Gamma[(n+1)/2]}{\Gamma(n/2)\sqrt{\pi n}}\left(1 + \frac{t^2}{n}\right)^{-(n+1)/2} \qquad -\infty < t < \infty$$

Esta distribución también es simétrica y cuando n tiende a infinito la distribución t tiende a la normal estándar, la diferencia entre estas distribuciones es que la t de Student tiene colas más pesadas, es decir se acumula más probabilidad en las colas.

Existen tablas de la distrubición t de Student y para leerlas se necesita la probabilidad de la cola derecha y los grados de libertad.
La expresión $t_{\alpha,g.l}$ representa al valor de la distribución t de Student que deja a la derecha una probabilidad igual a α y que tiene $g.l$ grados de libertad.

Uno de los problemas más frecuentes en estadística es la comparación entre medias de dos poblaciones y en el caso de que no se cuente con toda la información poblacional o bien el tamaño de la muestra sea pequeño se usa esta distribución y por esta razón la prueba de comparación de medias para muestras pequeñas se conoce también como "prueba t".

PARA LEER LAS TABLAS.

La tabla de distribución t de Student también tiene columnas y filas, en las columnas se localiza la probabilidad de cola derecha y en las filas se presentan los grados de libertad.
Si queremos buscar el valor de la distribución t de Student con 9 $g.l$ que deja a la derecha una cola de 0.025, buscamos en la tabla el número localizado en la intersección de la columna encabezada por 0.025 con la fila encabezada por 9 encontrando así el número 2.262.

Tabla de distribución *t*-Student.

Probabilidad de cola derecha

g.l	**0.025**
9	**2.2622**

Esto significa que $P\left(t^2_{0.025,9} > 2.2622\right) = 0.025$.

Ejercicio: Verificar en las tablas los siguientes valores:

a) $t^2_{0.025,25} = 2.06$

b) $t^2_{0.05,24} = 1.711$

c) $t^2_{0.005,10} = 3.169$

d) $t^2_{0.1,17} = 1.33$

Con Excel:
La función que usa Excel para la distribución t de Student está en base a dos colas, es decir reparte la distribución de interés entre dos, indicando así que se debe repartir en ambas colas (inferior y superior). Por lo tanto la instrucción que usaremos es "= *distr.t.inv(*2α*,g.l)*" donde α es la probabilidad en la cola derecha y $g.l$ son los grados de libertad de la distribución.
La instrucción "= *distr.t.inv(*$2 * 0.025, 25$*)*" devuelve el valor 2.06 lo cual significa que $P\left(t_{25} > 2.06\right) = 0.0.025$.

Para verificar los incisos anteriores en Excel podemos ayudarnos con el cuadro siguiente:

α	**2*α**	***g.l***	***t*-Student**	**Instrucción**	**Resultado**
0.025	0.05	25	2.060	= distr.t.inv(0.05, 25)	2.060
0.05	0.1	24	1.711	= distr.t.inv(0.1, 24)	1.711
0.005	0.01	10	3.169	= distr.t.inv(0.01, 10)	3.169
0.1	0.2	17	1.333	= distr.t.inv(0.2, 17)	1.333

3.5.3 DISTRIBUCION MUESTRAL F.

Se llama así en honor a R.A Fisher quién la estudió ampliamente. Tiene gran aplicación al probar hipótesis sobre dos o más distribuciones normales.
Supón que tienes una población de tamaño N y que consideras todas la muestras aleatorias posibles de tamaños n_1 y n_2 y para cada una de estas muestras encuentras las varianzas muestrales s_1^2 y s_2^2 luego defines la v.a $F = \frac{s_1^2}{s_2^2}$ y encuentras la f.d.p para esta v.a. Con este procedimiento se obtiene la distribución muestral llamada *F*, que varía con los grados de libertad. Observa que hay dos grados de libertad jugando en esta distribución, los que corresponden al numerador y los que corresponden al denominador.
Esta distribución se usa para hacer inferencias sobre las varianzas poblacionales cuando se tienen dos muestras aleatorias.

La forma de la f.d.p es: $f_X(x;\nu_1,\nu_2) = \frac{\Gamma\left[\frac{1}{2}(\nu_1+\nu_2)\right]\nu_1^{\nu_1/2}\nu_2^{\nu_2/2}}{\Gamma\left(\frac{1}{2}\nu_1\right)\Gamma\left(\frac{1}{2}\nu_2\right)} \frac{x^{\frac{\nu_1}{2}-1}}{(\nu_1 x+\nu_2)^{(\nu_1+\nu_2)/2}}$

Definición: Si *U* y *V* son dos variables aleatorias independientes tales que $U \sim \chi^2_{\nu_1}$ y $V \sim \chi^2_{\nu_2}$ entonces la v.a $F_{\nu_1,\nu_2} = \frac{U/\nu_1}{V/\nu_2}$ tiene una distribución llamada *F* con ν_1 $g.l$ en el numerador y ν_2 $g.l$ en el denominador y se escribe $F \sim F_{\nu_1,\nu_2}(x;\nu_1,\nu_2)$.

PARA LEER LAS TABLAS.

Se necesita la probabilidad de la cola derecha α, los grados de libertad del numerador ν_1 y los grados de libertad del denominador ν_2. La notación F_{α,ν_1,ν_2} representa el valor de la v.a F con ν_1 $g.l$ en el numerados y ν_2 $g.l$ en el denominador que deja a la derecha una cola (probabilidad) de α.
Por ejemplo $F_{0.025,30,14} = 2.73$ indica que la probabilidad de que la v.a F con 30 $g.l$ en el numerador y 14 $g.l$ en el denominador sea mayor que 2.73 es de 0.025, es decir $P(F_{30,14} > 2.73) = 0.025$.
Las tablas para esta distribución tienen columnas encabezadas por los grados de libertad correspondientes al numerador y filas encabezadas por los grados de libertad correspondientes al denominador. El valor de α en algunas tablas se presenta como encabezado de una sección, es decir:

Tabla de distribución F

α = 0.01	$g.l$ del numerador			
$g.l$ del denominador	1	2	…	20
1	4052.18	4.61		6208.73
2	98.50	4.83		99.45
⋮				
20	8.10	8.65		2.94

α = 0.05	$g.l$ del numerador			
$g.l$ del denominador	1	2	…	20
1	161.45	3.05		248.01
2	18.51	3.55		19.45
⋮				
20	4.35	6.94		2.12

Hay otras tablas que contemplan el valor de α dentro de cada valor para los grados de libertad, es decir algo como:

Tabla de distribución F

$g.l$ del denominador			$g.l$ del numerador 1	2	…	20
1	α =	0.01	4052.181	4999.500		6208.730
	α =	0.025	647.789	799.500		993.103
	α =	0.05	161.448	199.500		248.013
2	α =	0.01	98.503	99.000		99.449
	α =	0.025	38.506	39.000		39.448
	α =	0.05	18.513	19.000		19.446
⋮						
20	α =	0.01	8.096	5.849		2.938
	α =	0.025	5.871	4.461		2.464
	α =	0.05	4.351	199.500		2.124

Se debe localizar el apartado de la tabla correspondiente a α y buscar el número que se localiza en la intersección de la columna encabezada por ν_1 con la fila encabezada por ν_2.

Por ejemplo $F_{0.025,2,20} = 4.461$ como se muestra en la figura siguiente:

Tabla de distribución F

$g.l$ del denominador			$g.l$ del numerador 1	**2**	…	20
1	α =	0.01	4052.181	4999.500		6208.730
	α =	0.025	647.789	799.500		993.103
	α =	0.05	161.448	199.500		248.013
2	α =	0.01	98.503	99.000		99.449
	α =	0.025	38.506	39.000		39.448
	α =	0.05	18.513	19.000		19.446
⋮						
	α =	0.01	8.096	5.849		2.938
20	**α =**	**0.025**	5.871	**4.461**		2.464
	α =	0.05	4.351	199.500		2.124

Ejercicio: Verificar en tablas los siguientes valores:

a) $F_{0.005,15,21} = 3.43$

b) $F_{0.1,40,60} = 1.44$

c) $F_{0.05,60,120} = 1.43$

Nota: En las tablas de la distribución F se encuentran valores pequeños de α, si necesitamos encontrar valores de la distribución para probabilidades grandes usamos la igualdad $F_{1-\alpha,m,n} = \dfrac{1}{F_{\alpha,n,m}}$.

Por ejemplo:

a) $F_{0.975,14,30} = \dfrac{1}{F_{0.025,30,14}} = \dfrac{1}{2.73} = 0.37$

b) $F_{0.995,21,15} = \dfrac{1}{F_{0.005,15,21}} = \dfrac{1}{3.43} = 0.29$

c) $F_{0.9,60,40} = \dfrac{1}{F_{0.1,40,60}} = \dfrac{1}{1.44} = 0.69$

d) $F_{0.95,120,60} = \dfrac{1}{F_{0.05,60,120}} = \dfrac{1}{1.43} = 0.70$

Con Excel:

La instrucción que reporta los valores de la distribución F, con ν_1 $g.l$ en el numerador y ν_2 $g.l$ en el denominador, que dejan a la derecha una probabilidad igual a α es "$= distr.f.inv(\alpha, \nu_1, \nu_2)$".

Por ejemplo La instrucción "$= distr.f.inv(0.005, 15, 21)$" devuelve el valor 3.43 lo cual significa que $P(F_{15,21} > 3.43) = 0.005$.

Utilizaremos estas distribuciones en el capítulo 5 que trata de la estimación puntual y por intervalo. La distribución $t-$Student sirve para intervalos de confianza y pruebas de hipótesis de la media poblacional cuando tenemos falta de información poblacional o muestras chicas.

La distribución ji-cuadrada se utiliza en la construcción de intervalos de confianza y pruebas de hipótesis para la varianza de una población.

La distribución F se usa en intervalos de confianza y pruebas de hipótesis cuando se involucran dos varianzas poblacionales.

4 DISTRIBUCIÓN DE VARIABLES ALEATORIAS BIDIMENSIONALES.

OBJETIVO: Extender los conocimientos adquiridos para el caso de una variable aleatoria unidimensional al caso de las variables aleatorias bidimensionales.

4.1 DENSIDAD CONJUNTA Y MARGINAL.

En ciertas ocasiones es necesario estudiar el efecto conjunto de dos variables aleatorias, de ahí la necesidad de estudiar variables aleatorias bidimensionales representadas por $Z = (X, Y)$. Observa que:

a) Si X y Y son discretas, Z será una v. a. bidimensional discreta.

b) Si X y Y son continuas, Z será una v. a. bidimensional continua.

Definición: La función de probabilidad conjunta de una v. a. bidimensional discreta está dada por:

$$f_{X,Y}(x, y) = P(X = x,\ Y = y)$$

y cumple con las siguientes condiciones:

i) $P(X = x, Y = y) \geq 0 \quad \forall (x, y)$

ii) $\sum_{y}\sum_{x} P(X = x, Y = y) = 1$

Definición: La función de densidad de probabilidad conjunta de una v. a. bidimensional continua es una función que satisface las dos condiciones siguientes:

i) $f_{X,Y}(x, y) \geq 0 \quad \forall \quad (x, y) \in R$

ii) $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x, y)dxdy = 1$

Ahora el rango de valores posibles de la v.a bidimensional en el caso continuo serán regiones en el plano y las probabilidades de interés serán volúmenes bajo la curva en la región de interés.

Ejemplo 1: 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 el 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) ¿Cuánto vale 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 aleatorias unidimensionales:

X : Número de artículos producidos por la línea I.
Y : Número de artículos producidos por la línea II.

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

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.

f.d.p de Y

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

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

Ejemplo 2: 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

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.

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

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

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

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

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

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

f.d.p 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

Ejemplo 3: 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

$f_{X,Y}(x,y) = C \quad , \quad 5000 \le x \le 10{,}000 \quad y \quad 4000 \le y \le 9000$

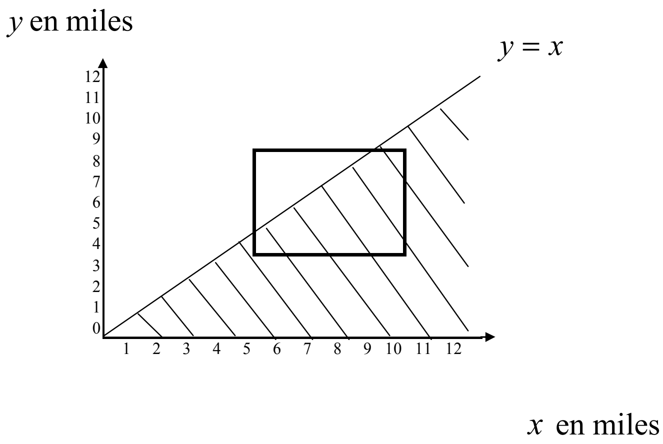
a) Encuentra el valor de la constante *C*.

b) Calcula $P(X \ge Y)$.

__Solución__

En problemas de este tipo lo que se acostumbra hacer es graficar la región en donde existe la f.d.p conjunta y localizar la región de interés para poder identificar los límites de integración en la integral doble que define a la probabilidad pedida.

En la siguiente figura el cuadradito representa la región donde está definida la f.d.p conjunta y la parte rayada que se intersecta con el cuadrado representa la región donde se cumple la condición $X \ge Y$.

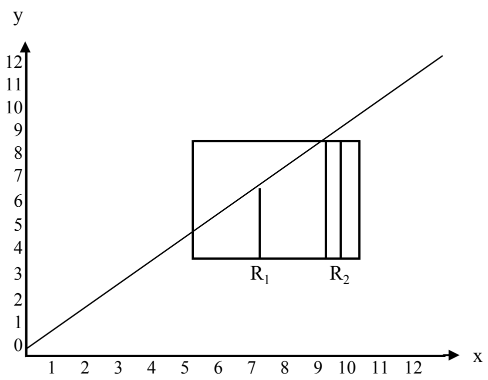


a) Sabemos que

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)dxdy = 1 \qquad \therefore \int_{4mil}^{9mil}\int_{5mil}^{10mil} Cdxdy = 1$$

$$C\left[x\Big|_{5mil}^{10mil} \quad y\Big|_{4mil}^{9mil} \right] = C[(5mil)(5mil)] = 1 \qquad \therefore C = \frac{1}{25{,}000{,}000}$$

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.



$$P(X \geq Y) = C\left[\int_{x=5{,}000}^{9{,}000}\int_{y=4{,}000}^{x} dydx + \int_{y=4{,}000}^{9{,}000}\int_{x=9{,}000}^{10{,}000} dxdy \right]$$

$$= C\left[\int_{x=5{,}000}^{9{,}000} \left(y\Big|_{4{,}000}^{x} \right)dx + \left(x\Big|_{9{,}000}^{10{,}000} \right)\left(y\Big|_{4{,}000}^{9{,}000} \right) \right]$$

$$= C\left[\int_{x=5{,}000}^{9{,}000} (x - 4{,}000)dx + 5{,}000{,}000 \right]$$

$$= C\left[\left(\frac{x^2}{2} - 4{,}000x \right)\Bigg|_{5{,}000}^{9{,}000} + 5{,}000{,}000 \right]$$

$$= \frac{1}{25{,}000{,}000}\left[\frac{81{,}000{,}000 - 25{,}000{,}000}{2} - 16{,}000{,}000 + 5{,}000{,}000 \right]$$

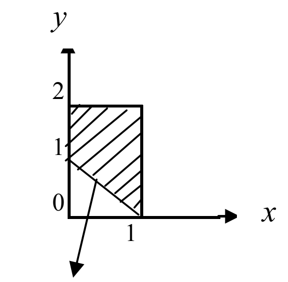
$$= \frac{17}{25} = 0.68$$

Ejemplo 4: 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 $P(X+Y\ge 1)$.

Solución



$y=1-x$

a) $$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)dxdy=\int_0^2\int_0^1 (x^2+\frac{y}{3}x)dxdy=\int_0^2\left[\frac{x^3}{3}+\frac{yx^2}{6}\right]_0^1 dy=\int_0^2\left(\frac{1}{3}+\frac{1}{6}y\right)dy=\left[\frac{1}{3}y+\frac{1}{12}y^2\right]_0^2=\frac{2}{3}+\frac{1}{3}=1$$

b) Si $x+y=1 \ \therefore\ \ y=1-x$

La probabilidad pedida es la integral en la región sombreada.

$$P(X+Y\ge 1)=\int_{x=0}^{x=1}\left[\int_{y=1-x}^{2}\left(x^2+\frac{y}{3}x\right)dy\right]dx=\int_{x=0}^{x=1}\left(x^2y+\frac{x}{3}\frac{y^2}{2}\right)\Bigg|_{y=1-x}^{2} dx$$

$$=\int_{x=0}^{x=1}\left(2x^2-(1-x)x^2+\frac{4}{6}x-\frac{x}{6}(1-x)^2\right)dx$$

$$=\int_{x=0}^{x=1}\left(2x^2-x^2+x^3+\frac{2}{3}x-\frac{x}{6}\left[1-2x+x^2\right]\right)dx$$

$$=\int_{x=0}^{x=1}\left(2x^2-x^2+x^3+\frac{2}{3}x-\frac{x}{6}+\frac{1}{3}x^2-\frac{x^3}{6}\right)dx$$

$$=\int_{x=0}^{x=1}\left(\frac{5}{6}x^3+\frac{4}{3}x^2+\frac{1}{2}x\right)dx=\left[\frac{5}{6}\frac{x^4}{4}+\frac{4}{3}\frac{x^3}{3}+\frac{1}{2}\frac{x^2}{2}\right]_0^1$$

$$=\left[\frac{5}{24}x^4+\frac{4}{9}x^3+\frac{1}{4}x^2\right]_0^1=\frac{45+96+54}{216}=\frac{195}{216}=\frac{65}{72}$$

4.2 DISTRIBUCIÓN CONJUNTA Y MARGINAL.

Definición: La función de distribución acumulada para la v. a bidimensional (X,Y) o bien la función de distribución acumulada conjunta esta dada por:

i) Caso discreto $F_{X,Y}(x,y) = P(X \le x, Y \le y)$

ii) Caso continuo $F_{X,Y}(x,y) = P(X \le x, Y \le y) = \int_{-\infty}^{y}\int_{-\infty}^{x} f_{x,y}(x,y)dxdy$

Nota: $\frac{\partial^2}{\partial x \partial y} F_{X,Y}(x,y) = f_{X,Y}(x,y)$

Ejemplo 1: Y_1 y Y_2 tienen la función de densidad conjunta dada por

$$f_{Y_1,Y_2}(y_1,y_2) = ky_1y_2 \text{ si } 0 \le y_1 \le 1,\ 0 \le y_2 \le 1$$

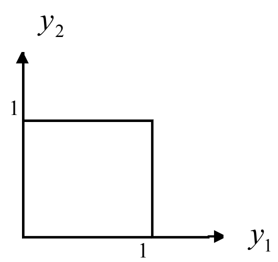
a) Determina el valor de k que la convierte en una función de densidad de probabilidad.

b) Encuentre la función $F_{Y_1,Y_2}(y_1,y_2)$

c) Calcula $P\left(Y_1 \le \frac{3}{4}, Y_2 \ge \frac{1}{2}\right)$

<u>Solución</u>

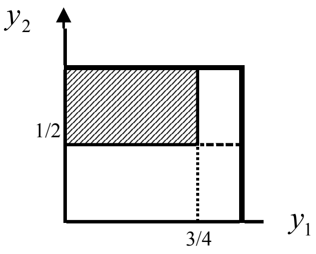
La f.d.p conjunta está definida en el cuadrado unitario siguiente



a) $k\int_0^1\int_0^1 y_1y_2dy_1dy_2 = k\left.\frac{y_1^{\,2}}{2}\right|_0^1\left.\frac{y_2^{\,2}}{2}\right|_0^1 = \frac{k}{4} = 1 \quad \therefore \quad k = 4$

b) $F_{Y_1,Y_2}(y_1,y_2) = 4\int_0^{y_2}\int_0^{y_1} y_1y_2dy_1dy_2 = 4\frac{y_1^{\,2}}{2}\frac{y_2^{\,2}}{2} = y_1^2y_2^2$ dentro del cuadrado $0 \le y_1 \le 1$, $0 \le y_2 \le 1$

c)



$$P\left(Y_1 \le \frac{3}{4}, Y_2 \ge \frac{1}{2}\right) = 4\int_0^{\frac{3}{4}} y_1 dy_1 \int_{\frac{1}{2}}^{1} y_2 dy_2$$

$$= 4\frac{y_1^{\,2}}{2}\Big|_0^{\frac{4}{3}} \frac{y_2^{\,2}}{2}\Big|_{\frac{1}{2}}^{1}$$

$$= \frac{4}{4}\left[y_1^{\,2}\Big|_0^{\frac{3}{4}} y_2^{\,2}\Big|_{\frac{1}{2}}^{1} \right]$$

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

$$= \left[\left(\frac{9}{16}\right)\left(\frac{3}{4}\right)\right] = \frac{27}{64} = 0.422$$

Si estamos interesados en la distribución de probabilidades de uno de los componentes de la v. a. bidimensional (X,Y) se dice que nos interesa la distribución marginal de X o bien la distribución marginal de la v. a. Y.

En el ejemplo de las líneas de producción (caso discreto) teníamos la distribución conjunta en una tabla, si ahora queremos la distribución marginal de X y de Y procedemos a sumar por renglones y por columnas reportando las dos tablas individuales como se muestra a continuación.

Función de probabilidad conjunta $P(X = x, Y = y)$

Y \ X	0	1	2	3	4	5	Suma
0	0	0.01	0.03	0.05	0.07	0.09	0.25
1	0.01	0.02	0.04	0.05	0.06	0.08	0.26
2	0.01	0.03	0.05	0.05	0.05	0.06	0.25
3	0.01	0.02	0.04	0.06	0.06	0.05	0.24
Suma	0.03	0.08	0.16	0.21	0.24	0.28	1.00

Distribución Marginal de X

$X = x$	$P(X = x)$
0	0.03
1	0.08
2	0.16
3	0.21
4	0.24
5	0.28

Distribución Marginal de Y

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

Definición: La función de probabilidad marginal de una variable aleatoria (X o Y) que forma parte de una v.a conjunta o bidimensional (X,Y) está dada por:

Caso Discreto

i) Distribución marginal de la v.a X: $f_X(x) = P(X = x) = \sum_y P(X = x, Y = y) = \sum_y f_{X,Y}(x,y) \quad \forall\ x$

ii) Distribución marginal de la v.a Y: $f_Y(y) = P(Y = y) = \sum_x P(X = x, Y = y) = \sum_x f_{X,Y}(x,y) \quad \forall\ y$

Caso Continuo

i) Distribución marginal de la v.a X: $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy$

ii) Distribución marginal de la v.a Y: $f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dx$

Ejemplo 2: 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.

<u>**Solución**</u>

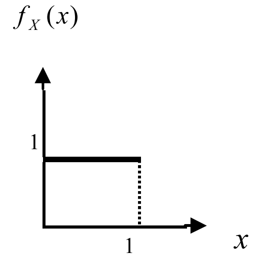
$$f_X(x) = \int_{y=0}^{1} f_{X,Y}(x,y)dy \quad \text{con} \quad 0 \le x \le 1$$

$$= \int_0^1 2(x + y - 2xy)dy = 2\left[xy + \frac{y^2}{2} - \frac{2xy^2}{2} \right]_0^1$$

$$= 2\left[x + \frac{1}{2} - x \right] = 1$$

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 $X \sim U(0,1)$

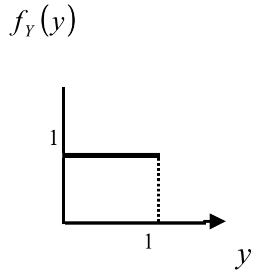


Para encontrar la marginal de Y se integra la conjunta respecto a X.

$$f_Y(y) = 2\int_0^1 (x + y - 2xy)dx \text{ con } 0 \le y \le 1$$

$$= 2\left[\frac{x^2}{2} + xy - \frac{2x^2 y}{2}\right]_0^1 = 2\left[\frac{1}{2} + y - y\right] = 1$$

Por lo tanto Y también tiene una f.d.p uniforme de 0 a 1, lo cual se escribe así $Y \sim U(0,1)$ $f_Y(y) = 1$ y su representación gráfica es



Ejemplo 3: Se dice que la v. a. bidimensional (X,Y) se distribuye uniformemente en la región *R* si tiene f.d.p. Encuentra el valor de la constante *C*.

$$f_{X,Y}(x,y) = \begin{cases} C & (x,y) \in R \\ 0 & \text{o.c.} \end{cases}$$

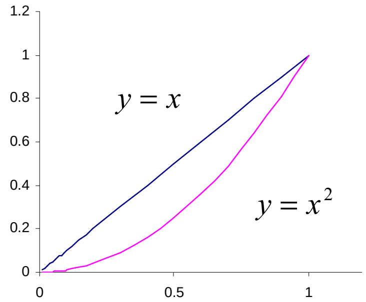
__**Solución**__

$$1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)dxdy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} C\, dxdy = C\iint_R dxdy$$

$$C\left[\text{Area de } R\right] = 1 \quad \therefore \quad C = \frac{1}{\text{Area de } R}$$

Siempre y cuando *R* sea una región de área finita.

Ejemplo 4: 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 conjunta y las marginales de X y de Y respectivamente.



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 $f_{X,Y}(x,y) = 6 \qquad (x,y) \in R$

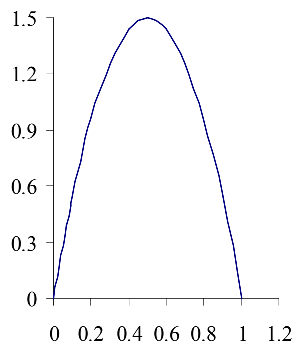
Cálculo de las marginales:

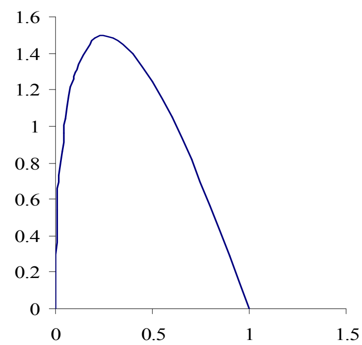
$$f_X(x) = \int_{y=x^2}^{x} 6dy = 6\left[y\Big|_{x^2}^{x}\right] = 6\left[x - x^2\right] \qquad 0 \le x \le 1$$

$$f_Y(y) = \int_{x=y}^{\sqrt{y}} 6dx = 6\left[x\Big|_{y}^{\sqrt{y}}\right] = 6\left[\sqrt{y} - y\right] \qquad 0 \le y \le 1$$

La representación gráfica de las f.d.p marginales para las variables X y Y respectivamente es:

$f_X(x)$ $\qquad\qquad$ $f_Y(y)$





Definición: La función de distribución condicional está dada por:

i) Caso discreto $f_{X|y}\left(x|y\right)=\dfrac{P(X=x,Y=y)}{P(Y=y)}$ con $P(Y=y)>0$

ii) Caso continuo $f_{X|y}\left(x|y\right)=\dfrac{f_{X,Y}(x,y)}{f_Y(y)}$ con $f_Y(y)>0$

Observación:

- En esta definición la variable aleatoria es X y está condicionada al valor particular y de la v.a Y.
- Las dos funciones anteriores cumplen las condiciones de la f. d. p

$$f_{X|y}\geq 0 \quad y \quad \begin{cases} \sum_x f_{X|y}\left(x|y\right)=1 \\ \int_{-\infty}^{\infty} f_{X|y}\left(x|y\right)dx=1 \end{cases}$$

Tarea: Demuestra el punto 2 de la observación.

Definición: La esperanza condicional de la v.a X se define como antes en términos de su f.d.p es decir

$$E\left(X|y\right)=\begin{cases} \sum_x x f_{X|y}\left(x|y\right) & \text{Caso Discreto} \\ \int_{-\infty}^{\infty} x f_{X|y}\left(x|y\right)dx & \text{Caso Continuo} \end{cases}$$

Ejemplo 5: Si (X,Y) tiene f.d.p conjunta $f_{X,Y}(x,y)=x^2+\dfrac{yx}{3}$ $0\leq x\leq 1$, $0\leq y\leq 2$, encuentra las distribuciones condicionales $f_{X|y}\left(x|y\right)$ y $f_{Y|x}\left(y|x\right)$ y verificar que son f.d.p.

<u>**Solución**</u>

Primero se calculan las f.d.p marginales para las variables X y Y respectivamente.

$$f_X(x)=\int_0^2 (x^2+\frac{xy}{3})dy=\left[x^2y+\frac{x}{3}\frac{y^2}{2}\right]_{y=0}^{y=2}=2x^2+\frac{2}{3}x$$

$$f_Y(y)=\int_0^1 (x^2+\frac{xy}{3})dx=\left[\frac{x^3}{3}+\frac{y}{3}\frac{x^2}{2}\right]_0^1=\frac{1}{3}+\frac{1}{6}y$$

Siguiendo la definición de la función de distribución de probabilidad condicional se tiene que:

$$f_{X|y}\left(x|y\right)=\frac{f_{X,Y}(x,y)}{f_Y(y)}=\frac{x^2+\frac{xy}{3}}{\frac{1}{3}+\frac{1}{6}y}=\frac{\frac{3x^2+xy}{3}}{\frac{2+y}{6}}=\frac{6(3x^2+xy)}{3(2+y)}$$

Y la condicional de la variable aleatoria Y dado el valor particular x de la variable aleatoria X, es:

$$f_{Y|x}(y|x)=\frac{f_{X,Y}(x,y)}{f_X(x)}=\frac{x^2+\frac{xy}{3}}{2x^2+\frac{2}{3}x}=\frac{\frac{3x^2+xy}{3}}{\frac{6x^2+2x}{3}}=\frac{3x^2+xy}{6x^2+2x}=\frac{x(3x+y)}{x(6x+2)}$$

Es decir; las f.d.p condicionales son:

$$f_{X|y}\left(x|y\right)=\frac{6x^2+2xy}{2+y} \qquad 0\le x\le 1 \qquad 0\le y\le 2$$

$$f_{Y|x}(y|x)=\frac{3x+y}{6x+2} \qquad 0\le x\le 1 \qquad 0\le y\le 2$$

4.3 Independencia de dos o más variables aleatorias.

Definición: Sea (X,Y) una v. a. bidimensional. Se dice que las variables aleatorias X y Y son independientes si:

i) Caso discreto $P\left(X=x,Y=y\right)=P(X=x)P(Y=y)$

ii) Caso continuo $f_{X,Y}(x,y)=f_X(x)f_Y(y)$

Teorema: Sea (X,Y) una v. a. bidimensional discreta. Entonces X y Y son independientes si y sólo si se tiene que:

i) Caso discreto $P(X=x|Y=y)=P(X=x)$

ii) Caso continuo $f_{X|y}(x,y)=f_X(x)$

Nota._ Equivalentemente las variables aleatorias X y Y son independientes si la f.d.p conjunta se puede escribir como producto de las f.d.p marginales.

Demostración del teorema:

i) Supóngase que $X \perp Y$

Por definición

$$P(X = x|Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)} = \frac{P(X = x)P(Y = y)}{P(Y = y)} = P(X = x)$$

ii) Tarea.

Ejemplo 1: 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

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:

$$0.2 = 0.4(0.5) \qquad 0.2 = 0.4(0.5)$$
$$0.08 = 0.4(0.2) \qquad 0.08 = 0.4(0.2)$$
$$0.12 = 0.4(0.3) \qquad 0.12 = 0.4(0.3)$$

Por lo tanto X y Y sí son independientes.

Ejemplo 2: Sean X y Y la duración de dos dispositivos electrónicos, con f. d. p conjunta dada por

$$f_{X,Y}(x,y) = e^{-(x+y)} \qquad y \geq 0 \quad , \qquad x \geq 0$$

¿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:

$$f_X(x) = \int_0^\infty e^{-x}e^{-y}dy = e^{-x}\left(-e^{-y}\right)\Big|_0^\infty = e^{-x}(1) = e^{-x}$$

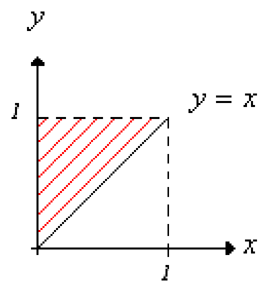
$$f_Y(y) = \int_0^\infty e^{-x}e^{-y}dx = e^{-y}\left(-e^{-x}\right)\Big|_0^\infty = e^{-y}$$

$f_{X,Y}(x,y) = e^{-(x+y)} = e^{-x}e^{-y} = e^{-x}\ e^{-y}$ entonces X y Y sí son independientes.

Nota: Si la función de densidad conjunta se puede factorizar como el producto de una función de una de las variables y otra función de la otra variable y el recorrido de una de las variables no depende del recorrido de la otra variable entonces las variables aleatorias son independientes.

Ejemplo 3: Sea (X,Y) una v. a. bidimensional con f. d. p conjunta $f_{X,Y}(x,y) = 8xy$ en $0 \leq x \leq y \leq 1$.
¿Son independientes X y Y?

<u>Solución</u>



No son independientes ya que a pesar de que la f.d.p conjunta se puede factorizar en una función que depende sólo de X y otra que depende sólo de Y, el recorrido de X depende del recorrido de Y.

Recomendación: Primero se analiza al recorrido de las variables, si estos son independientes, se procede a encontrar las marginales para ver si se cumple que la f.d.p conjunta es igual al producto de las marginales.

Definición: La esperanza de la v.a bidimensional (X, Y) está dada por:

i) Caso discreto: $E(X, Y) = \sum_{y} \sum_{x} xyP(X = x,\ Y = y)$

ii) Caso continuo: $E(X, Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy\ f_{X,Y}(x, y)dxdy$

Teorema.- Si $g(X,Y)$ es una función de la v.a bidimensional (X,Y), entonces la esperanza de la función está dada, como en el caso unidimensional, por la expresión:

i) Caso discreto: $E[g(X, Y)] = \sum_{y} \sum_{x} g(x, y)P(X = x,\ Y = y)$

ii) Caso continuo: $E[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y)\, f_{X,Y}(x, y)dxdy$

Ejemplo 4: Sea (X,Y) una v.a bidimensional con f.d.p conjunta dada por $f_{X,Y}(x, y)$. Encuentra $E[g(X, Y)]$ si $g(X,Y) = X$.

<u>Solución</u>

$$
\begin{aligned}
E[g(X, Y)] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y)\, f_{X,Y}(x, y)dxdy \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x\ f_{X,Y}(x, y)dxdy \\
&= \int_{-\infty}^{\infty} xdx \int_{-\infty}^{\infty} f_{X,Y}(x, y)dy \\
&= \int_{-\infty}^{\infty} xdxf_X(x) \\
&= \int_{-\infty}^{\infty} xf_X(x)dx = E(X)
\end{aligned}
$$

Ejemplo 5: 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.

f.d.p conjunta de (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 $E(Y_1 - Y_2)$

Solución

a)

f.d.p 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:

$$\frac{4}{9} = \binom{n}{0} p^0 q^n \quad \text{con} \quad n = 2$$

$$\frac{4}{9} = q^2 \text{ entonces } q = \sqrt{\frac{4}{9}} = \frac{2}{3} \text{ y por lo tanto } p = \frac{1}{3}$$

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) $E(Y_1 - Y_2) = \sum_{y_2=0}^{2}\sum_{y_1=0}^{2}(y_1 - y_2)P(Y_1 = y_1,\ Y_2 = y_2)$

$$= (0-0)\frac{1}{9} + (1-0)\frac{2}{9} + (2-0)\frac{1}{9} + (0-1)\frac{2}{9} + (1-1)\frac{2}{9}$$

$$+ (2-1)0 + (0-2)\frac{1}{9} + (1-2)0 + (2-2)0$$

$$= \frac{2}{9} + \frac{2}{9} - \frac{2}{9} - \frac{2}{9} = 0$$

Ejemplo 6: Y_1, Y_2 corresponden a las proporciones de tiempo, en un día de trabajo, que los empleados I y II, respectivamente, ocupan realmente en hacer sus tareas asignadas. La función de densidad conjunta para Y_1 y Y_2 está dada por

$$f_{Y_1,Y_2}(y_1, y_2) = \begin{cases} y_1 + y_2 & 0 \le y_1 \le 1\ ,\ 0 \le y_2 \le 1 \\ 0 & \text{o.c} \end{cases}$$

El operador I tiene mayor productividad que el operador II, y una medida de la productividad total de los dos empleados está dada por $30Y_1 + 25Y_2$. Calcula la productividad esperada.

Solución

$$E(30Y_1 + 25Y_2) = \int_{y_2=0}^{1}\int_{y_1=0}^{1}(30y_1 + 25y_2)(y_1 + y_2)dy_1dy_2$$

$$= 30\int_0^1\int_0^1 y_1^2 dy_1dy_2 + 55\int_0^1\int_0^1 y_1y_2dy_1dy_2 + 25\int_0^1\int_0^1 y_2^2 dy_1dy_2$$

$$= 30\left[\int_0^1 y_1^2dy_1\int_0^1 dy_2\right] + 55\left[\int_0^1 y_1dy_1\int_0^1 y_2dy_2\right] + 25\left[\int_0^1 dy_1\int_0^1 y_2^2dy_2\right]$$

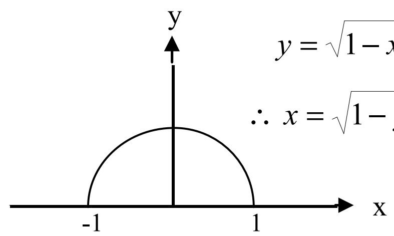
$$= 30\left[\frac{y_1^3}{3}(y_2)\right]_0^1 + 55\left[\frac{y_1^2}{2}\left(\frac{y_2^2}{2}\right)\right]_0^1 + 25\left[y_1\left(\frac{y_2^3}{3}\right)\right]_0^1$$

$$= 30\frac{1}{3}(1) + 55\left(\frac{1}{2}\right)\left(\frac{1}{2}\right) + 25(1)\left(\frac{1}{3}\right)$$

$$= 10 + \left(\frac{55}{4}\right) + \left(\frac{25}{3}\right) = 32.08$$

Ejemplo 7: Supón que (X,Y) se distribuye de manera uniforme sobre el semicírculo del diagrama. De tal modo que $f_{X,Y}(x,y)=\frac{2}{\pi}$ 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.



<u>Solución</u>

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) $$E(X|y)=\int_{-1}^{1}xf_{X|y}(x|y)dx=\int_{-1}^{1}x\frac{1}{2\sqrt{1-y^2}}dx$$

$$=\frac{1}{2\sqrt{1-y^2}}\frac{x^2}{2}\Big|_{-1}^{1}=0$$

$$E(Y|x)=\int_0^1 yf_{Y|x}(y|x)dy=\int_0^1 y\frac{1}{\sqrt{1-x^2}}dy$$

$$=\frac{1}{\sqrt{1-x^2}}\frac{y^2}{2}\Big|_0^1=\frac{1}{2\sqrt{1-x^2}}$$

Ejemplo 8: Dadas las siguientes distribuciones conjuntas determina si X y Y son independientes.

a) $g(X,Y)=4xye^{-\left(x^2+y^2\right)} \qquad x\geq 0,\ y\geq 0$

b) $f(X,Y)=3x^2y^{-3} \qquad 0\leq x\leq y\leq 1$

c) $f(X,Y)=6(1+x+y)^{-4} \qquad x\geq 0,\ y\geq 0$

<u>Solución</u>

a) Independientes.
b) No son independientes porque el recorrido de x depende del recorrido de y.
c) No son independientes.

Ejemplo 9: X y Y corresponden a la vida útil, expresada en horas, de componentes de tipo I y tipo II, respectivamente, en un sistema electrónico. La función de densidad conjunta para X y Y está dada por $f_{X,Y}(x,y)=\frac{1}{8}xe^{-(x+y)/2}$ en $x>0,\ y>0$

Un modo de medir la eficiencia relativa de los dos componentes es el cálculo de la razón $\frac{Y}{X}$. Encuentra la eficiencia relativa esperada.

<u>Solución</u>

$$E\left(\frac{Y}{X}\right)=\int_0^\infty\int_0^\infty \frac{y}{x}\frac{1}{8}xe^{-(x+y)/2}dxdy$$

$$=\int_0^\infty\int_0^\infty \frac{y}{8}e^{-x/2}e^{-y/2}dxdy$$

$$=\frac{1}{8}\int_0^\infty e^{-x/2}dx\int_0^\infty ye^{-y/2}dy$$

$$=\frac{1}{8}\left(-2e^{-x/2}\Big|_0^\infty\right)\left(-2ye^{-y/2}-4e^{-y/2}\right)\Big|_0^\infty=1$$

Ejemplo 10: 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

$$f_{X_1|x_2} = \frac{f_{X_1,X_2}(x_1,x_2)}{f_{X_2}(x_2)} = \frac{6x_1^2 x_2}{2x_2} = 3x_1^2$$

$$f_{X_2|x_1} = \frac{f_{X_1,X_2}(x_1,x_2)}{f_{X_1}(x_1)} = \frac{6x_1^2 x_2}{3x_1^2} = 2x_2$$

Finalmente las esperanzas condicionales son:

$$E(X_1|x_2) = \int_0^1 x_1 3x_1^2 dx_1 = 3\int_0^1 x_1^3 dx_1 = \left.\frac{3x_1^4}{4}\right|_0^1 = \frac{3}{4}$$

$$E(X_2|x_1) = \int_0^1 x_2 f_{X_2|x_1}(x_2 \mid x_1) dx_2 = 2\int_0^1 x_2^2 dx_2 = 2\left.\frac{x_2^3}{3}\right|_0^1 = \frac{2}{3}$$

5 ESTIMACION

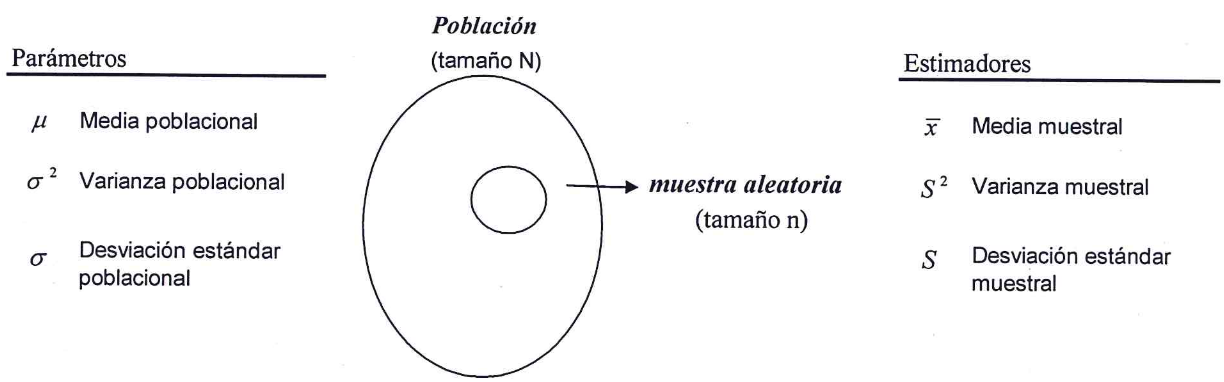
Objetivo: Aprender la construcción y características de los estimadores de los parámetros más comunes, de manera puntual y por intervalo.

Las mediciones hechas de una población reciben el nombre de parámetros, por ejemplo la media poblacional μ, la varianza poblacional σ^2 y la desviación estándar poblacional σ.

Conocer el valor de un parámetro implica realizar un censo lo cual además de ser costoso en ocasiones es casi imposible. Dado que es necesario tener idea de por donde está el valor del parámetro se acostumbra obtener un pedacito de la población, al cual se le llama muestra aleatoria (m.a), se busca que tal muestra sea lo más parecida posible a la población de donde salió para poder decir cosas sobre la población en base al estudio de la muestra. Las mediciones realizadas en una muestra aleatoria se llaman estimadores. Ejemplos de estimadores son la media muestral $\bar{x}$, la varianza muestral S^2 y la desviación estándar muestral S.

Con $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$, $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$ y $S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$, como se observa los estimadores son funciones de la m.a.

Una vez que se tiene la m.a y se sustituyen los valores en la fórmula de un estimador ya se tiene un valor numérico al cual se le llama estimación.

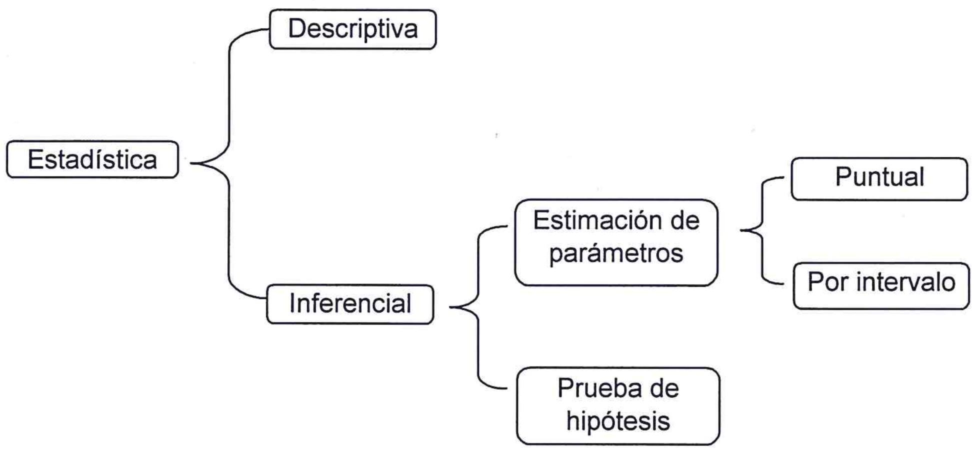


La elección de la muestra aleatoria no es tarea fácil, existe la teoría del muestreo la cual estudia los diferentes tipos de muestreo y su análisis.

5.1 ESTIMACIÓN DE PARÁMETROS.

La parte de las matemáticas que trata de resolver problemas en base al análisis realizado sobre datos que han sido obtenidos de la manera más adecuada posible (en base a un buen diseño de muestreo) se conoce como **estadística**.

La estadística descriptiva organiza, resume y presenta los datos, mientras que la estadística inferencial toma decisiones respecto a una población en base a todo un estudio que hace sobre una parte de tal población, conocida como **muestra**.



5.1.1 ESTIMACIÓN PUNTUAL

Generalmente los parámetros de las distribuciones que describen adecuadamente a un conjunto de datos son desconocidos por lo que surge la necesidad de "*estimarlos*".

Definición 1: Un **estimador** es una función de las variables aleatorias $X_1, X_2, \ldots, X_n$. Es también una variable aleatoria y su función de distribución de probabilidad puede obtenerse a partir de la distribución conjunta de $X_1, X_2, \ldots, X_n$.

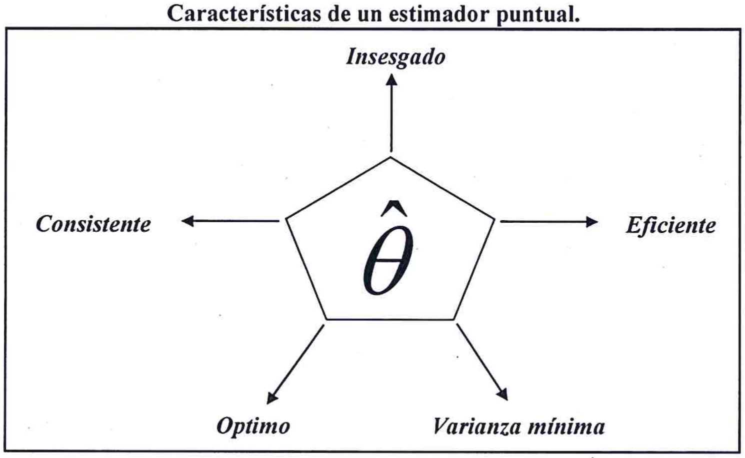
Ejemplo 1: Supón que el parámetro de interés es θ. Entonces $\hat{\theta} = f(X_1, X_2, \ldots, X_n)$ es un estimador de θ.

Existen dos tipos de estimadores; **puntual** si solo se presenta un número como posible valor del parámetro θ, y **por intervalo** si se presenta un conjunto de números dentro del cual puede estar el valor de θ.
Por ejemplo cuando decimos; edad promedio de 35 años, estatura promedio de 1.65m, calificación promedio de 9.3, etc. son estimaciones puntuales.
Una vez que se ha obtenido la muestra aleatoria y sustituido en la función del estimador se obtiene un número al cual se le conoce como estimación.

Se sabe que la media muestral $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ es un estimador de la media poblacional μ. Si se ha seleccionado la m.a de tamaño 3 dada por: $x_1 = 20$, $x_2 = 18$ y $x_3 = 21$ se tiene que una estimación para la media poblacional es $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{20+18+21}{3} = 19.67$.

A todo estimador puntual se le piden ciertas características para considerarlo bueno. Estas características sirven para decidir con cual estimador quedarnos en caso de ser necesario.

Características de un estimador puntual.



Como ya se dijo $\hat{\theta}$ es una v.a y tiene una distribución de probabilidad. Si $\hat{\theta}_1$ y $\hat{\theta}_2$ son estimadores diferentes del parámetro θ tales que $E(\hat{\theta}_1) = \theta$ y $E(\hat{\theta}_2) > \theta$.

¿Qué estimador es "mejor"?

Si tenemos que elegir entre uno de los dos, nos decidiríamos por el primero es decir $\hat{\theta}_1$ ya que en promedio le pega al verdadero valor del parámetro; esto significa que si se repite muchas veces el muestreo obteniendo estimaciones $\hat{\theta}_{11}$, $\hat{\theta}_{12}$, ... , $\hat{\theta}_{1k}$ gran número de estos valores están cerca del parámetro de interés, sin embargo no ocurre lo mismo si usamos $\hat{\theta}_2$ para estimar al parámetro ya que sólo una pequeña proporción de estos valores están cerca de θ.

Pero si las distribuciones de los dos estimadores están centradas en θ con $V(\hat{\theta}_1) < V(\hat{\theta}_2)$. ¿Qué estimador es mejor? Bajo estas circunstancias $\hat{\theta}_1$ es mejor ya que tiene una menor dispersión, esto garantiza que la proporción de estimadores del parámetro que se acercan al parámetro es mayor que la proporción correspondiente de los valores de $\hat{\theta}_2$.

Definición 2: $\hat{\theta}$ es un estimador **insesgado** de θ si $E(\hat{\theta}) = \theta$.

Definición 3: ***i)*** El error cuadrático medio del estimador está dado por $ECM(\hat{\theta}) = E(\theta - \hat{\theta})^2$.

ii) La eficiencia de $\hat{\theta}_1$ relativa a $\hat{\theta}_2$ se define como $Er = \frac{ECM(\hat{\theta}_1)}{ECM(\hat{\theta}_2)}$.

Observaciones:

1) El error cuadrático medio es un medio para comparar estimadores, así, si $ECM(\hat{\theta}_1) < ECM(\hat{\theta}_2)$ decimos que $\hat{\theta}_1$ es "mejor" que $\hat{\theta}_2$.

2) Si $Er < 1$, se dice que $\hat{\theta}_1$ es más **eficiente** que $\hat{\theta}_2$.

3) $\boxed{Sesgo(\hat{\theta}) = E(\hat{\theta}) - \theta}$

$$
\begin{aligned}
ECM(\hat{\theta}) &= E(\theta - \hat{\theta})^2 \\
&= E\left[\theta - \hat{\theta} + E(\hat{\theta}) - E(\hat{\theta})\right]^2 \\
&= E\left[\left(E(\hat{\theta}) - \hat{\theta}\right) - \left(E(\hat{\theta}) - \theta\right)\right]^2 \\
&= E\left[\left(E(\hat{\theta}) - \hat{\theta}\right)^2 - 2\left(E(\hat{\theta}) - \hat{\theta}\right)\left(E(\hat{\theta}) - \theta\right) + \left(E(\hat{\theta}) - \theta\right)^2\right] \\
&= E\left[E(\hat{\theta}) - \hat{\theta}\right]^2 - 2E\left[E(\hat{\theta}) - \hat{\theta}\right]Sesgo(\hat{\theta}) + E\left[Sesgo(\hat{\theta})\right]^2 \\
&= V(\hat{\theta}) - 2Sesgo(\hat{\theta})\left[E(\hat{\theta}) - E(\hat{\theta})\right] + Sesgo^2(\hat{\theta})
\end{aligned}
$$

$$\boxed{ECM(\hat{\theta}) = V(\hat{\theta}) + Sesgo^2(\hat{\theta})}$$

Si el estimador es insesgado, se tiene que $Sesgo(\hat{\theta}) = 0$ y entonces $ECM(\hat{\theta}) = V(\hat{\theta})$.

El estimador insesgado de varianza mínima es aquel que tiene varianza igual a la cota de Cramér Rao, CCR. Los demás estimadores tienen varianza mayor a dicha cota.

$$\boxed{V(\hat{\theta}) \geq \frac{1}{nE\left[\frac{\partial}{\partial\theta}\ln f_X(x;\theta)\right]^2} = CCR}$$

Si $V(\hat{\theta}) = \dfrac{1}{nE\left[\frac{\partial}{\partial\theta}\ln f_X(x;\theta)\right]^2}$ se dice que se tiene el estimador insesgado de **varianza mínima** para el parámetro $\hat{\theta}$.

Ejemplo 2: 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 $\overline{X}$ y la varianza muestral S^2 son estimadores puntuales de la media y la varianza poblacional respectivamente. ¿Son insesgados estos estimadores?

Solución

La media muestral está dada por la expresión $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$. Entonces su esperanza es:

$E(\overline{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \mu$ $\therefore$ $\overline{X}$ sí estima de manera insesgada a la media poblacional.

La varianza muestral esta 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 $E\left(S^2\right)=\dfrac{1}{n-1}\left[\sum_{i=1}^{n}E\left(x_i^2\right)-nE\left(\bar{x}^2\right)\right]$

Recuerda que $\overline{X}\sim N\left(\bar{x};\mu_{\bar{x}}=\mu,\sigma_{\bar{x}}^2=\dfrac{\sigma^2}{n}\right)$ y que $V(X)=E\left(X^2\right)-E^2(X)$ y por tanto $E\left(X^2\right)=V(X)+E^2(X)$ Utilizando toda esta información tenemos que

$$
\begin{aligned}
E\left(S^2\right) &= \frac{1}{n-1}\left[n\left(\sigma^2+\mu^2\right)-n\left(\frac{\sigma^2}{n}+\mu^2\right)\right] \\
&= \frac{1}{n-1}\left(n\sigma^2+n\mu^2-\sigma^2-n\mu^2\right) \\
&= \frac{\sigma^2(n-1)}{n-1}=\sigma^2
\end{aligned}
$$

Entonces S^2 estima de manera insesgada a la varianza poblacional.

Ejemplo 3: Supón que se tiene una m.a de tamaño $2n$ de una población denotada por X con media μ y varianza σ^2 y que $\hat{\mu}_1=\dfrac{1}{2n}\sum_{i=1}^{2n}X_i$ y $\hat{\mu}_2=\dfrac{1}{n}\sum_{i=1}^{n}X_i$ son dos estimadores para la media poblacional μ. ¿Cuál estimador recomendarías?

<u>Solución</u>

$$E\left(\hat{\mu}_1\right)=E\left[\frac{1}{2n}\sum_{i=1}^{2n}X_i\right]=\mu$$

$$E\left(\hat{\mu}_2\right)=E\left[\frac{1}{n}\sum_{i=1}^{n}X_i\right]=\mu$$

Ambos estimadores son insesgados, entonces hay que considerar otro criterio. Veamos cuales son sus errores cuadráticos medios.

$$ECM(\hat{\mu}_1) = V(\overline{X}) = \frac{\sigma^2}{2n}$$

$$ECM(\hat{\mu}_2) = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

$$Er = \frac{\sigma^2}{2n}\frac{n}{\sigma^2} = \frac{1}{2} < 1$$

Entonces recomendaríamos $\hat{\mu}_1$ porque estima de manera más eficiente a la media poblacional.

Ejemplo 4: Si $X_1, X_2, \ldots, X_7$ es una m.a de una población que tiene media μ y varianza σ^2. Considera los siguientes estimadores de μ:

$$\hat{\theta}_2 = \frac{1}{2}(2X_1 - X_6 + X_4) \qquad \hat{\theta}_1 = \frac{1}{7}\sum_{i=1}^{7} X_i$$

a) ¿Alguno de estos estimadores es insesgado?
b) ¿Cuál es mejor? ¿Por qué?

Solución

a) $E(\hat{\theta}_1) = \frac{1}{7}(7\mu) = \mu$ y $E(\hat{\theta}_2) = \frac{1}{2}(2-1+1)\mu = \mu$. Ambos estiman de manera insesgada a la media poblacional.

b) $ECM(\hat{\theta}_1) = V(\overline{X}) = \frac{\sigma^2}{7}$

$ECM(\hat{\theta}_2) = V(\hat{\theta}_2) = \frac{1}{4}(6\sigma^2) = \frac{3}{2}\sigma^2$

Entonces $\hat{\theta}_1$ es más eficiente que $\hat{\theta}_2$.

Ejemplo 5: Considera que se toman tres muestras aleatorias de tamaños 10, 8 y 6 de una población con media μ y varianza σ^2. S_1^2, S_2^2 y S_3^2 las varianzas correspondientes de las muestras. ¿Será S^2 un estimador insesgado de la varianza poblacional? $S^2 = \frac{1}{24}(10S_1^2 + 8S_2^2 + 6S_3^2)$

Solución

Ya vimos que $E(S^2) = \sigma^2$ entonces $E(S^2) = \frac{1}{24}(10+8+6)\sigma^2 = \sigma^2$, por tanto sí es un estimador insesgado del la varianza poblacional.

Ejemplo 6: Supón que Y_1, Y_2, Y_3 forman una muestra aleatoria de una distribución exponencial con función de densidad $f_Y(y;\theta) = \frac{1}{\theta} exp\left(-\frac{y}{\theta}\right)$ con $y > 0$. Considera los siguientes estimadores para θ:

$\hat{\theta}_1 = Y_1$, $\hat{\theta}_2 = \frac{Y_1 + Y_2}{2}$, $\hat{\theta}_3 = \frac{Y_1 + 2Y_2}{3}$, $\hat{\theta}_4 = \overline{Y}$.

a) ¿Cuáles de estos son estimadores insesgados?
b) Considerando sólo a los estimadores insesgados ¿Cuál tiene menor varianza?

Solución

a) Todos son insesgados.

b) $V(\hat{\theta}_1) = \sigma^2$, $V(\hat{\theta}_2) = \frac{\sigma^2}{2}$, $V(\hat{\theta}_3) = \frac{5\sigma^2}{9}$, $V(\hat{\theta}_4) = \frac{\sigma^2}{3}$. Entonces $\hat{\theta}_4$ tiene menor varianza.

Ejemplo 7: 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

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 $f_X(x;\mu,\sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$

$$\ln f_X(x;\mu,\sigma^2) = \ln\frac{1}{\sigma\sqrt{2\pi}} - \frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2$$

$$\frac{\partial}{\partial\mu}\ln f_X(x;\mu,\sigma^2) = \frac{1}{2\sigma^2}[2(x-\mu)] = \frac{x-\mu}{\sigma^2}$$

$$E\left[\frac{\partial}{\partial\mu}\ln f_X(x;\mu,\sigma^2)\right]^2 = E\left[\frac{1}{\sigma^4}(x-\mu)^2\right] = \frac{1}{\sigma^4}E(x-\mu)^2 = \frac{1}{(\sigma^2)^2}V(X) = \frac{1}{\sigma^2}$$

Entonces $CCR = \frac{\sigma^2}{n} = V(\overline{X})$ por lo tanto $\overline{X}$ es el estimador insesgado de varianza mínima para μ.

Definición 4:
i) Se dice que $\hat{\theta}^*$ es un estimador **óptimo** de θ si $ECM(\hat{\theta}^*) \le ECM(\hat{\theta}) \quad \forall\hat{\theta}$.
ii) Se dice que $\hat{\theta}_n$ (estimador de θ basado en una muestra de tamaño *n*) es **consistente** para θ si $\lim_{n\to\infty} ECM(\hat{\theta}_n) = 0$.

Ejemplo 8: $\overline{X}$ es un estimador óptimo y consistente de μ para $N(\mu,\sigma^2)$ ya que

$$V(\overline{X}) = ECM(\overline{X}) = CCR \text{ y } \lim_{n\to\infty} ECM(\overline{X}) = \lim_{n\to\infty}\frac{\sigma^2}{n} = 0$$

MÉTODO DE LOS MOMENTOS PARA ESTIMAR PARÁMETROS.

Este método da por hecho que los momentos muestrales son una buena aproximación de los momentos poblacionales, por lo tanto se igualan ambos momentos y de ahí se despeja el estimador del parámetro de interés.

$\mu_k = E(X^k)$ k-ésimo momento poblacional de la v.a X.

$m_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$ k-ésimo momento muestral de la v.a X.

$\boxed{\mu_k = m_k}$ y de aquí se despeja el estimador.

Al estimador del parámetro θ obtenido por el método de los momentos se le pone una tilde de la forma $\tilde{\theta}$.

Ejemplo 9: Sea $X \sim N(x;\mu,\sigma^2)$ estima $\underline{\theta} = (\mu,\sigma^2)$ en base a una muestra aleatoria de tamaño n usando el método de los momentos.

Solución

Tenemos dos parámetros a estimar por lo tanto necesitamos los dos primeros momentos muestrales y los dos primeros momentos poblacionales.
Se igualan los primeros momentos; $\tilde{\mu} = \bar{x}$
Se igualan también los segundos momentos

$\mu_2 = E(X^2) = \sigma^2 + \mu^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 = m_2$ entonces $\boxed{\hat{\mu} = \bar{x}}$

$$\tilde{\sigma}^2 + \tilde{\mu}^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2$$

$$\tilde{\sigma}^2 + (\bar{x})^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2$$

$$\tilde{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - (\bar{x})^2 = \frac{1}{n}\left[\sum_{i=1}^{n} x_i^2 - n(\bar{x})^2\right]$$

$$\boxed{\tilde{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

Ejemplo 10: Determina los estimadores por momentos para los parámetros α y β de la distribución gamma, en base a una m.a de tamaño n.

Solución

Igualando los dos primeros momentos poblacionales con los muestrales tenemos un sistema de dos ecuaciones con dos incógnitas a resolver.

$$\mu_1 = \frac{\tilde{\alpha}}{\tilde{\beta}} = \overline{X} = m_1$$

$$\mu_2 = \frac{\tilde{\alpha}}{\tilde{\beta}^2} + \frac{\tilde{\alpha}^2}{\tilde{\beta}^2} = \frac{1}{n}\sum_{i=1}^{n} X_i^2 = m_2$$

Despejando $\tilde{\alpha}$ de la primera ecuación y sustituyendo en la segunda se tiene:

$$\tilde{\alpha} = \tilde{\beta}\overline{X}$$

$$\frac{\tilde{\beta}\overline{X}}{\tilde{\beta}^2} + \frac{\tilde{\beta}^2\overline{X}^2}{\tilde{\beta}^2} = \frac{1}{n}\sum_{i=1}^{n} X_i^2$$

$$\frac{\overline{X}}{\tilde{\beta}} + \overline{X}^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$$

$$\frac{\overline{X}}{\tilde{\beta}} = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i^2 - \overline{X}\right)^2$$

Finalmente, los estimadores por el método de los momentos para los parámetros de la distribución Gamma son: $\tilde{\beta} = \dfrac{n\overline{X}}{\sum_{i=1}^{n}\left(X_i^2 - \overline{X}\right)^2}$ y $\tilde{\alpha} = \tilde{\beta}\overline{X} = \dfrac{n\overline{X}^2}{\sum_{i=1}^{n}\left(X_i^2 - \overline{X}\right)^2}$

Ejemplo 11: Sea X una v.a binomial con parámetros n y p desconocidos, encuentra los estimadores por el método de los momentos para tales parámetros en base a una muestra aleatoria de tamaño k.

Solución

$$\mu_1 = np = \overline{X} = m_1$$

$$\mu_2 = np(1-p) + n^2p^2 = \frac{1}{k}\sum_{i=1}^{k} X_i^2 = m_2$$

$$\overline{X}(1-\tilde{p}) + \overline{X}^2 = \frac{1}{k}\sum_{i=1}^{k} X_i^2$$

$$\overline{X} - \tilde{p}\overline{X} + \overline{X}^2 = \frac{1}{k}\sum_{i=1}^{k} X_i^2$$

$$\overline{X} - \tilde{p}\overline{X} = \frac{1}{k}\sum_{i=1}^{k} X_i^2 - \overline{X}^2 = \frac{1}{k}\sum_{i=1}^{k}\left(X_i - \overline{X}\right)^2$$

$$\tilde{p}\overline{X} = \overline{X} - \frac{1}{k}\sum_{i=1}^{k}\left(X_i - \overline{X}\right)^2$$

$$\boxed{\tilde{p} = 1 - \frac{1}{k\overline{X}}\sum_{i=1}^{k}\left(X_i - \overline{X}\right)^2}$$

$$\boxed{\tilde{n} = \frac{\overline{X}}{\tilde{p}} = \frac{k\overline{X}^2}{k\overline{X} - \sum_{i=1}^{k}\left(X_i - \overline{X}\right)^2}}$$

Ejemplo 12: Sea una población con distribución Geométrica con parámetro p. Encuentra el estimador para el parámetro usando el método de los momentos y en base a una muestra aleatoria de tamaño n.

Solución

$$\mu_1 = \frac{1}{\tilde{p}} = \overline{X} = m_1 \quad \therefore \quad \tilde{p} = \frac{1}{\overline{X}}$$

MÉTODO DE MÁXIMA VEROSIMILITUD.

Definición 5: La función de verosimilitud de n variables aleatorias se define como su densidad conjunta $f_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots,x_n)$. En particular si $X_1,X_2,\ldots,X_n$ es una muestra aleatoria de la densidad $f(x;\theta)$ entonces la función de verosimilitud de la muestra está dada por

$$L(\theta) = f_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots,x_n) = f(x_1;\theta)f(x_2;\theta)\cdots f(x_n;\theta)$$

$$\boxed{L(\theta) = \prod_{i=1}^{n} f(x_i;\theta)}$$

Definición 6: Se dice que $\hat{\theta}$ es un estimador máximo verosímil si maximiza la función de verosimilitud.

Recuerda que para maximizar se deriva parcialmente respecto al parámetro de interés, se iguala a cero y se despeja el valor que hace posible la ecuación.
Por facilidad de cálculo y dado que las funciones $L(\theta)$ y $\ln L(\theta)$ se maximizan en el mismo punto, se acostumbra encontrar al estimador máximo verosímil resolviendo la ecuación $\frac{\partial}{\partial\theta}\ln L(\theta) = 0$

Cuando se tiene un vector de parámetros a estimar por ejemplo $\underline{\theta} = (\theta_1,\theta_2,\ldots,\theta_k)$ la manera de proceder es resolver las k ecuaciones siguientes:

$$\frac{\partial}{\partial\theta_1}\ln L(\underline{\theta}) = \frac{\partial}{\partial\theta_1}\ln L(\theta_1,\theta_2,\ldots,\theta_k) = 0$$

$$\frac{\partial}{\partial \theta_2} \ln L(\underline{\theta}) = \frac{\partial}{\partial \theta_2} \ln L(\theta_1, \theta_2, \ldots, \theta_k) = 0$$

$$\vdots$$

$$\frac{\partial}{\partial \theta_k} \ln L(\underline{\theta}) = \frac{\partial}{\partial \theta_k} \ln L(\theta_1, \theta_2, \ldots, \theta_k) = 0$$

Ejemplo 13: Encuentra el estimador máximo verosímil de p en base a una m.a de tamaño n de una población Bernoulli.

Solución

$$f_X(x;p) = p^x (1-p)^{1-x} \text{ si } x = 0,\ 1$$

i) La función de verosimilitud de una m.a de tamaño n

$$L(p) = \prod_{i=1}^{n} f(x_i, p) = \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i} = p^{\sum_{i=1}^{n} x_i}(1-p)^{n-\sum_{i=1}^{n} x_i}$$

ii) Sacar 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) Sacar 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) Igualar a cero y resolver 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.

$$\frac{\sum_{i=1}^{n} x_i}{\hat{p}} - \frac{n - \sum_{i=1}^{n} x_i}{1-\hat{p}} = 0$$

$$\frac{\sum_{i=1}^{n} x_i}{\hat{p}} = \frac{n - \sum_{i=1}^{n} x_i}{1-\hat{p}}$$

$$\frac{1-\hat{p}}{\hat{p}} = \frac{n - \sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i}$$

$$\frac{1}{\hat{p}} - 1 = \frac{n}{\sum_{i=1}^{n} x_i} - 1$$

$$\frac{1}{\hat{p}} = \frac{n}{\sum_{i=1}^{n} x_i} \therefore \underline{\hat{p} = \bar{x}}$$

Ejemplo 14: Sea $X_1, X_2, \ldots, X_n$ una m.a de una densidad Normal con media μ y varianza σ^2 encuentra el estimador máximo verosímil de $\underline{\theta} = (\mu, \sigma^2)$.

Solución

$$f_X\left(x;\mu,\sigma^2\right)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad \text{si } x>0$$

i) La función de verosimilitud de una m.a de tamaño n es:

$$L\left(\mu,\sigma^2\right)=\prod_{i=1}^{n} f\left(x_i,\mu,\sigma^2\right)=\prod_{i=1}^{n}\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x_i-\mu}{\sigma}\right)^2}$$

$$=\left(\sigma\sqrt{2\pi}\right)^{-n}e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}$$

ii) Sacar el logaritmo natural de la función de verosimilitud.

$$\ln L\left(\mu,\sigma^2\right)=-n\ln\sigma-n\ln\sqrt{2\pi}-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

iii) Sacar las derivadas parciales respecto a μ y respecto a σ^2.

$$\frac{\partial}{\partial\mu}\ln L\left(\mu,\sigma^2\right)=\frac{2}{2\sigma^2}\sum_{i=1}^{n}2(x_i-\mu)=\frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu)$$

$$\frac{\partial}{\partial\sigma^2}\ln L\left(\mu,\sigma^2\right)=-\frac{n}{2\sigma^2}+\frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2$$

iv) Igualar a cero y resolver el sistema para encontrar los estimadores máximo verosímiles de los parámetros de interés.

$$0=\frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu)=\frac{1}{\sigma^2}\left[\sum_{i=1}^{n}x_i-n\hat{\mu}\right] \quad \therefore \quad \hat{\mu}=\bar{x}$$

$$0=-\frac{n}{2\sigma^2}+\frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2$$

$$n=\frac{1}{\hat{\sigma}^2}\sum_{i=1}^{n}(x_i-\mu)^2 \quad \therefore \quad \hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2$$

Observa que el estimador máximo verosímil para la media poblacional de la distribución Normal es la media muestral que además es un estimador insesgado, óptimo y consistente. Y que el estimador máximo verosímil para la varianza poblacional de una distribución Normal es un estimador sesgado pero consistente.

Definición 7: Supón que $X_1,X_2,\ldots,X_n$ es una muestra aleatoria de tamaño n en orden de aparición, si ordenamos esta muestra de menor a mayor tenemos lo que se conoce como estadísticas de orden y se denota $X_{(1)},X_{(2)},\ldots,X_{(n)}$ tales que $X_{(1)}\le X_{(2)}\le\ldots\le X_{(n)}$.

Ejemplo 15: Se toma una muestra aleatoria de 5 estudiantes y se registra su calificación en el segundo parcial de probabilidad y estadística, siendo las calificaciones en orden de aparición $x_1 = 10$, $x_2 = 5.3$, $x_3 = 8$, $x_4 = 6.7$, $x_5 = 1$, las estadísticas de orden correspondientes son $x_{(1)} = 1$, $x_{(2)} = 5.3$, $x_{(3)} = 6.7$, $x_{(4)} = 8$, $x_{(5)} = 10$.

Ejemplo 16: Se toma una muestra aleatoria de tamaño n de una población distribuida uniformemente en el intervalo $(0, a)$ ¿Cuál es el estimador máximo verosímil para a?

<u>**Solución**</u>

i) La función de verosimilitud está dada por

$$L(a) = \prod_{i=1}^{n} f(x_i, a) = \prod_{i=1}^{n} \frac{1}{a} = \frac{1}{a^n}$$

ii) Encontrar el valor de a que maximiza la función de verosimilitud.

Observa que esta función es máxima cuando a toma el valor más pequeño permisible que es precisamente la última estadística de orden, es decir $\hat{a} = X_{(n)}$

Ejemplo 17: 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?

<u>**Solución**</u>

i) La función de verosimilitud está dada por

$$L(\lambda) = \prod_{i=1}^{n} \frac{\lambda^{x_i} e^{-\lambda}}{x_i!} = e^{-\lambda n} \frac{\lambda^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!}$$

$$\ln L(\lambda) = -\lambda n + \sum_{i=1}^{n} x_i \ln(\lambda) - \ln \prod_{i=1}^{n} x_i!$$

$$\frac{\partial}{\partial \lambda} \ln L(\lambda) = -n + \sum_{i=1}^{n} \frac{x_i}{\lambda}$$

ii) Igualando a cero y resolviendo se encuentra el estimador máximo verosímil.

$$n = \sum_{i=1}^{n} \frac{x_i}{\lambda} \quad \therefore \quad \hat{\lambda} = \frac{1}{n} \sum_{i=1}^{n} x_i = \bar{x}$$

b) Valor esperado y varianza del estimador.

$$E\left(\overline{X}\right)=\lambda \text{ y } V\left(\overline{X}\right)=\frac{\lambda}{n}$$

c) $\lim\limits_{n\to\infty} V\left(\overline{X}\right)=\lim\limits_{n\to\infty}\frac{\lambda}{n}=0$ entonces sí es consistente.

Ejemplo 18: Si $X_1, X_2, \ldots, X_n$ es una muestra aleatoria de una población con f.d.p $f_X(x;\theta)=\frac{r}{\theta}x^{r-1}e^{-x^r/\theta}$ con $\theta>0$ y $x>0$ con r constante y positiva. Encuentra el estimador máximo verosímil para θ.

Solución

i) La función de verosimilitud está dada por

$$L(\theta)=\prod_{i=1}^{n}\frac{r}{\theta}x_i^{r-1}e^{-x_i^r/\theta}=\frac{r^n}{\theta^n}e^{-\frac{1}{\theta}\sum_{i=1}^{n}x_i^r}\prod_{i=1}^{n}x_i^{r-1}$$

ii) Sacar el logaritmo natural de la función de verosimilitud.

$$\begin{aligned}\ln L(\theta)&=\ln\left(\frac{r^n}{\theta^n}\right)-\frac{1}{\theta}\sum_{i=1}^{n}x_i^r+\sum_{i=1}^{n}\ln\left(x_i^{r-1}\right)\\&=n\ln(r)-n\ln(\theta)-\frac{1}{\theta}\sum_{i=1}^{n}x_i^r+(r-1)\sum_{i=1}^{n}\ln(x_i)\end{aligned}$$

$$\frac{\partial}{\partial\theta}\ln L(\theta)=-\frac{n}{\theta}+\frac{1}{\theta^2}\sum_{i=1}^{n}x_i^r$$

iii) Igualando a cero y resolviendo se tiene que

$\frac{n}{\hat{\theta}}=\frac{1}{\hat{\theta}^2}\sum_{i=1}^{n}x_i^r$ despejando al estimador del parámetro se tiene que $\hat{\theta}=\frac{1}{n}\sum_{i=1}^{n}x_i^r$

Ejemplo 19: Supón que $X_1, X_2, \ldots, X_m$ es una m.a de la producción por acre de la variedad A de trigo, la cual tiene una distribución Normal con media μ_1 y varianza σ^2 y que $Y_1, Y_2, \ldots, Y_n$ es una m.a de la producción por acre de la variedad B de trigo, la cual se distribuye como una Normal con media μ_2 y varianza σ^2. Si $X \perp Y$ encuentra el estimador máximo verosímil para la varianza común σ^2 si se desconocen las medias poblacionales.

Solución

$X \sim N\left(x;\mu_1,\sigma^2\right)$ y $Y \sim N\left(y;\mu_2,\sigma^2\right)$

Sabemos que el estimador máximo verosímil para la media poblacional está dado por la media muestral es decir $\hat{\mu}_1 = \overline{X}$ y $\hat{\mu}_2 = \overline{Y}$

La función de verosimilitud para la muestra conjunta está dada por

$$\begin{aligned}
L(\mu_1, \mu_2, \sigma^2) &= f(X_1, X_2, \ldots, X_m, Y_1, Y_2, \ldots, Y_n) \\
&= \prod_{i=1}^{m} f(x_i; \mu_1, \sigma^2) \prod_{i=1}^{n} f(y_i; \mu_2, \sigma^2) \\
&= \prod_{i=1}^{m} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x_i-\mu_1)^2} \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(y_i-\mu_2)^2} \\
&= \prod_{i=1}^{m} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x_i-\mu_1)^2} \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(y_i-\mu_2)^2} \\
&= (2\pi)^{-\frac{m+n}{2}} (\sigma^2)^{-\frac{m+n}{2}} \exp\left(-\frac{1}{2\sigma^2} \left[\sum_{i=1}^{m} (X_i - \mu_1)^2 + \sum_{i=1}^{n} (Y_i - \mu_2)^2 \right] \right)
\end{aligned}$$

Sacar el logaritmo natural a la función de máximo verosímilitud

$$\ln L(\mu_1, \mu_2, \sigma^2) = \ln(2\pi)^{-\frac{m+n}{2}} - \frac{m+n}{2} \ln(\sigma^2) - \frac{1}{2\sigma^2} \left[\sum_{i=1}^{m} (X_i - \mu_1)^2 + \sum_{i=1}^{n} (Y_i - \mu_2)^2 \right]$$

Derivando parcialmente respecto a la varianza

$$\frac{\partial}{\partial\sigma^2} \ln L(\mu_1, \mu_2, \sigma^2) = -\frac{m+n}{2\sigma^2} + -\frac{1}{2\sigma^4} \left[\sum_{i=1}^{m} (X_i - \mu_1)^2 + \sum_{i=1}^{n} (Y_i - \mu_2)^2 \right]$$

Igualando a cero y despejando se tiene que el estimador máximo verosímil para la varianza común es:

$$\hat{\sigma}^2 = \frac{1}{m+n} \left[\sum_{i=1}^{m} (X_i - \overline{X})^2 + \sum_{i=1}^{n} (Y_i - \overline{Y})^2 \right]$$

5.1.2 ESTIMACIÓN POR INTERVALO.

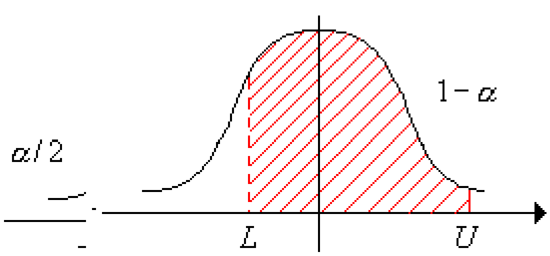
Cuantas veces hemos escuchado expresiones como "*tenía entre 25 y 30 años*", "*yo pienso que saco entre 8 y 10 en el examen de mate*", "*espero llegar entre 10:30 y 10:45*", "*tiene una estatura entre 1.75m y 1.85m*", etc. Todas estas expresiones en realidad son estimaciones por intervalo y muchas veces una estimación de este tipo nos da más información que una estimación puntual. Y la manera de hacer estimaciones por intervalo es construyendo intervalos de confianza.

5.2 INTERVALOS DE CONFIANZA.

Un intervalo de confianza (IC) es un intervalo de la forma $[L,\ U]$ que se construye de tal manera que exista una probabilidad grande de que el parámetro de interés se encuentre dentro del intervalo. Si α es un número pequeño entonces $(1-\alpha)$ es un número grande y si $P(L \le \theta \le U) = 1-\alpha$, se dice que $[L,\ U]$ es un intervalo para el parámetro θ con una confianza de $(1-\alpha)100\%$.

CASO 1._ IC para la media poblacional, cuando la muestra es grande y la varianza conocida. ($n \ge 30$ y σ^2 conocida)

Queremos estimar la media poblacional, como es de suponerse el estimador puntual idóneo es la media muestral $\overline{X}$, en base a una muestra aleatoria de tamaño n de una densidad con media μ desconocida y varianza σ^2 conocida se tiene que el TCL garantiza que $\overline{X} \sim N\left(\mu, \dfrac{\sigma^2}{n}\right)$ entonces $Z = \dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$



Queremos encontrar L y U tales que $P(L \le Z \le U) = 1-\alpha$. ¿Cómo escoger L y U de tal manera que el intervalo sea lo más angosto posible?

Aprovechando la simetría de la distribución Normal tomamos $L = -U$

La construcción del intervalo es como sigue:

$$1-\alpha = P\left(-Z_{\alpha/2} \le Z \le Z_{\alpha/2}\right)$$

$$= P\left(-Z_{\alpha/2} \le \frac{\overline{X} - \mu}{\sigma / \sqrt{n}} \le Z_{\alpha/2} \right)$$

$$= P\left(-\frac{\sigma}{\sqrt{n}} Z_{\alpha/2} \le \overline{X} - \mu \le \frac{\sigma}{\sqrt{n}} Z_{\alpha/2} \right) \text{ Multiplicando por } -1 \text{ y reacomodando}$$

$$= P\left(-\frac{\sigma}{\sqrt{n}} Z_{\alpha/2} \le \mu - \overline{X} \le \frac{\sigma}{\sqrt{n}} Z_{\alpha/2} \right) \text{ Sumando } \overline{X}$$

$$= P\left(\overline{X} - \frac{\sigma}{\sqrt{n}} Z_{\alpha/2} \le \mu \le \overline{X} + \frac{\sigma}{\sqrt{n}} Z_{\alpha/2} \right)$$

El intervalo para μ con una confianza de $(1-\alpha)100\%$ está dado por $\boxed{\left[\overline{X} - \frac{\sigma}{\sqrt{n}} Z_{\alpha/2},\ \overline{X} + \frac{\sigma}{\sqrt{n}} Z_{\alpha/2} \right]}$

Ejemplo 1: Considera el intervalo de confianza para μ con σ conocida $\overline{X} - \frac{\sigma}{\sqrt{n}} Z_{\alpha_2} \le \mu \le \overline{X} + \frac{\sigma}{\sqrt{n}} Z_{\alpha_1}$ donde $\alpha_1 + \alpha_2 = \alpha$, si $\alpha = 0.05$ obtén el intervalo para $\alpha_1 = \alpha_2 = \alpha / 2 = 0.025$. Después determina el intervalo para $\alpha_1 = 0.01$ y $\alpha_2 = 0.04$. ¿Cuál es el intervalo más angosto? ¿Hay ventajas para un intervalo de confianza simétrico?

<u>**Solución**</u>

a) $$\left[\overline{X} - \frac{\sigma}{\sqrt{n}} Z_{0.025},\ \ \overline{X} + \frac{\sigma}{\sqrt{n}} Z_{0.025} \right] = \left[\overline{X} - \frac{\sigma}{\sqrt{n}} 1.96,\ \overline{X} + \frac{\sigma}{\sqrt{n}} 1.96 \right]$$

b) $$\left[\overline{X} - \frac{\sigma}{\sqrt{n}} Z_{0.04},\ \ \overline{X} + \frac{\sigma}{\sqrt{n}} Z_{0.01} \right] = \left[\overline{X} - \frac{\sigma}{\sqrt{n}} 1.75,\ \overline{X} + \frac{\sigma}{\sqrt{n}} 2.35 \right]$$

La longitud del I.C en a) es $2(1.96)\frac{\sigma}{\sqrt{n}} = 3.92\frac{\sigma}{\sqrt{n}}$

Y la longitud del I.C en b) es $(1.75 + 2.35)\frac{\sigma}{\sqrt{n}} = 4.10\frac{\sigma}{\sqrt{n}}$. La ventaja del intervalo simétrico es que es más angosto.

La precisión de la estimación por intervalo esta dada por la longitud del intervalo, mientras más angosto sea el I.C se considera más precisa la estimación. Si el intervalo es simétrico a medida que la confianza aumenta la precisión disminuye.

Ejemplo 2: Es habitual la manifestación tardía de lesiones después de la exposición a dosis suficientes de radiación. Se obtuvieron los datos siguientes del tiempo en días que transcurren entre la exposición a la radiación y la aparición de eritema (enrojecimiento de la piel) de intensidad máxima:

16	12	14	16	13	9	15	7
20	19	11	14	9	13	11	3

8	21	16	16	12	16	14	20
7	14	18	14	18	13	11	16
18	16	11	13	14	16	15	15

a) Aunque el tiempo en el cual aparece el eritema se redondea al día más cercano, en realidad es una v.a continua cuya distribución es Normal con desviación estándar de 4 días. Construye un IC al 95% para la media de tiempo para la aparición del eritema.
b) ¿Te sorprendería la afirmación de que $\mu = 17$ días? Explica tu respuesta.

Ejemplo 3: En estudios, se ha comprobado que la v.a X, el tiempo de procesamiento requerido para una multiplicación en una nueva computadora tridimensional, tiene distribución normal, con media μ y desviación estándar de 2 μs (microsegundos). Se toma una muestra aleatoria de 16 observaciones.
a) ¿Cuál es la distribución de $\overline{X}$?
b) Se obtiene los datos siguientes:

42.65	45.15	39.32	44.44
41.63	41.54	41.59	45.68
46.50	41.35	44.37	40.27
43.87	43.79	43.28	40.70

A partir de estos datos, encuentra una estimación insesgada para la media poblacional.
c) Construye un IC al 95% para la media poblacional. ¿Te sorprendería la afirmación de que el tiempo promedio requerido para procesar una multiplicación en el sistema sea de 42.2 μs ? Explica.

Ejemplo 4: En un estudio de diversos sistemas de cómputo, se considera la v.a X, el número de archivos que se ha almacenado. La experiencia indica que la desviación estándar es de 5 archivos. Para los datos siguientes:

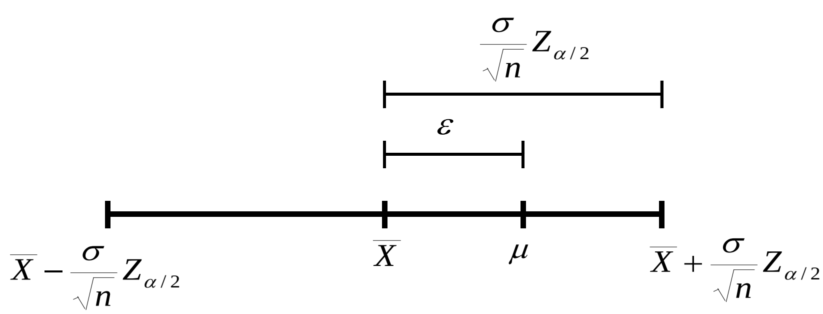
7	8	4	5	9	9
4	12	8	1	8	7
3	13	2	1	17	7
12	5	6	2	1	13
14	10	2	4	9	11
3	5	12	6	10	7

a) Encuentra una estimación insesgada para el número promedio de archivos almacenados.
b) ¿Cuál es la distribución aproximada de la media muestral?
c) Construye un IC al 98% para el número promedio de archivos almacenados.
d) En una descripción del tamaño del sistema, un ejecutivo afirma que el número promedio de archivos almacenados es mayor de 10 ¿Qué opinas al respecto?

5.3 ERROR ESTÁNDAR Y TAMAÑO DE LA MUESTRA.

Todo estadístico tienen su varianza, por ejemplo la media muestral tiene varianza igual a la varianza de la población de donde salió la muestra dividido por el tamaño de la muestra, $\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}$, y la raíz cuadrada positiva de esta varianza es lo que se conoce como **error estándar**, en este caso el error estándar de la media muestral está dado por $\varepsilon\varepsilon = \dfrac{\sigma}{\sqrt{n}}$.

Considera ahora el intervalo de confianza para una media poblacional con varianza conocida, el cual está centrado en $\overline{X}$ con límite inferior $\overline{X} - \dfrac{\sigma}{\sqrt{n}} Z_{\alpha/2}$ y límite superior $\overline{X} + \dfrac{\sigma}{\sqrt{n}} Z_{\alpha/2}$. Por la manera en que fue construido el I.C se tiene que $(1-\alpha)100$ de las veces el parámetro cae dentro del intervalo como se muestra en la figura siguiente.



La distancia entre el parámetro y la estimación es un error $\varepsilon = \left|\mu - \overline{X}\right|$ entonces $(1-\alpha)100$ de las veces se tiene que $\varepsilon \le \dfrac{\sigma}{\sqrt{n}} Z_{\alpha/2}$ o equivalentemente $\sqrt{n} \le \dfrac{\sigma\, Z_{\alpha/2}}{\varepsilon}$. El intervalo más corto se obtiene cuando n es lo más grande posible, es decir cuando $\boxed{n = \left[\dfrac{\sigma\, Z_{\alpha/2}}{\varepsilon}\right]^2}$

Ejemplo 5: Se sabe que la vida en horas de una bombilla eléctrica de 75 watts se distribuye aproximadamente en forma Normal, con desviación estándar $\sigma = 25$ hrs. Una muestra aleatoria de 20 bombillas tiene una vida media de $\overline{x} = 1014$hrs.

a) Construye un intervalo de confianza de dos lados al 95% respecto a la vida media.
b) Construye un I.C inferior del 95% respecto a la vida media.
c) Se desea que el ancho total del intervalo de confianza respecto a la vida media sea de ocho horas. ¿Qué tamaño de muestra debe utilizarse?

<u>Solución</u>

$n = 20$
$\sigma = 25$
$\overline{x} = 1014$hrs.

a) $1-\alpha = 0.95$ entonces $\alpha = 0.05$ y $\dfrac{\alpha}{2} = 0.025$ por lo tanto $z_{\alpha/2} = 1.96$ El IC pedido es:

$$\boxed{\overline{X} \pm \frac{\sigma}{\sqrt{n}} Z_{\alpha/2} = 1014 \pm \frac{25}{\sqrt{20}} 1.96 = [1003.04, 1024.96]}$$

b) $P(L \leq \mu) = 0.95$ Para este caso que se desea un intervalo de una sola cola se tiene que $z_{\alpha} = 1.645$

Por lo tanto $L = \overline{X} - \frac{\sigma}{\sqrt{n}} Z_{\alpha} = 1014 - \frac{25}{\sqrt{20}} 1.645 = [1004.804, \infty)$

c) $1-\alpha = 0.95$ y $z_{\alpha/2} = 1.96$.

Si el ancho del intervalo es de 8 hrs entonces la mitad del intervalo es de 4 hrs, es decir $\frac{\sigma}{\sqrt{n}} Z_{\alpha/2} = 4$,

despejando el tamaño de la muestra se tiene que $n = \left(\frac{\sigma}{4} Z_{\alpha/2}\right)^2 = \left[\frac{25}{4}(1.96)\right]^2 = 150$ bombillas.

Ejemplo 6: Un experto en eficiencia desea determinar el tiempo promedio que le toma hacer tres agujeros en una abrazadera metálica.
¿De qué tamaño se necesita la muestra para tener una confianza de 95% de que la media de la muestra está dentro de 15 seg, respecto de la media verdadera? Supón que por estudios previos se sabe que $\sigma = 40$ segundos.

<u>Solución</u>

$\alpha = 5\%$
$\varepsilon = 15$ seg
$\sigma = 40$ seg
$z_{\alpha/2} = 1.96$

$\varepsilon = \left|\mu - \overline{X}\right| = 15$ seg

$$n = \left[\frac{\sigma Z_{\alpha/2}}{\varepsilon}\right]^2 = \left[\frac{40(1.96)}{15}\right]^2 = 27 \text{ abrazaderas.}$$

Ejemplo 7: La lectura de pantallas digitales bajo luz brillante es problemática. Un grupo de ingenieros pretende diseñar un filtro para optimizar la brillantez y el contraste de color. A tal efecto, se plantea estimar el número promedio de pies-candela en la cabina de mando de aviones comerciales, donde se usará el filtro. Se realiza un estudio piloto preliminar y se obtiene una desviación estándar estimada de 500 pies-candela. ¿Cuán grande debe ser la muestra para estimar la media poblacional a no más de 50 pies-candela con una confianza del 95%?
Respuesta n=385

CASO 2._ I.C para una media de una distribución Normal con varianza desconocida.
(Para con σ^2 desconocida)

Si $X_1,\ldots,X_n$ es una m.a de una $N(\mu,\ \sigma^2)$ con parámetros desconocidos, no conocemos la varianza poblacional por lo tanto la estimamos con la varianza muestral S^2.

Sabemos que $\overline{X} \sim f_{\overline{X}}(\mu,\ \sigma^2/n) \approx f_{\overline{X}}(\mu,\ S^2/n)$ el estadístico $t = \dfrac{\overline{X}-\mu}{S/\sqrt{n}}$ tiene una distribución t de Student con $(n-1)$ grados de libertad.

El I.C con una confianza de $(1-\alpha)100\%$ para la media poblacional se construye de la manera siguiente:

$$1-\alpha = P\left(-t_{\alpha/2} \le t \le t_{\alpha/2}\right)$$
$$= P\left(-t_{\alpha/2} \le \frac{\overline{X}-\mu}{S/\sqrt{n}} \le t_{\alpha/2}\right)$$
$$= P\left(-\frac{S}{\sqrt{n}}t_{\alpha/2} \le \overline{X}-\mu \le \frac{S}{\sqrt{n}}t_{\alpha/2}\right) \text{ Multiplicando por } -1 \text{ y acomodando}$$
$$= P\left(-\frac{S}{\sqrt{n}}t_{\alpha/2} \le \mu-\overline{X} \le \frac{S}{\sqrt{n}}t_{\alpha/2}\right) \text{ Sumando } \overline{X}$$
$$= P\left(\overline{X}-\frac{S}{\sqrt{n}}t_{\alpha/2} \le \mu \le \overline{X}+\frac{S}{\sqrt{n}}t_{\alpha/2}\right)$$

El I.C es $\boxed{\left[\overline{X}-\dfrac{S}{\sqrt{n}}t_{\alpha/2,n-1}\ ,\ \overline{X}-\dfrac{S}{\sqrt{n}}t_{\alpha/2,n-1}\right]}$

Si $n \ge 30$ este intervalo suele construirse como $\boxed{\left[\overline{X}-\dfrac{S}{\sqrt{n}}Z_{\alpha/2}\ ,\ \overline{X}-\dfrac{S}{\sqrt{n}}Z_{\alpha/2}\right]}$

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

Construye un intervalo de confianza al 98% para el verdadero promedio de cantidad de saturación de gas residual.
Respuesta: [33.53, 43.79]

Considera los siguientes 1000 intervalos de confianza al 96% 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 a μ. ¿Qué clase de variable aleatoria es Y?)
Respuesta: 950; .8714

Ejemplo 8: Los conductores metálicos o tubos huecos se usan en el cableado eléctrico. En la prueba de tubos de 1 pulgada, se obtienen los datos siguientes respecto del diámetro exterior en pulgadas del tubo:

1.281	1.293	1.287	1.286
1.288	1.293	1.291	1.295
1.292	1.291	1.290	1.296
1.289	1.289	1.286	1.291
1.291	1.288	1.289	1.286

a) Calcula la media y la varianza muestral en base a la muestra.
b) Si el muestreo se hizo de una distribución normal. Construye un IC al 95% para el diámetro externo medio de tubos de este tipo.
c) Los fabricantes de este tipo de tuvo afirman que el diámetro exterior medio es de 1.29 pulgadas. ¿Qué opinas al respecto?

CASO 3._ I.C para la diferencias de medias con varianzas conocidas.
(Para $\mu_1 - \mu_2$ con σ_1^2 y σ_2^2 conocidas)

Sea $X_1 \sim f_{X_1}(x_1;\theta)$ con $E(X_i) = \mu_1$ desconocida y $V(X_1) = \sigma_1^2$ conocida, y $X_2 \sim f_{X_2}(x_2;\theta)$ con $E(X_2) = \mu_2$ desconocida y $V(X_2) = \sigma_2^2$ conocida.

Si las densidades son Normales o bien los tamaños de las muestras aleatorias n_1 y n_2 son grandes, el TCL garantiza que $\overline{X}_1 \sim N\left(\mu_1, \dfrac{\sigma_1^2}{n_1}\right)$ y $\overline{X}_2 \sim N\left(\mu_2, \dfrac{\sigma_2^2}{n_2}\right)$ además una diferencia de Normales es Normal con media igual a la diferencia de las medias y varianza igual a la suma de las varianzas.

$$\left(\overline{X}_1 - \overline{X}_2\right) \sim N\left(\mu_1 - \mu_2, \ \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right) \text{ por lo tanto } \quad Z = \frac{\left(\overline{X}_1 - \overline{X}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \sim N(0, 1)$$

La construcción del I.C es como sigue:

$$1 - \alpha = P\left(-Z_{\alpha/2} \le Z \le Z_{\alpha/2}\right)$$

$$= P\left(-Z_{\alpha/2} \le \frac{\left(\overline{X}_1 - \overline{X}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \le Z_{\alpha/2}\right)$$

$$= P\left(-\sqrt{\frac{\sigma_1^{\ 2}}{n_1} + \frac{\sigma_2^{\ 2}}{n_2}} Z_{\alpha/2} \le \left(\overline{X}_1 - \overline{X}_2\right) - \left(\mu_1 - \mu_2\right) \le \sqrt{\frac{\sigma_1^{\ 2}}{n_1} + \frac{\sigma_2^{\ 2}}{n_2}} Z_{\alpha/2} \right)$$ Multiplicando por -1 y

acomodando

$$= P\left(-\sqrt{\frac{\sigma_1^{\ 2}}{n_1} + \frac{\sigma_2^{\ 2}}{n_2}} Z_{\alpha/2} \le \left(\mu_1 - \mu_2\right) - \left(\overline{X}_1 - \overline{X}_2\right) \le \sqrt{\frac{\sigma_1^{\ 2}}{n_1} + \frac{\sigma_2^{\ 2}}{n_2}} Z_{\alpha/2} \right)$$ Restando $\left(\overline{X}_1 - \overline{X}_2\right)$

$$= P\left(\left(\overline{X}_1 - \overline{X}_2\right) - \sqrt{\frac{\sigma_1^{\ 2}}{n_1} + \frac{\sigma_2^{\ 2}}{n_2}} Z_{\alpha/2} \le \mu_1 - \mu_2 \le \left(\overline{X}_1 - \overline{X}_2\right) + \sqrt{\frac{\sigma_1^{\ 2}}{n_1} + \frac{\sigma_2^{\ 2}}{n_2}} Z_{\alpha/2} \right)$$

El IC para la diferencia de medias es de la forma

$$\boxed{\left[\left(\overline{X}_1 - \overline{X}_2\right) - \sqrt{\frac{\sigma_1^{\ 2}}{n_1} + \frac{\sigma_2^{\ 2}}{n_2}} Z_{\alpha/2} \ ,\ \left(\overline{X}_1 - \overline{X}_2\right) + \sqrt{\frac{\sigma_1^{\ 2}}{n_1} + \frac{\sigma_2^{\ 2}}{n_2}} Z_{\alpha/2} \right]}$$

Observación: Si $n_1 = n_2 = n$ el I.C queda

$$\left[\left(\overline{X}_1 - \overline{X}_2\right) - \sqrt{\frac{\sigma_1^{\ 2} + \sigma_2^{\ 2}}{n}} Z_{\alpha/2},\ \left(\overline{X}_1 - \overline{X}_2\right) - \sqrt{\frac{\sigma_1^{\ 2} + \sigma_2^{\ 2}}{n}} Z_{\alpha/2} \right]$$

y entonces $\varepsilon \le \frac{\sqrt{\sigma_1^{\ 2} + \sigma_2^{\ 2}}}{\sqrt{n}} Z_{\alpha/2}$, despejando al tamaño de la muestra se tiene que

$$\boxed{n = \left(\frac{Z_{\alpha/2}}{\varepsilon} \right)^2 \left(\sigma_1^{\ 2} + \sigma_2^{\ 2} \right)}$$

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 al 95% para el verdadero promedio de porosidad de cierta veta, si el promedio de porosidad para 20 especimenes de la veta fue de 4.85.

b) Calcula un IC al 98% para el verdadero promedio de porosidad de otra veta con base en 16 especimenes 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 un tamaño muestral si la longitud del intervalo de 95% debe ser 0.40?

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. Para los siguientes datos: $n_1 = 6$, $\overline{x}_1 = 115.7$, $s_1 = 5.03$, $n_2 = 6$, $\overline{x}_1 = 129.3$ y $s_2 = 5.38$. Utiliza estos datos para calcular un intervalo de confianza al 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. ¿Sugiere el intervalo que se dispone de información precisa acerca del valor de esta diferencia?

Respuesta: [-20.3, -6.9]

Un artículo en la revista Consumer Reports, de noviembre de 1983, comparó varios tipos de baterías. El promedio de duraciones de baterías AA alcalinas marca Duracell y de baterías AA alcalinas marca Eveready Energizar se dieron como 4.1 y 4.5 h, respectivamente. Supón que éste es el promedio de duraciones poblacionales. Sea $\bar{x}$ la duración promedio muestral de 100 baterías Duracell y $\bar{y}$ la duración promedio muestral de 100 baterías Eveready. ¿Cuál es el valor medio de $\bar{x}-\bar{y}$ (esto es, dónde está centrada la distribución de $\bar{x}-\bar{y}$)? Respuesta: -0.4 h;

Caso 4 I.C para $(\mu_1-\mu_2)$ con σ_1^2 y σ_2^2 desconocidas, con densidades Normales con muestras pequeñas.

$$\left.\begin{array}{l} X_1 \sim N\left(x_1;\mu_1,\ \sigma_1^2\right) \\ \\ X_2 \sim N\left(x_2;\mu_2,\ \sigma_2^2\right) \end{array}\right\} \text{con } \mu_i \text{ desconocida y } \sigma_i^2 \text{ desconocida } i=1,2.$$

Estimamos las varianzas poblacionales con las muestrales, es decir $\hat{\sigma}_1^2 = S_1^2$ y $\hat{\sigma}_2^2 = S_2^2$.

4i) Suponiendo que las varianzas son iguales $\sigma_1^2 = \sigma_2^2 = \sigma^2$.

Un estimador insesgado para la varianza común y aprovechando la información muestral es

$$S^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}$$

Pues
$$E\left(S^2\right) = \frac{(n_1-1)\sigma^2 + (n_2-1)\sigma^2}{n_1+n_2-2} = \sigma^2$$

Se sabe que $\bar{X}_1 \sim N\left(\mu_1,\ \frac{S_1^2}{n_1}\right)$ y $\bar{X}_2 \sim N\left(\mu_2,\ \frac{S_2^2}{n_2}\right)$

$$E\left(\bar{X}_1-\bar{X}_2\right) = \mu_1-\mu_2 \ \text{ y } \ V\left(\bar{X}_1-\bar{X}_2\right) = \frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}$$

Considerando varianzas iguales se tiene que $V\left(\bar{X}_1-\bar{X}_2\right) = S^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)$

Y el estadístico t se distribuye como una t de Student con (n_1+n_2-2) grados de libertad lo cual se escribe; $t=\dfrac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{S\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}\sim t_{n_1+n_2-2}$

Queremos que $P\left(-t_{\alpha/2}\le \dfrac{\overline{X}_1-\overline{X}_2-(\mu_1-\mu_2)}{S\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}\le t_{\alpha/2}\right)=1-\alpha$ entonces el I.C es de la forma

$$\boxed{\left[\overline{X}_1-\overline{X}_2-S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\,t_{\alpha/2,n_1+n_2-2}\ ,\ \overline{X}_1-\overline{X}_2-S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\,t_{\alpha/2,n_1+n_2-2}\right]}$$

4ii) Suponiendo varianzas poblacionales diferentes, $\sigma_1^{\ 2}\ne\sigma_2^{\ 2}$ entonces el estadístico

$$t^*=\frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{\sqrt{\dfrac{S_1^{\ 2}}{n_1}+\dfrac{S_2^{\ 2}}{n_2}}}\sim t_\nu$$

Donde los grados de libertad υ están dados por $\nu=\dfrac{\left(\dfrac{S_1^{\ 2}}{n_1}+\dfrac{S_2^{\ 2}}{n_2}\right)^2}{\dfrac{\left(S_1^{\ 2}/n_1\right)^2}{n_1+1}+\dfrac{\left(S_2^{\ 2}/n_2\right)^2}{n_2+1}}-2$

CASO 5 Estimación por intervalo para una proporción p .

La proporción de elementos de interés en la población es $p=\dfrac{\sum_{i=1}^{N}X_i}{N}$ con N tamaño de la población y

$$X_i=\begin{cases}1 \text{ si tienen la característica de interés}\\ 0 \text{ en otro caso}\end{cases}$$

el estimador para p es la proporción muestral $\hat{p}=\dfrac{\sum_{i=1}^{n}X_i}{n}$, observa que esta proporción muestral es una media muestral para una población Bernoulli con media p y varianza pq. El TCL garantiza que $\hat{p}\sim N\left(\mu_{\hat{p}}=\mu,\sigma_{\hat{p}}^2=\dfrac{\sigma^2}{n}\right)$ es decir $\hat{p}\sim N\left(\mu_{\hat{p}}=p,\sigma_{\hat{p}}^2=\dfrac{pq}{n}\right)$.

La construcción del Intervalo con una confianza de $(1-\alpha)100\%$ es como sigue:

$$1-\alpha = P\left(-Z_{\alpha/2} \le Z \le Z_{\alpha/2}\right)$$
$$= P\left(-Z_{\alpha/2} \le \frac{\hat{p}-\mu_{\hat{p}}}{\sigma_{\hat{p}}} \le Z_{\alpha/2}\right)$$
$$= P\left(-Z_{\alpha/2} \le \frac{\hat{p}-p}{\sqrt{\frac{pq}{n}}} \le Z_{\alpha/2}\right)$$
$$= P\left(-Z_{\alpha/2}\sqrt{\frac{pq}{n}} \le \hat{p}-p \le Z_{\alpha/2}\sqrt{\frac{pq}{n}}\right)$$
$$= P\left(\hat{p}-Z_{\alpha/2}\sqrt{\frac{pq}{n}} \le p \le \hat{p}+Z_{\alpha/2}\sqrt{\frac{pq}{n}}\right)$$

El intervalo para p con una confianza o de $(1-\alpha)100\%$ está dado por

$$\boxed{\left[\hat{p}-Z_{\alpha/2}\sqrt{\frac{pq}{n}},\, \hat{p}+Z_{\alpha/2}\sqrt{\frac{pq}{n}}\right]}$$

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. Obten un intervalo de confianza al 99% para la verdadera proporción p de compradoras adultas que solicitan vale en tales situaciones.
Respuesta: [.182, .316]

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 al 90% para la verdadera proporción de cánceres de pecho que serían detectados por la técnica GST (dado que n es pequeña, el intervalo será muy amplio).

CASO 6 Estimación por intervalo para una diferencia de proporciones $p_1 - p_2$.

Recuerda que el TCL garantiza que $\hat{p}_i \sim N\left(\mu_{\hat{p}_i} = \mu_i, \sigma^2_{\hat{p}_i} = \frac{\sigma_i^2}{n_i}\right)$ es decir

$\hat{p}_i \sim N\left(\mu_{\hat{p}_i} = p_i, \sigma^2_{\hat{p}_i} = \frac{p_i q_i}{n_i}\right)$. Para $i = 1,\ 2$.

Y que una diferencia de Normales también es Normal con media igual a la resta de las medias y varianza igual a la suma de las varianzas. Es decir;

$$\left(\hat{p}_1 - \hat{p}_2\right) \sim N\left(\mu_{\hat{p}_1-\hat{p}_2} = p_1 - p_2,\ \sigma^2_{\hat{p}_1-\hat{p}_2} = \frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}\right)$$

por lo tanto $Z = \dfrac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\dfrac{p_1q_1}{n_1} + \dfrac{p_2q_2}{n_2}}} \sim N(0,\ 1)$

$$1-\alpha = P\left(-Z_{\alpha/2} \le Z \le Z_{\alpha/2}\right)$$

$$= P\left(-Z_{\alpha/2} \le \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}}} \le Z_{\alpha/2} \right)$$

$$= P\left(-Z_{\alpha/2}\sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}} \le (\hat{p}_1 - \hat{p}_2) - (p_1 - p_2) \le Z_{\alpha/2}\sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}} \right)$$

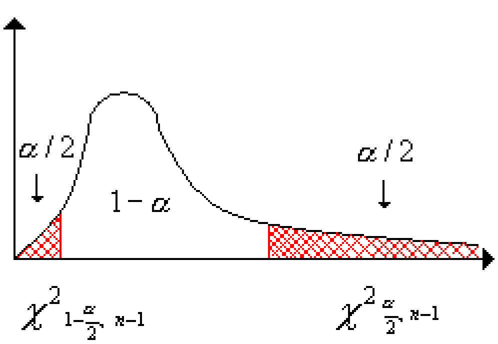
$$= P\left((\hat{p}_1 - \hat{p}_2) - Z_{\alpha/2}\sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}} \le p_1 - p_2 \le (\hat{p}_1 - \hat{p}_2) + Z_{\alpha/2}\sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}} \right)$$

El IC para la diferencia de medias es de la forma

$$\boxed{\left[(\hat{p}_1 - \hat{p}_2) - Z_{\alpha/2}\sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}}\ ,\ (\hat{p}_1 - \hat{p}_2) + Z_{\alpha/2}\sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}} \right]}$$

CASO 7 Intervalo de confianza para σ^2.

Si $X \sim N(\mu,\ \sigma^2)$ el estimador puntual para σ^2 es la varianza muestral S^2 y el estadístico $\chi^2 = \dfrac{(n-1)S^2}{\sigma^2}$ se distribuye como una ji-Cuadrada con $(n-1)$ g.l.



Procedemos como siempre, con una confianza de $(1-\alpha)100\%$

$$1-\alpha = P\left(\chi^2_{1-\frac{\alpha}{2},n-1} \le \chi^2 \le \chi^2_{\frac{\alpha}{2},n-1}\right)$$
$$= P\left(\chi^2_{1-\frac{\alpha}{2},n-1} \le \frac{n-1}{\sigma^2}S^2 \le \chi^2_{\frac{\alpha}{2},n-1}\right)$$
$$= P\left(\frac{\chi^2_{1-\frac{\alpha}{2},n-1}}{(n-1)S^2} \le \frac{1}{\sigma^2} \le \frac{\chi^2_{\frac{\alpha}{2},n-1}}{(n-1)S^2}\right)$$
$$= P\left(\frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2},n-1}} \le \sigma^2 \le \frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2},n-1}}\right)$$

El I.C al $(1-\alpha)100\%$ para σ^2 de una densidad normal es $\boxed{\left[\frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2},n-1}}\ ,\ \frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2},n-1}}\right]}$

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.

Se hicieron las siguientes observaciones de resistencia a la fractura de placas base 18% de acero mariginizado al níquel:

69.5	71.9	72.6	73.1	73.3	73.5	75.5	75.7
75.8	76.1	76.2	76.2	77.0	77.9	78.1	79.6
79.7	79.9	80.1	82.2	83.7	93.7		

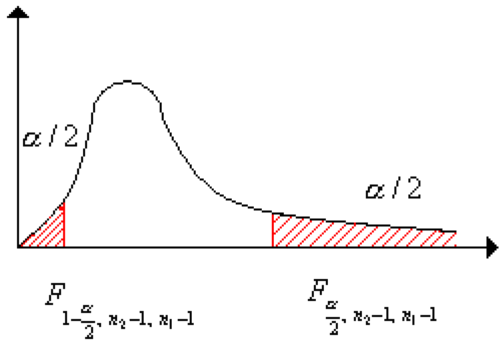
Calcula un intervalo de confianza al 99% para la desviación estándar de la distribución de la resistencia a la fractura. ¿Es valido este intervalo, cualquiera que sea la naturaleza de la distribución? Explica.
Respuesta: [3.6, 8.1]; no

CASO 8 I.C **para** $\dfrac{\sigma_1^{\,2}}{\sigma_2^{\,2}}$ **de densidades Normales.**

Si $X_1 \sim N\left(\mu_1,\ \sigma_1^{\,2}\right)$ y $X_2 \sim N\left(\mu_2,\ \sigma_2^{\,2}\right)$

Se toma una m.a de X_1 de tamaño n_1 cuya varianza muestral es $S_1^{\,2}$ y una m.a de X_2 de tamaño n_2 y la varianza muestral para ésta es $S_2^{\,2}$.

El estadístico $F = \dfrac{S_2^{\,2}/\sigma_2^{\,2}}{S_1^{\,2}/\sigma_1^{\,2}}$ se distribuye como una F con (n_2-1) grados de libertad en el numerador y (n_1-1) g.l en el denominador.



El I.C al $(1-\alpha)100\%$ se construye de la forma siguiente:

$$1-\alpha = P\left(F_{1-\frac{\alpha}{2},n_2-1,n_1-1} \le F \le F_{\frac{\alpha}{2},n_2-1,n_1-1} \right)$$

$$= P\left(F_{1-\frac{\alpha}{2},n_2-1,n_1-1} \le \frac{S_2^{\,2}/\sigma_2^{\,2}}{S_1^{\,2}/\sigma_1^{\,2}} \le F_{\frac{\alpha}{2},n_2-1,n_1-1} \right)$$

$$= P\left(\frac{S_1^{\,2}}{S_2^{\,2}} F_{1-\frac{\alpha}{2},n_2-1,n_1-1} \le \frac{\sigma_1^{\,2}}{\sigma_2^{\,2}} \le \frac{S_1^{\,2}}{S_2^{\,2}} F_{\frac{\alpha}{2},n_2-1,n_1-1} \right)$$

Finalmente el IC es $\boxed{\left[\left(\frac{S_1^{\,2}}{S_2^{\,2}}\right) F_{1-\frac{\alpha}{2},n_2-1,n_1-1} \quad , \quad \left(\frac{S_1^{\,2}}{S_2^{\,2}}\right) F_{\frac{\alpha}{2},n_2-1,n_1-1} \right]}$

Se encontró que la desviación estándar muestral de concentración de sodio en sangre entera (mEq/L) para $n_1 = 20$ anguilas marinas fue de $s_1 = 40.5$, mientras que la desviación estándar muestral de concentración para $n_2 = 20$ anguilas de agua dulce fue de $s_2 = 32.1$. Si se supone normalidad de las dos distribuciones de concentración, prueba con una confianza del 90% si la información sugiere cualquier diferencia entre varianzas de concentración para los dos tipos de anguilas.

5.5 PRUEBA DE HIPÓTESIS.

La hipótesis es un supuesto que se tiene de la realidad, basado en la teoría o en la experiencia y que se pretende demostrar mediante un experimento o una prueba usando la información proporcionada por una muestra aleatoria de la población de interés.

Los elementos que se necesitan en una prueba de hipótesis son cuatro:

Elemento 1: Plantear la hipótesis. En todo experimento existen dos hipótesis que compiten como explicación de los resultados, una es la hipótesis nula denotada por H_0 y la otra es la hipótesis alternativa, H_a. La hipótesis alternativa puede ser direccional, es decir indica el sentido de la desigualdad, en tal caso se dice que la prueba es de una cola. Si la hipótesis alternativa no indica el sentido de la desigualdad se llama no direccional y en tal caso se tendrá una prueba de dos colas.

<table>
<tr><th></th><th></th><th>Hipótesis nula
H_0</th><th>Hipótesis alternativa
H_a</th></tr>
<tr><td rowspan="2">H_a
direccional</td><td rowspan="2">Prueba de una cola</td><td>$\theta \leq \theta_0$</td><td>$\theta > \theta_0$</td></tr>
<tr><td>$\theta \geq \theta_0$</td><td>$\theta < \theta_0$</td></tr>
<tr><td>H_a no
direccional</td><td>Prueba de dos colas</td><td>$\theta = \theta_0$</td><td>$\theta \neq \theta_0$</td></tr>
</table>

La hipótesis nula es la contraparte lógica de la hipótesis alternativa, de tal manera que si la una es falsa, la otra es verdadera. Las hipótesis deben ser mutuamente excluyentes y exhaustivas.

La manera de escribir cada uno de los contrastes de hipótesis del cuadro anterior es como se muestra a continuación

Prueba de cola superior.

$$H_0 : \theta \leq \theta_0$$
$$H_a : \theta > \theta_0$$

Prueba de cola inferior.

$$H_0 : \theta \geq \theta_0$$
$$H_a : \theta < \theta_0$$

Prueba de dos colas.

$$H_0 : \theta = \theta_0$$
$$H_a : \theta \neq \theta_0$$

Elemento 2: Seleccionar un estadístico de prueba ε_p. Un estadístico es una función de las mediciones muestrales, que sirve para tomar la decisión estadística de rechazar o no la hipótesis nula ya que la distribución del estadístico de prueba se encuentra suponiendo que la hipótesis nula es verdadera.

La distribución del estadístico de prueba depende del parámetro sobre el cual se hace la prueba de hipótesis, si la prueba es sobre una o dos medias con muestras grandes o una o dos proporciones poblacionales el ε_p tiene una distribución normal estándar $N(0,1)$, si la muestra es chica y la varianza poblacional desconocida el estadístico de prueba tiene una distribución t de Student, si queremos hacer una prueba de hipótesis para una varianza la distribución Ji-cuadrada es la indicada y para el caso dos varianzas poblacionales el ε_p se distribuye como un F. La probabilidad de la colita o colitas delimitadas por el ε_p se conoce como $p-value$.

En el cuadro siguiente se presenta el valor del $p-value$ para diferentes estadísticos de prueba con diferente nivel de significancia.

Prueba	ε_p	$p-value$
De cola superior para una media poblacional con una m.a de tamaño 87.	$Z = 2.46$	0.0069
De cola inferior para una diferencia de medias con varianzas poblacionales desconocidas pero iguales en base a dos muestras aleatorias de tamaños 20 y 25.	$t = -1.83$	0.037
De dos colas para una varianza poblacional, considerando una m.a de 80 casos.	$\chi^2 = 102.67$	0.038
De cola superior para una comparación de dos varianzas poblacionales con muestras aleatorias de 100 y 125 casos.	$F = 1.47$	0.021

Elemento 3: Encontrar el punto crítico P_c definido por el nivel de significancia α de la prueba. El punto crítico es el que marca a partir de donde se considera que los valores del ε_p están lo suficientemente lejos para pensar que hay muy poca probabilidad de que la hipótesis nula sea verdadera.

Los valores extremos de la distribución delimitados por el P_c (cola superior, cola inferior o dos colas) forman lo que se llama la Región de Rechazo (RR).

El nivel α de la prueba es un valor pequeño dado por el investigador, y en una prueba lateral el punto crítico es aquel valor que deja una probabilidad igual a α en la cola superior o en la cola inferior de la distribución según sea el caso. Si la prueba es bilateral entonces el P_c está formado por dos valores, uno en cada extremo de la distribución y en cada una de las colas hay una probabilidad de $\alpha/2$.

Ejemplo 1: Para una prueba de hipótesis de cola superior de una media poblacional con varianza conocida se tienen que el $\varepsilon_p \sim N(0,1)$. Si se sabe que el investigador desea realizar la prueba con un nivel $\alpha = 0.05$ se tiene que $P_c = Z_\alpha = Z_{0.05} = 1.65$.

Ejemplo 2: En caso de una prueba de cola inferior sobre una media poblacional cuando el tamaño de muestra es chico, digamos $n = 20$ y la varianza poblacional desconocida, se tiene que el estadístico de prueba se distribuye como una t de Student y por lo tanto el punto crítico para una significancia de $\alpha = 0.03$ es $-t_{\alpha,n-1} = -t_{0.03,19} = -2.35$.

Ejemplo 3: Para una prueba de dos colas sobre la varianza σ^2, con un nivel de significancia $\alpha = 0.05$ se tiene que el estadístico de prueba tiene una distribución ji-cuadrada, por lo tanto el punto crítico se forma de dos partes $\chi^2_{1-\alpha/2,n-1} = \chi^2_{0.975,n-1}$ y $\chi^2_{\alpha/2,n-1} = \chi^2_{0.025,n-1}$, observa que estos dos valores que forman el P_c delimitan las dos colas extremas de la distribución ji-cuadrada.
Si el tamaño de la muestra es $n=65$ el punto crítico se forma por los números $\chi^2_{0.975,64} = 43.78$ y $\chi^2_{0.025,n-1} = 88$.

Ejemplo 4: En el caso de una prueba de cola inferior para probar una hipótesis sobre dos varianzas poblacionales, se tiene que el estadístico de prueba se distribuye como una F.
Si la significancia de la prueba es $\alpha = 0.01$ en base a dos muestras aleatorias de tamaños 40 y 50 se tiene que el punto crítico es $F_{0.99,39,49} = 0.48$.

Elemento 4: Establecer la regla de decisión. Hay dos maneras de tomar la decisión, la primera es comparando el estadístico de prueba con el punto crítico y la segunda es comparando el $p-value$ con el nivel de significancia de la prueba α.

***Primera*:** Si el estadístico de prueba cae en la región de rechazo, es decir si $\varepsilon_p > P_c$ para una prueba de cola superior, $\varepsilon_p < P_c$ para una prueba de cola inferior o $\varepsilon_p > P_c$ o $\varepsilon_p < P_c$ para una prueba de dos colas entonces se rechaza la hipótesis nula, de lo contrario se acepta.

***Segunda*:** Si $p-value \le \alpha$ se rechaza H_0
Si $p-value > \alpha$ no se rechaza H_0.

Nota._ Generalmente la hipótesis alternativa es la que se desea apoyar en base a la información contenida en la muestra.

Recuerda que si tuviéramos certeza de las cosas la estadística no tendría razón de ser, debemos apoyarnos en probabilidades y por lo mismo nos podemos equivocar, pero se busca que esto ocurra el menor número de veces. Siempre que se hace una prueba de hipótesis puede ocurrir que nuestra decisión sea correcta pero también puede darse el caso de tomar una decisión incorrecta como se muestra en el cuadro siguiente.

	Estado Real	
Decisión	**H_0 es verdadera**	**H_0 es falsa**
Rechazar H_0	**Error tipo I**	✓ Decisión correcta
No rechazar H_0	✓ Decisión correcta	**Error tipo II**

Las probabilidades asociadas a estos errores son:

$$\alpha = P(\text{Error tipo I}) = P(\text{Rechazar } H_o \mid H_o \text{ es verdadera})$$

$$\beta = P(\text{Error tipo II}) = P(\text{No rechazar } H_o \mid H_0 \text{ es falsa})$$

La ***confianza*** es la probabilidad de no rechazar la hipótesis nula dado que es verdadera, es decir

$$1-\alpha = P(\text{No rechazar } H_o \mid H_o \text{ es verdadera})$$

$$\boxed{\text{Confianza de la prueba} = 1-\alpha}$$

y la ***potencia de la prueba*** es la probabilidad de rechazar la hipótesis nula dado que es falsa.

$$\text{Potencia de la prueba} = P(\text{rechazar } H_o \mid H_0 \text{ es verdadera})$$

$$\boxed{\text{Potencia de la prueba} = 1-\beta}$$

Observa que si DC es el evento de tomar una decisión correcta y ε es el evento de cometer un error, cuando se hace una prueba de hipótesis se tienen que $P(DC)+P(\varepsilon)=1$ y como el error puede ser de dos tipos se tiene que $P(\varepsilon)=\alpha+\beta$, no se sabe cuanto vale la probabilidad de cometer un error pero una vez que ésta es fija se observa que a medida que α crece β disminuye y cuando β crece α decrece, es decir existe una relación inversa entre la probabilidad de cometer un error tipo I y la probabilidad de cometer un error tipo II.

El nivel α de la prueba lo establece el investigador al principio del experimento, es el nivel al cual desea limitar la probabilidad de cometer un error de tipo I, su valor depende de la naturaleza del experimento, generalmente se usan valores muy pequeños (confianzas grandes 95% o más) cuando el estudio es de mucha precisión o está relacionado con cuestiones de salud, y toma valores no tan pequeños cuando se está en la etapa exploratoria del experimento.

Ejemplo 5: Se enseño a dos grupos de niños de la escuela primaria a leer por dos métodos diferentes, 50 por cada método. Al terminar el periodo de instrucción, una prueba de lectura dio los siguientes resultados: $\bar{x}_1 = 74, \bar{x}_2 = 71, s_1 = 9, s_2 = 10$. ¿Cuál es el nivel de significancia de la prueba si deseas verificar si hay evidencia de una diferencia real entre las dos medias poblacionales? ¿Cuál sería tu conclusión si deseas utilizar un valor de α igual a 0.05?

Solución

X_1 : Calificación en la lectura de los niños bajo el método 1.
X_2 : Calificación en la lectura de los niños bajo el método 2.

$n_1 = n_2 = 50$
$\bar{x}_1 = 74, \bar{x}_2 = 71$
$s_1 = 9, s_2 = 10$

Contraste de hipótesis: $H_o : \mu_1 = \mu_2$
$H_a : \mu \neq \mu_2$

Estadístico de prueba: $Z = \dfrac{\overline{X}_1 - \overline{X}_2}{\sqrt{\dfrac{S_1^2 + S_2^2}{n}}} = \dfrac{74-71}{\sqrt{\dfrac{81+100}{50}}} = 1.576 \approx 1.58$

Región de rechazo (RR): Rechazar H_0 si $|Z| > Z_{\alpha/2}$ o equivalentemente si $p < \dfrac{\alpha}{2}$

El nivel de confianza es $(1-\alpha)100\% = 95\%$ y $z_{\alpha/2} = 1.96$, como $1.58 < 1.96$ no se rechaza la hipótesis nula, equivalentemente para $z = 1.58$ se tiene que $p = 0.0571 > 0.025$ por lo tanto no se rechaza H_0.

Se concluye que en base a la muestra no existe evidencia de que exista una diferencia real entre los dos métodos de enseñanza para la lectura en los niños de esa primaria.

Ejemplo 6: En un estudio para investigar los efectos del empleo de una joven modelo en la publicidad para automóviles, se mostraron a 100 hombres las fotografías de dos automóviles de igual valor y tamaño, pero de diferente tipo. Uno de los coches se mostró a 50 individuos (grupo A) con una modelo (el otro sin modelo) y a los otros 50 individuos (grupo B) se presentaron ambos automóviles sin la modelo. En el grupo A, 37 individuos consideraron más caro el automóvil mostrado con la modelo, mientras que en el grupo B, 23 individuos consideraron el mismo automóvil más costoso. ¿Indican estos resultados que el emplear una modelo aumenta el costo estimado de un automóvil? Encuentra el valor p asociado e indica tu conclusión para una prueba con un nivel de significancia igual a 0.05.

Solución

X_A : Número de hombres que consideran más caro el coche más caro (con la modelo).
X_B : Número de hombres que consideran más caro el coche más caro (sin la modelo).

$n_A = n_B = 50$

$\hat{p}_A = \frac{37}{50} = 0.74, \ \hat{p}_B = \frac{23}{50} = 0.46$

Contraste de hipótesis: $H_o : p_A \le p_B$

$H_a : p_A > p_B$

Estadístico de prueba: $Z = \dfrac{\hat{p}_A - \hat{p}_B}{\sqrt{\dfrac{\hat{p}_A\hat{q}_A + \hat{p}_B\hat{q}_B}{n}}} = \dfrac{0.74 - 0.46}{\sqrt{\dfrac{0.74(0.26) + 0.46(0.54)}{50}}} = 2.98$

Región de rechazo: Rechazar H_0 si $Z > Z_\alpha$ o equivalentemente si $p < \alpha$

El nivel de confianza es $(1-\alpha)100\% = 95\%$ y $z_\alpha = 1.645$, como $2.98 > 1.645$ se rechaza la hipótesis nula, equivalentemente para $z = 2.98$ se tiene que $p = 0.0014 < 0.05$ por lo tanto se rechaza H_0.

Se concluye que en base a la muestra existe evidencia de que la modelo sí influye en la percepción de los hombres respecto al precio del coche.

Nota._ Cuando se rechaza la hipótesis nula se dice que la prueba es significativa, esto se debe a que la igualdad siempre va en H_0 y la desigualdad en H_a. Por ejemplo, si hacemos una prueba de hipótesis para comparar dos medias poblacionales μ_1 y μ_2 y las medias muestrales son $\bar{x}_1 = 10$ y $\bar{x}_2 = 13$, numéricamente es cierto que 10 es menor que 13 pero lo que debes tener presente es que estos valores son información de muestras aleatorias de las poblaciones de interés, pudo haber ocurrido que para el mismo experimento se tuviera que $\bar{x}_1 = 13$ y $\bar{x}_2 = 10$, o incluso $\bar{x}_1 = 10$ y $\bar{x}_2 = 10$, si la conclusión sobre las medias poblacionales se hiciera en base a estas tres situaciones muestrales tendríamos que $\mu_1 < \mu_2$ en el primer caso, $\mu_1 > \mu_2$ en el segundo caso y $\mu_1 = \mu_2$ en el tercer caso, es decir estaríamos concluyendo tres cosas diferentes para la mismas poblaciones. Entonces la pregunta es ¿estos 3 puntos de diferencia entre el 10 y el 13 son realmente una diferencia significativa? ¿Se puede considerar que las medias poblacionales μ_1 y μ_2 son realmente diferentes? ¿O en realidad esta diferencia no es significativa y se debe sólo a un error muestral y en realidad las medias poblacionales se pueden considerar iguales?

5.7 Prueba de hipótesis para la media.

Caso 1._ Muestras grandes.

Una muestra grande de una distribución Normal con varianza conocida o desconocida.

Supóngase una población de una distribución Normal con media μ desconocida y varianza σ^2. En base a una muestra aleatoria suficientemente grande de tamaño n y media muestral $\bar{x}$ se puede hallar ausencia o presencia de evidencia estadística para apoyar una hipótesis que se tiene sobre el posible valor del parámetro desconocido.

Los contrastes de hipótesis que se pueden presentar son :

Contraste *i*	Contraste *ii*	Contraste *iii*	
$H_0 : \mu = \mu_0$	$H_0 : \mu = \mu_0$	$H_0 : \mu = \mu_0$	
$H_a : \mu > \mu_0$	$H_a : \mu < \mu_0$	$H_a : \mu \neq \mu_0$	(1)

Dado que la muestra es grande y proviene de una distribución Normal, se tiene que la media muestral se distribuye como una Normal con $\bar{x} \sim N\left(\mu_{\bar{x}} = \mu, \sigma^2_{\bar{x}} = \sigma^2 / n\right)$ y el estadístico de la prueba será $Z = \dfrac{\bar{x} - \mu_0}{\sigma_{\bar{Y}}} = \dfrac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$.

En caso de que la varianza población sea desconocida se estima con la varianza poblacional y dado que la muestra es grande se tiene que el estadístico $Z = \dfrac{\bar{x} - \mu_0}{S / \sqrt{n}}$ también se distribuye como una Normal estándar.

Si la confianza de la prueba de hipótesis es de $(1-\alpha)100\%$ entonces se define la **región de rechazo** como los valores de Z que están lo suficientemente lejos de μ_0 como para hacernos pensar que en realidad la media de la población es diferente a μ_0, el valor que se toma como límite de esta región de rechazo se conoce como valor crítico y se denota Z_α, $Z_{1-\alpha}$ y $Z_{\alpha/2}$ respectivamente para los contrastes de hipótesis (1). El valor Z_α es el punto de la distribución Normal estándar tal que deja a la derecha una probabilidad de α.

Una vez que se tiene el estadístico de la prueba y el valor crítico se procede a tomar la decisión de rechazar o no la hipótesis nula H_0 lo cual sea hará de acuerdo a los siguientes criterios para cada uno de los tres contrastes presentados en (1):

i) Si $Z > Z_\alpha$ rechazar H_0 a favor de H_a.
ii) Si $Z < Z_\alpha$ rechazar H_0 a favor de H_a.
iii) Si $Z < Z_{1-\alpha/2}$ o $Z > Z_{\alpha/2}$ rechazar H_0 a favor de H_a.

Ejemplo 7: El voltaje de salida en cierto circuito eléctrico debe ser iguala 130 según se especifica. Una muestra de 40 lecturas independientes para este circuito dio una media muestral de 128.6 y una desviación estándar de 2.1. Prueba la hipótesis de que el voltaje de salida promedio es 130 frente a la hipótesis de que alternativa de que es menor que 130. Utiliza un nivel de significancia de 5%.

<u>Solución</u>

Datos $n = 40$, $\bar{x} = 128.6$, $S^2 = (2.1)^2$

Contraste de hipótesis: $H_0 = 130$

$H_a < 130$

Estadístico de prueba: $Z = \dfrac{\bar{x} - \mu_0}{S/\sqrt{n}} = \dfrac{128.6 - 130}{2.1/\sqrt{40}} = -4.22$

Decisión: El valor critico es $Z_{0.05} = -1{,}645$ Debido a que el estadístico de la prueba si cae en la región de rechazo $(-4.22 < -1.645)$ se concluye que sí hay evidencia estadística de que el voltaje promedio esta por debajo de lo especificado.

CASO2._ MUESTRAS CHICHAS.
Una muestra pequeña de una distribución Normal, con varianza desconocida.

En este caso existe menos información así que este hecho se debe ver reflejado en alguna parte, cuando se calcula el intervalo de confianza se ve muy claro que queda más ancho que el intervalo correspondiente a una muestra de una población con varianza desconocida que el correspondiente a una muestra de una población con varianza conocida, y en el caso de las pruebas de hipótesis la región de rechazo será más angosta para el caso de la varianza desconocida que para el caso de la varianza conocida.

Esto ocurre porque el estadístico de prueba $t = \dfrac{\bar{x} - \mu_0}{S/\sqrt{n}}$ se distribuye cono una t de Student con $(n-1)$ grados de libertad. Y los criterios de rechazo son:

i) Si $t > t_\alpha$ rechazar H_0 a favor de H_a.
ii) Si $t < t_\alpha$ rechazar H_0 a favor de H_a.
iii) Si $t < t_{1-\alpha/2}$ o $t > t_{\alpha/2}$ rechazar H_0 a favor de H_a.

PRUEBA t PARA UNA MUESTRA.

Cuando se conoce la distribución de los datos así como sus parámetros, el estadístico de prueba adecuado es el Z que representa una distribución Normal, esta situación es casi imposible en la vida real y por lo general se desconoce el valor de los parámetros de la población de donde ha salido la muestra, el estadístico correcto es el t que corresponde a una distribución t de Student, la cual también es una distribución simétrica muy parecida a la Normal pero con colas más pesadas, lo cual tiene como consecuencia producir intervalos de confianza más amplios, cosa que resulta muy lógica ya que es el precio que se tiene que pagar por la falta de información respecto a los parámetros. La prueba de hipótesis que involucra al estadístico de prueba t, comúnmente se conoce como prueba t.

Ejemplo 8: Se analizan 26 observaciones de tiempo de escape, en segundos, para trabajadores petroleros, en un ejercicio simulado, a partir de las cuales se calculan el promedio y la desviación estándar muestrales, que fueron 370.69 y 24.26, respectivamente. Supón que los investigadores habían creído, a priori, que el tiempo real promedio de escape era de 6 minutos, a lo sumo. ¿Los datos contradicen esta creencia a priori? Asumiendo normalidad, prueba las hipótesis apropiadas con un nivel de significancia de 0.05.

<u>**Solución**</u>

$n = 26$
$\bar{x} = 370.69 seg = 6.18 \min$
$s = 24.26 seg = 0.40 \min$
$\alpha = 0.05$

Contraste de hipótesis: $H_0 : \mu \leq 6$
$H_a : \mu > 6$

Estadístico de prueba: $t = \dfrac{\bar{x} - \mu_0}{\dfrac{s}{\sqrt{n}}} = \dfrac{\sqrt{n}\left(\bar{x} - \mu_0\right)}{s} = \dfrac{\sqrt{26}\left(6.18 - 6\right)}{0.4} = 2.25$

$p - value$: 0.0167

Punto crítico: $t_{\alpha, n-1} = t_{0.05, 25} = 2.059$

Decisión: $2.25 > 2.059$ es decir el estadístico de prueba cae en la región de rechazo entonces se rechaza H_0. Equivalentemente 0.0167<0.05 por lo tanto se concluye que la información de la muestra sí contradicen la creencia a priori de que el tiempo real promedio de escape era a lo más de 6 minutos.

Prueba de hipótesis para la diferencia de media.
Muestras grandes.
Dos muestras grandes independientes de distribuciones Normales con varianzas conocidas o desconocidas.

Los contrastes de hipótesis que se pueden hacer en estos casos son:

Contraste *i*	Contraste *ii*	Contraste *iii*	
$H_0 : \mu_1 = \mu_2$	$H_0 : \mu_1 = \mu_2$	$H_0 : \mu_1 = \mu_2$	
$H_a : \mu_1 > \mu_2$	$H_a : \mu_1 < \mu_2$	$H_a : \mu_1 \neq \mu_2$	(1)

Observa que la expresión $\mu_1 = \mu_2$ es equivalente a la expresión $\mu_1 - \mu_2 = 0$ y la expresión $\mu_1 > \mu_2$ es equivalente a $\mu_1 - \mu_2 > 0$ y la expresión $\mu_1 < \mu_2$ es equivalente a $\mu_1 - \mu_2 < 0$ y la expresión $\mu_1 \neq \mu_2$ es equivalente a la expresión $\mu_1 - \mu_2 \neq 0$.

Sean μ_1 y σ_1^2 la media y la varianza de la población 1, y μ_2 y σ_2^2 la media y la varianza de la población 2, y sean $\bar{x}_1$ y $\bar{x}_2$ las medias muestrales de tales poblaciones. El estadístico de prueba es de la forma

$Z = \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$ en el caso de varianzas poblacionales conocidas y $Z = \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}}$ para el caso de varianzas poblacionales desconocidas.

PRUEBA *T* PARA MUESTRAS INDEPENDIENTES.

Se compara las medias de dos grupos de casos, idealmente los sujetos deben asignarse aleatoriamente a dos grupos, de forma que cualquier diferencia en la respuesta sea debida al tratamiento (o falta de tratamiento) y no a otros factores.
Si quisiéramos comparar por ejemplo el ingreso medio para hombres y mujeres, en sentido estricto no podríamos aplicar esta prueba ya que el sexo no se asigna aleatoriamente. En este caso primero se establecen los grupos, uno de mujeres y otro de hombres y luego se comparan los ingresos medios.
Lo ideal es tomar personas y asignarlas aleatoriamente, a uno de dos grupos y después comparar el ingreso medio. Sin embargo a menudo se asume independencia y se hace la comparación entre hombres y mujeres.

El estadístico de prueba depende de si las varianzas de los grupos pueden o no considerarse iguales. Por lo tanto es necesario como primer paso hacer una prueba de hipótesis para verificar igualdad de varianzas.

Ejemplo 9: Dos secciones de un curso de estadística son sometidas a un mismo examen final. De las calificaciones obtenidas se extrae una muestra aleatoria de tamaño 9 en el grupo "A", y otra de tamaño 4 en el grupo "B".

Grupo "A":	65, 68, 72, 75, 82, 85, 87, 91, 95
Grupo "B":	50, 59, 71, 80

Con un nivel de significancia de 0.05 ¿podría decirse que los dos grupos tienen las mismas calificaciones promedio? Supón que provienen de poblaciones normales con varianzas iguales.

<u>Solución</u>

Contraste de hipótesis: $H_0 : \mu_A = \mu_B$
$H_a : \mu_A \neq \mu_B$

Estadístico de prueba:

$$s_p = \sqrt{\frac{(n_A-1)S_A^2+(n_B-1)S_B^2}{n_A+n_B-2}} = \sqrt{\frac{(9-1)110.25+(4-1)174}{9+4-2}} = \sqrt{\frac{(8)110.25+(3)174}{11}} = \sqrt{124.64} = 11.3$$

$$t = \frac{\bar{x}_A-\bar{x}_B}{s_p\sqrt{\frac{1}{n_A}+\frac{1}{n_B}}} = \frac{80-65}{11.3\sqrt{\frac{1}{9}+\frac{1}{4}}} = 2.21$$

Punto crítico: $t_{\alpha/2, n_1+n_2-2} = t_{0.025,11} = 2.59$

Criterio de decisión: $2.21 < 2.59$ entonces el ε_p no está en la región de rechazo.

Con una confianza del 95% sí puede decirse que los dos grupos tienen la misma calificación promedio.

5.8 PRUEBA DE HIPÓTESIS PARA LA VARIANZA.

Las pruebas de hipótesis también son de utilidad cuando queremos averiguar si existen elementos estadísticos para rechazar o no la hipótesis de que la varianza de una población es diferente a un valor en particular llamado σ_0^2, o bien para comparar las varianzas de dos poblaciones.

Se supone que las muestras aleatorias con las que trabajaremos provienen de poblaciones con distribución Normal, con medias μ_1 y μ_2 y varianzas σ_1^2 y σ_2^2 respectivamente.

En este caso la hipótesis nula es $H_0 : \sigma^2 = \sigma_0^2$, y la hipótesis alternativa puede ser cualquiera de las tres siguientes a) $H_a : \sigma^2 > \sigma_0^2$, b) $H_a : \sigma^2 < \sigma_0^2$ o c) $H_a : \sigma^2 \neq \sigma_0^2$.

El estadístico de la prueba es $\chi^2 = \frac{(n-1)S^2}{\sigma_0^2}$, el cual se distribuye como una ji-cuadrada con $(n-1)$ grados de libertad, y se rechazara la hipótesis nula si a) $\chi^2 > \chi_\alpha^2$, b) $\chi^2 < \chi_{1-\alpha}^2$ o c) $\chi^2 > \chi_{\alpha/2}^2$ o $\chi^2 < \chi_{1-\alpha/2}^2$, respectivamente para cada una de las opciones de la hipótesis alternativas del párrafo anterior.

PRUEBA DE HIPÓTESIS PARA EL COCIENTE DE DOS VARIANZAS.

Para comparar dos varianzas poblaciones es necesario tomar muestras independientes de cada una de las poblaciones y usar el cociente de las varianzas ya que si este cociente es menor que uno entonces la varianza en el denominador es mayor que la del denominador y si el cociente es mayor que uno entonces la varianza del numerador es mayor que la del denominador.

La hipótesis nula es de la forma $H_0 : \sigma_1^2 = \sigma_2^2$ y la hipótesis alternativa puede ser cualquiera de las siguientes posibilidades: a) $H_a : \sigma_1^2 > \sigma_2^2$, b) $H_a : \sigma_1^2 < \sigma_2^2$, c) $H_a : \sigma_1^2 \neq \sigma_2^2$.

Bajo la hipótesis nula el estadístico de la prueba $F = \dfrac{S_1^2}{S_2^2}$ se distribuye como una F_{n_1-1,n_2-1} con $(n_1 - 1)$ grados de libertad en el numerador y $(n_2 - 1)$ grados de libertad en el denominador.
Se rechaza la hipótesis nula si a) $F > F_{\alpha,n_1-1,n_2-1}$ con α la cola derecha de la distribución F, es decir $P(F > F_{\alpha,n_1-1,n_2-1}) = \alpha$ siendo la confianza de la prueba $(1-\alpha)100\%$, b) $F < F_{\alpha,n_1-1,n_2-1}$ o c) Para la hipótesis alternativa bilateral $F > F_{\alpha/2,n_1-1,n_2-1}$ o $F < F_{1-\alpha/2,n_1-1,n_2-1}$.

Ejemplo 10: Un fabricante de máquinas para empacar detergente afirma que su máquina podría llenar con un peso dado las cajas con un rango no mayor de 0.4 onzas. La media y la varianza de una muestra de ocho cajas de 3 libras fueron iguales a 3.1 y 0.018, respectivamente. Prueba la hipótesis de que la varianza de la población de las mediciones de los pesos es $\sigma^2 = 0.01$ frente a la alternativa $\sigma^2 > 0.01$. Utiliza una significancia de $\alpha = 0.05$.

<u>**Solución**</u>

Los datos del problema son $n = 8$, $\bar{x} = 3.1$, $S^2 = 0.018$, $\alpha = 0.05$

Contraste de hipótesis: $H_0 : \sigma^2 = 0.01$

$H_a : \sigma^2 > 0.01$

Estadístico de la prueba: $\chi^2 = \dfrac{(n-1)S^2}{\sigma_0^2} = \dfrac{(8-1)0.018}{0.01} = 12.6$

Toma de decisión: Rechazar la hipótesis nula si $\chi^2 > \chi_\alpha^2$ buscando el valor en tablas se tiene que $\chi_{0.05,7}^2 = 14.0671$ con 0.05 la significancia de la prueba y 7 son los grados de libertad $(n-1)$.

Pero 12.6<14.0671 por lo tanto se concluye que **no** existe evidencia estadística para **rechazar la hipótesis nula**, entonces no se rechaza el hecho de que la varianza del peso de las cajas sea 0.01 libras.

Ejemplo 11: Se registraron los precios, al cierre de las operaciones, de dos acciones comunes durante un periodo de 16 días. Las medias y varianzas fueron $\bar{x}_1 = 40.33$, $S_1^2 = 1.54$ y $\bar{x}_2 = 42.54$, $S_2^2 = 2.96$ respectivamente. ¿Presentan estos datos suficiente evidencia para indicar una diferencia en variabilidad para los precios al cierre de las operaciones, de las dos acciones para las poblaciones asociadas con las dos muestras? Considera $\alpha = 0.02$.

<u>**Solución**</u>

Datos del problema $n_1 = n_2 = 16$, $\overline{x}_1 = 40.33$, $S_1^2 = 1.54$, $\overline{x}_2 = 42.54$, $S_2^2 = 2.96$

Contraste de hipótesis: $H_0 : \sigma_1^2 = \sigma_2^2$

$H_a : \sigma_1^2 \neq \sigma_2^2$

Estadístico de la prueba: $F = \dfrac{S_1^2}{S_2^2} = \dfrac{1.54}{2.96} = 0{,}52$

Toma de decisión: Rechazar la hipótesis nula si $F > F_{\alpha/2, n_1-1, n_2-1}$ o $F < F_{1-\alpha/2, n_1-1, n_2-1}$.

Los grados de libertad son 15 tanto para el numerador como para el denominador y $F_{0.01,15,15} = 3.52$, $F_{0.01,15,15} = 0.280$ por lo tanto **No se rechaza la hipótesis nula** es decir no existe evidencia estadística de que las varianzas sean distintas.

PRUEBA DE HIPÓTESIS SOBRE UNA PROPORCIÓN DE POBLACIONES NORMALES.

Observa que la proporción es también una media pero de variables bernoulli. Si $\hat{p}$ es la proporción muestral, $\hat{q} = 1 - \hat{p}$ y n el tamaño de la muestra el estadístico de prueba es

$Z = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}}$ en el caso de una muestra.

CASO DE DOS PROPORCIONES.

Y para el caso de dos muestras independientes el estadístico de prueba es

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\dfrac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \dfrac{\hat{p}_2(1-\hat{p}_2)}{n_2}}}$$