

CAPITULO VII
VARIABLE ALEATORIA

Definición. Una variable aleatoria es una función X que asigna un número real a cada elemento de un espacio muestral " Ω ".

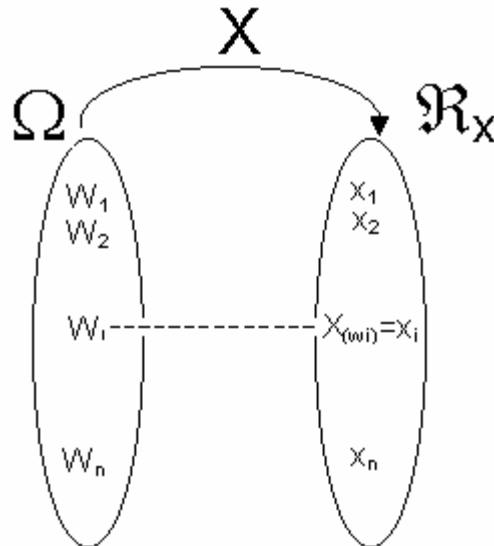
Las variables aleatorias pueden ser discretas ó continuas.

VARIABLE ALEATORIA DISCRETA

Definición. Una variable aleatoria es discreta si el número de valores posibles de la variable es un número finito ó infinito numerable.

El dominio de la variable aleatoria X es el espacio muestral Ω y el rango es un subconjunto de números reales denotado por \mathcal{R}_X .

En otras palabras, una variable aleatoria es discreta si su rango \mathcal{R}_X es un conjunto finito ó infinito numerable.



Ejemplo 1. Un experimento aleatorio E : Lanzamiento de tres monedas, el espacio muestral asociado a " E " es:

$$\Omega = \{(C,C,C), (C,C,S), (C,S,C), (C,S,S), (S,C,C), (S,C,S), (S,S,C), (S,S,S)\}$$

Se definen las siguientes variables aleatorias:

X : Número de caras

Y : Número de sellos menos número de caras.

w_i	$X(w_i)$	$Y(w_i)$
(C,C,C)	+3	-3
(C,C,S)	+2	-1
(C,S,C)	+2	-1

(S,C,C)	+2	-1
(C,S,S)	+1	+1
(S,C,S)	+1	+1
(S,S,C)	+1	+1
(S,S,S)	0	+3

$R_x = \{0, 1, 2, 3\}$
 $R_y = \{-3, -1, +1, +3\}$

Se observa que ambas variables aleatorias X e Y son del tipo discreto.

Ejemplo 2. En una caja existen 4 bolas negras, 3 blancas y 5 rojas. Se describe el experimento E como: Se extraen 2 bolas de dicha caja uno por una con reemplazo. Se define la V.A. X : número de bolas blancas. En el siguiente cuadro se muestra el espacio muestral y el rango de X.

w_i	(N,N)	(N,B)	(B,N)	(N,R)	(B,B)	(B,R)	(R,N)	(R,B)	(R,R)
$X(w_i)$	0	1	0	1	2	1	0	1	0

$R_x = \{0,1,2\}$

Ejemplo 3. E: Dos equipos A y B participan en un torneo que consiste en varios juegos. Las reglas del torneo indican que no hay empates y que gana el torneo el primer equipo que gana dos juegos seguidos o un total de tres.

Se define las variables aleatorias:

X : número de juegos en el torneo.
 Y : número de juegos que gana el equipo A en el torneo

Diagrama de árbol para representar los posibles resultados del experimento aleatorio "E".



W_i	$X(w_i)$	$Y(w_i)$
(A,A)	2	2
(A,B,A,A)	4	3
(A,B,A,B,A)	5	3
(A,B,A,B,B)	5	2
(A,B,B)	3	1
(B,B)	2	0
(B,A,A)	3	2
(B,A,B,A,A)	5	3
(B,A,B,A,B)	5	2
(B,A,B,B)	4	1

$$R_x = \{2,3,4,5\}$$

Según el rango de X e Y , estas variables son del tipo discreto.

FUNCION DE PROBABILIDAD

Sea X una variable aleatoria discreta con rango R_x , a esta variable se asocia una función $f(x)$:

$$f(x) = P(X=x) = \sum_{\{w_i \in \Omega / X(w_i)=x\}} P(\{w_i\})$$

La sumatoria es sobre todos los puntos muestrales (w_i) del espacio muestral Ω , tales que $X(w_i) = x$

La función $f(x)$ define la función de probabilidad o cuantía de la variable aleatoria discreta X, la cual cumple lo siguiente:

$$(i) f(x) > 0 \quad \text{para todo } x \in R_x$$

$$(ii) \sum_{x \in R_x} f(x) = \sum_{x \in R_x} P(X=x) = 1$$

$$(iii) f(x) = P(X=x) = 0, \quad \text{si } x \text{ no pertenece a } R_x$$

Ejemplo 4. En una caja existen 4 bolas negras, 3 bolas blancas y 5 rojas. Se extraen 2 bolas de dicha caja, una por una y con reemplazo. Se define la V.A.D. X: número de bolas blancas. Hallar la función de probabilidad de X.

En base al ejemplo 2, $R_x = \{0,1,2\}$

$$f(0) = P(X=0) = P(N,N) + P(N,R) + P(R,N) + P(R,R)$$

$$P(X=0) = (4/12)(4/12) + (4/12)(5/12) + (5/12)(4/12) + (5/12)(5/12)$$

$$P(X=0) = 9/16$$

$$f(1) = P(X=1) = P(N,B) + P(B,N) + P(B,R) + P(R,B)$$

$$P(X=1) = (4/12)(3/12) + (3/12)(4/12) + (3/12)(5/12) + (5/12)(3/12)$$

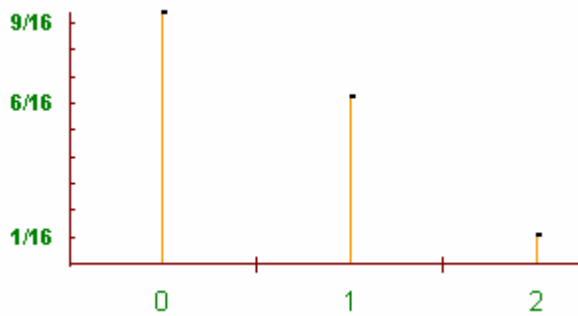
$$P(X=1) = 6/16$$

$$f(2) = P(X=2) = P(B,B) = (3/12)(3/12) = 1/16$$

Tabla de probabilidades

X	0	1	2
f(x)=P(X=x)	9/16	6/16	1/16

$f(x) = P(X=x)=0$ para todo x que no pertenece a R_x



Ejercicios :

$$(i) P(0 \leq X \leq 1) = P(X=0) + P(X=1) = 9/16 + 6/16 = 15/16$$

$$(ii) P(X > 0) = P(X=1) + P(X=2) = 6/16 + 1/16 = 7/16$$

$$(iii) P(0.5 \leq X < 1.5) = P(X=1) = 6/16$$

$$(iv) P(X > 0 | X < 1.5) = P(0 < X < 1.5) / P(X < 1.5) = P(X=1) / [P(X=0) + P(X=1)]$$

$$P(X > 0 | X < 1.5) = [6/16] / [15/16] = 6/15$$

Valor Esperado de una Variable Aleatoria Discreta X: E[X]

Sean x_1, x_2, \dots, x_k ; valores posibles de una V.A.D. X, las cuales tienen probabilidad de ocurrencia $P(X=x_1), P(X=x_2), \dots, P(X=x_k)$ respectivamente; luego el valor esperado de X ó media de X ó esperanza matemática de X está dado por:

$$E[X] = \mu = \sum_{i=1}^k x_i P(X = x_i)$$

Propiedades :

1. $E[k] = k$; k es constante
2. $E[X \pm a] = E[X] \pm a$; a es constante
3. $E[aX] = a E[X]$; a es constante
4. $E[aX + b] = a E[X] + b$; a,b constantes

TEOREMA . "Valor esperado de una función de una V.A.D. X" . Si $H(X)$ es una función de una V.A.D. x, entonces:

$$E[H(X)] = \sum_{x_i \in R_x} H(x_i) P(X=x_i)$$

Ejemplos:

$$(i) \quad E[X^2] = \sum x_i^2 P(X=x_i)$$

$$(ii) \quad E[3X^2 + 2] = \sum (3x_i^2 + 2) P(X=x_i)$$

$$(iii) \quad E[2X-6] = \sum (2x_i - 6) P(X=x_i)$$

Variancia de una Variable Aleatoria Discreta X: V(X)

X es una V.A.D. que tiene valores posibles: x_1, x_2, \dots, x_k ; con probabilidades de ocurrencia $P(X=x_1), P(X=x_2), \dots, P(X=x_k)$ respectivamente; luego la variancia de X esta dado por:

$$V(X) = \sigma^2 = E[(X-E[X])^2] = E[(X-\mu)^2]$$

$$E[(X-\mu)^2] = \sum_{i=1}^k (X-\mu)^2 P(X = x_i)$$

ó

$$V(X) = E[X^2] - \mu^2 = \sum_{i=1}^k x_i^2 P(X = x_i)$$

Propiedades

1. $V(k) = 0$
2. $V(X \pm a) = V[X]$
3. $V(aX) = a^2V(X)$
4. $V(aX + b) = a^2V(X)$

k, a, b son constantes.

Ejercicio 1. En base al ejemplo 4, hallar $E[X]$, $V(X)$, $E[3X+2]$, $V(3X+2)$.

X	0	1	2
f(x)	9/16	6/16	1/16

$$E[X] = \sum x_i P(X=x_i) = \sum x_i f(x_i) = 0 (9/16) + 1 (6/16) + 2 (1/16)$$

$$E[X] = \mu = 1/2$$

el número de bolas blancas que se esperaría extraer de esa caja que contiene 12 bolas es de 0.5 .

$$V(X) = E[X^2] - \mu^2 = \sum x_i^2 P(X=x_i) - \mu^2 = \sum x_i^2 f(x_i) - \mu^2$$

$$V(X) = 0^2(9/16) + 1^2(6/16) + 2^2(1/16) - (1/2)^2 = 3/8$$

$$E[3X + 2] = 3 E[X] + 2 = 3(1/2) + 2 = 7/2$$

$$V(3X + 2) = 3^2V(X) = 9(3/8) = 27/8$$

Ejercicio 2. Suponga que la demanda (D) diaria de un artículo es una variable aleatoria con la siguiente función de probabilidad:

$$f(d) = P(D=d) = \begin{cases} 2^d c / d! & \text{para } d=1,2,3,4 \\ 0 & \text{para otros valores} \end{cases}$$

Hallar la constante c y la demanda esperada del artículo.

Tabla de probabilidades

d_i	1	2	3	4
$P(D=d_i)$	2c	2c	8c/6	4c/6

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

$P(D=d)$ es una función de probabilidad, por lo tanto cumple:

$$\sum_{d_i \in R_D} P(D=d_i) = P(D=1) + P(D=2) + P(D=3) + P(D=4) = 1$$

$$2c + 2c + 8c/6 + 4c/6 = 1 \implies c = 1/6$$

$$P(D=1) = (2)(1/6) = 2/6$$

$$P(D=2) = (2)(1/6) = 2/6$$

$$P(D=3) = (8/6)(1/6) = 8/36$$

$$P(D=4) = (4/6)(1/6) = 4/36$$

Luego la demanda esperada del artículo $E(D)$ es:

$$E[D] = \sum_{d_i \in R_D} d_i P(D=d_i) = 1(2/6) + 2(2/6) + 3(8/36) + 4(4/36)$$

$$E[D] = 2.11 \quad ; \text{ el número de artículos que se venden en un día se espera que sea 2.11}$$

VARIABLE ALEATORIA CONTINUA

Definición. Una variable aleatoria X es continua si el número de valores posibles de la variable es un número infinito no numerable o cuando el rango de X (R_x) es un intervalo de los números reales.

Ejemplo: Sea X la variable aleatoria que representa la longitud de una varilla de acero obtenida de la línea de producción de un fabricante.

La variable X es una V.A. continua pues su rango son todos los puntos de un intervalo, como por ejemplo $R_x = [3.5 \text{ pulgadas}, 3.9 \text{ pulgadas}]$

FUNCION DE DENSIDAD DE PROBABILIDADES

Sea X una V.A. continua con rango R_x ; la función de densidad de probabilidades asociada a la variable aleatoria continua X, es una función $f(x)$ integrable que satisface lo siguiente:

$$1. f(x) \geq 0 \quad \text{para todo } X \in (-\infty, +\infty)$$

$$2. \int_{-\infty}^{+\infty} f(x) dx = 1$$

Ejemplo 1. Verifique que $f(x)$ es una función de densidad.

$$f(x) = \begin{cases} 1/2 & 0 \leq x \leq 1 \\ 3/4 - x/4 & 1 < x \leq 3 \\ 0 & \text{otros casos} \end{cases}$$

primera condición:

$$\text{si } x < 0 \quad f(x) = 0$$

$$\text{si } 0 \leq x \leq 1 \quad f(x) = 1/2 > 0$$

$$\text{si } 1 < x \leq 3 \quad 0 \leq f(x) < 2/4$$

$$\text{si } x > 3 \quad f(x) = 0$$

Cumple la primera condición.

Segunda condición:

$$\int_{-\infty}^{+\infty} f(x) dx = \int_{-\infty}^0 f(x) dx + \int_0^1 f(x) dx + \int_1^3 f(x) dx + \int_3^{+\infty} f(x) dx$$

$$\int_{-\infty}^0 f(x) dx = 0$$

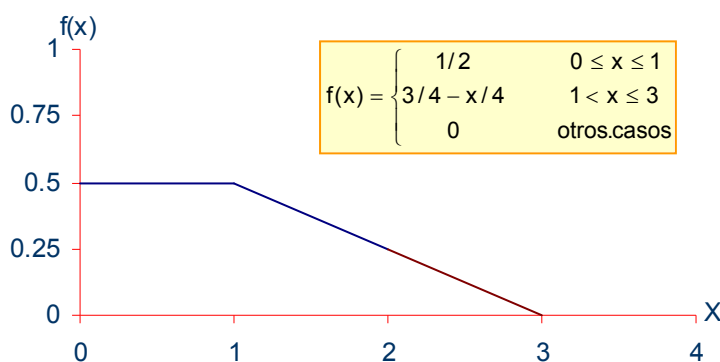
$$\int_3^{+\infty} f(x) dx = 0$$

$$\int_0^1 f(x) dx = \int_0^1 1/2 dx = 1/2$$

$$\int_1^3 f(x) dx = \int_1^3 (3/4 - x/4) dx = 1/2$$

$$\int_{-\infty}^{+\infty} f(x) dx = 0 + 1/2 + 1/2 + 0 = 1$$

Cumple la segunda condición; luego es una función de densidad.



Observe que el área debajo de la curva es igual a la unidad.

Probabilidad de un evento. Si se desea calcular la probabilidad de que la variable aleatoria tome valores entre a y b , donde $[a, b]$ es un subconjunto de R_x , es decir, la probabilidad de un evento $A = \{ x / a \leq x \leq b \}$, se define como:

$$P(A) = P(a \leq x \leq b) = \int_a^b f(x)dx$$

Características de la probabilidad de un evento

(i) $f(x)$ es una función de densidad y no representa una probabilidad, solamente cuando se integra ésta función entre dos puntos produce una probabilidad.

$$(ii) P(a \leq x \leq a) = P(x = a) = \int_a^a f(x)dx = 0$$

La probabilidad en un punto es cero. El que la probabilidad sea cero en el caso de V.A. continua no quiere decir que ese evento ($x=a$) es imposible de que ocurra. En el caso discreto si la probabilidad en un punto es cero, nos indicaría que ese evento ($x=a$) es imposible de que ocurra.

$$(iii) P(a < x < b) = P(a < x \leq b) = P(a \leq x < b) = P(a \leq x \leq b) = \int_a^b f(x)dx$$

ESPERANZA DE UNA V.A. CONTINUA.

Definición .- Sea X una V.A. continua con función de densidad $f(x)$, entonces el valor esperado (media) de la V.A. X se define por:

$$\mu = E[X] = \int_{-\infty}^{+\infty} x f(x)dx$$

Teorema. "La esperanza de una función de una V. A. continua X ". Sea X una V.A. continua con función de densidad $f(x)$, e $Y = H(x)$ una función de X , el valor esperado de Y esta dado por:

$$E[Y] = E[H(x)] = \int_{-\infty}^{+\infty} H(x)f(x)dx$$

Las propiedades para la esperanza de una v.a. continua son las mismas que para la v.a. discreta.

1. $E[k] = k$
2. $E[X \pm a] = E[X] \pm a$
3. $E[aX] = a E[X]$
4. $E[aX + b] = a E[X] + b$

k, a, b son constantes

Variación de una variable aleatoria continua.

X una v.a. continua con función de densidad $f(x)$, la variación de X se define por:

$$\sigma^2 = V(x) = E\left[(x-\mu)^2\right] = \int_{-\infty}^{+\infty} (x-\mu)^2 f(x)dx$$

$$\sigma^2 = V(x) = E[X^2] - \mu^2 = \int_{-\infty}^{+\infty} x^2 f(x)dx - \mu^2$$

Las propiedades de la variación en el caso continuo son las mismas que en el caso discreto:

1. $V(k) = 0$
2. $V(X \pm a) = V[X]$
3. $V(aX) = a^2V(X)$
4. $V(aX + b) = a^2V(X)$

k, a, b son constantes.

Mediana de una v.a. continua

"Me" es un valor particular en X tal que $P(X \leq Me) = P(X \geq Me) = 0.5$

Ejemplo. sea X una v.a. continua con función de densidad f(x):

$$f(x) = \begin{cases} kx & , \quad 0 \leq x \leq 6 \\ 0 & , \quad \text{Otros.Valores.de.X} \end{cases}$$

Se tiene además la siguiente función de x dado por $Y = 10 + 2X$

Hallar:

- a) Valor de K
- b) $E[X]$
- c) $Me(X)$ {mediana de X}
- d) $E[Y]$

Solucion. f(x) es una función de densidad, entonces:

$$\int_{-\infty}^{+\infty} f(x)dx = 1$$

a) $K = ?$

$$\int_{-\infty}^{+\infty} f(x)dx = \int_{-\infty}^0 0dx + \int_0^6 kxdx + \int_6^{+\infty} 0dx = 0 + 18k + 0 = 1$$

$$K = 1/18$$

$$f(x) = \begin{cases} x/18 & \quad 0 \leq x \leq 6 \\ 0 & \quad \text{Otros.valores.de.X} \end{cases}$$

b) $E[X] = ?$

$$E[x] = \int_{-\infty}^{+\infty} xf(x)dx = \int_{-\infty}^0 xf(x)dx + \int_0^6 xf(x)dx + \int_6^{+\infty} xf(x)dx$$

$$E[x] = 0 + \int_0^6 x(x/18)dx + 0 = \frac{x^3}{54} \Big|_0^6 = 4$$

c) $Me(X) = ?$

$Me(X) = x_0$ tal que $P(X \leq x_0) = 0.5$

$$P(x \leq x_0) = \int_{-\infty}^0 0 dx + \int_0^{x_0} (x/18) dx = \frac{x^2}{36} \Big|_0^{x_0} = \frac{x^2}{36} = 0.5$$

$$x_0 = \sqrt{18}$$

entonces $Me(x) = \sqrt{18}$

d) $E[Y] = E[10+2X] = 10 + 2 E[X] = 10 + 2(4) = 18$

Ejemplo. Un combustible para cohetes va a contener cierta cantidad (X) de un compuesto en particular. El fabricante tendrá una utilidad neta en el combustible (por galón) cuya función de densidad de X es:

$T(x) = S/. 0.05$ soles por galón si $3 \leq x \leq 6$

$T(x) = S/. -0.10$ soles por galón en otro caso.

X es una v.a y tiene una función de densidad:

$$f(x) = \begin{cases} 1/2 - x/18 & 3 \leq x \leq 9 \\ 0 & \text{en otro caso} \end{cases}$$

Hallar:

- Cantidad esperada del compuesto por galón de combustible
- Variación de la cantidad del compuesto por galón de combustible.
- La utilidad neta esperada.

Solucion:

a) $E[X] = ?$

$$E[x] = \int_{-\infty}^{+\infty} xf(x)dx = \int_3^9 x(1/2 - x/18)dx = \left(\frac{x^2}{4} - \frac{x^2}{54} \right) \Big|_3^9 = 5$$

b) $V(X) = ?$

$$V[x] = \int_{-\infty}^{+\infty} x^2 f(x)dx - \mu^2 = \int_3^9 x^2 (1/2 - x/18)dx - 5^2 = 2$$

c) $E[T(x)]$

$$\int_{-\infty}^{+\infty} T(x)f(x) = \int_{-\infty}^3 (-10)(0)dx + \int_3^6 (5)(1/2 - x/18)dx + \int_6^{+\infty} (-10)(1/2 - x/18)dx = 1.25$$

PRINCIPALES DISTRIBUCIONES DE VARIABLES ALEATORIAS DISCRETAS

1. DISTRIBUCION BINOMIAL

Una variable binomial se deriva a partir del siguiente experimento aleatorio: E: realizar "r" pruebas de bernoulli.

Pruebas de Bernoulli. Son experimentos que tienen las siguientes características:

1. Para cada prueba se tiene dos resultados posibles, éxito y fracaso (E y F), (la ocurrencia y la no ocurrencia de un evento particular).

Probabilidad de éxito = p
 Probabilidad de fracaso = q = 1-p

2. El valor de "p" no se altera a través de las repeticiones de la prueba.
3. Las pruebas realizadas son independientes.

Ejemplo: E: Lanzamiento de una moneda "r" veces.

El lanzamiento corresponde a una prueba de bernoulli.

Exito = cara ==> p = 1/2
 Fracaso = sello ==> q = 1/2

Sea la v.a. discreta X: número de caras (éxitos)

1ra Prueba 2da Prueba r ésima Prueba

$\Omega_1 = \{E,F\}$ $\Omega_2 = \{E,F\}$ $\Omega_r = \{E,F\}$
 $\Omega_1 = \{C,S\}$ $\Omega_2 = \{C,S\}$ $\Omega_r = \{C,S\}$

Espacio muestral total = $\Omega = \Omega_1 \times \Omega_2 \times \dots \times \Omega_r$

Nro de elementos de $\Omega = 2^r$, $\Omega = \{r \text{-adas ordenadas}\}$.

La probabilidad de una r -ada ordenada cualquiera :

$P(E,E,F,E,F,\dots,F,E) = p \cdot p \cdot q \cdot p \cdot q \cdot \dots \cdot q \cdot p$

$= p^x q^{r-x}$, x = nro de exitos
 r-x = nro. de fracasos

$P(\text{éxitos} = x) = P(X = x) = P_{x,r-x}^r p^x q^{r-x}$

Luego la función de una variable aleatoria con distribución binomial es:

$$P(X = x) = \binom{r}{x} p^x q^{r-x} ; x=0,1,2, \dots, r$$

Siendo "r" el número de pruebas de Bernoulli.

Experanza y variancia de una v.a. X que se distribuye como binomial

$$E[X] = \mu = rp$$

$$V(X) = \sigma^2 = rpq$$

Ejemplo 1: Una granja debe trasladarse a otro lugar, la probabilidad de encontrar un animal con un peso mayor de 80 kgr. es de $1/3$, y la probabilidad de encontrar un animal con un peso de a lo más 80 kgrs. es de $2/3$. Los animales de la granja son colocados en camiones en un número de cinco por camión:

- ¿Cuál es la probabilidad de encontrar por lo menos 4 animales que pesen a lo más 80 kgrs. ?
- ¿Cuál es el número medio de animales que pesen a lo más 80 kgrs. y se encuentren colocados en el camión ?

Solución:

X = Número de animales con un peso inferior o igual a 80 kgr.

$$P(X = x) = \binom{5}{x} (2/3)^x (1/3)^{5-x} ; x=0,1,2, \dots, 5$$

$$a) \quad P(X \geq 4) = P(X=4) + P(X=5)$$

$$P(X \geq 4) = \binom{5}{4} (2/3)^4 (1/3)^1 + \binom{5}{5} (2/3)^5 (1/3)^0$$

$$P(X \geq 4) = 112/234$$

$$b) \quad E[X] = rp = 5(2/3) = 10/3$$

Ejemplo 2 : Las ventas de detergente (en miles de kilos) en las tiendas de la ciudad de Arequipa el mes de Julio constituyen una variable aleatoria cuya función de densidad es:

$$f(x) = \begin{cases} 2 - 2x & , \quad 0 < x < 1 \\ 0 & , \quad \text{de otro modo} \end{cases}$$

Si al azar y con reemplazo se eligen 10 tiendas de la ciudad de Arequipa ¿Cuál es la probabilidad de que por lo menos en dos de ellas se hayan vendido más de 300 kilos de detergente ?

Solución :

V.A. discreta Y : número de tiendas que han vendido más de 300 kilos de detergente.

$$P(\text{éxito}) = p = P(x > 0.3) = \int_{0.3}^{+\infty} f(x)dx = \int_{0.3}^1 (2 - 2x)dx = 0.49$$

Entonces $Y \sim \text{Binomial} (r = 10; p=0.49)$

con función de probabilidad:

$$P(Y = y) = \binom{10}{y} (0.49)^y (0.51)^{10-y} ; y=0,1,\dots,10$$

$$P(Y \geq 2) = 1 - \left[\binom{10}{0} (0.49)^0 (0.51)^{10} + \binom{10}{1} (0.49)^1 (0.51)^9 \right]$$

$$P(Y \geq 2) = 1 - 0.001190442 - 0.011437409 = 0.987372$$

2. DISTRIBUCION DE POISSON

La distribución de Poisson se deduce de un proceso de Poisson o como el límite de la distribución binomial.

Primera deducción : como un Proceso de Poisson

El proceso aleatorio (Proceso de Poisson) esta formado por la ocurrencia de eventos discretos que se genera en un intervalo continuo (unidad de medida : longitud, área, volúmen), el número de ocurrencias es independiente de una unidad de medida a otra. La probabilidad de más de una ocurrencia en una unidad definida es muy pequeña que es despreciable en comparación de la probabilidad de una ocurrencia en la unidad de medida. El valor esperado de la variable es proporcional al tamaño de la unidad definida, esto se cumple si la unidad es muy pequeña.

Al Proceso de Poisson esta asociado un parámetro λ

La función de probabilidad de poisson:

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

x: Número de eventos discretos en "t" unidades de medida.

λ : es el número esperado o promedio de eventos discretos en una unidad de medida.

t: número de unidades de medida

e: 2.71828 (base del logaritmo neperiano)

La variable x así definida es una V.A.D. distribuida como Poisson con la función de probabilidad dada anteriormente.

Los siguientes casos aleatorios son Procesos de Poisson:

- Número de manchas por metro cuadrado de cierto objeto. Son eventos discretos porque se puede encontrar 0,1,2 ó muchas más manchas en un metro cuadrado. El metro cuadrado es la unidad de medida y es un intervalo continuo porque tiene infinitos puntos.
- El número de glóbulos rojos en una muestra de sangre.
- El número de vehículos que llegan a una estación de servicio durante una hora en un día determinado. Los eventos son 0, 1, 2,... vehículos. El intervalo continuo es 1 hora.

Ejemplo. La experiencia a demostrado que el número promedio de llamadas que llegan a un conmutador de una central es de 2 llamadas por minuto.

- a. ¿Cuál es la probabilidad de recibir 3 llamadas en 2 minutos ?.

- b. Si el conmutador puede recibir un máximo de 5 llamadas en 2 minutos. ¿Cuál es la probabilidad de que no pueda contestar todas las llamadas que entren en un periodo de 2 minutos ?

Solución:

V.A.D. X: número de llamadas que se presenten en 2 minutos.

Si son en promedio 2 llamadas por minuto, entonces en 2 minutos serían 4 llamadas, así: $\lambda t = 2(2) = 4$

$$a. P(X = x) = \frac{e^{-4} 4^x}{x!}$$

$$P(X = 3) = \frac{e^{-4} 4^3}{3!} = 0.195367$$

$$c. P(X > 5) = 1 - P(X \leq 5)$$

$$P(X \leq 5) = \frac{e^{-4} 4^0}{0!} + \frac{e^{-4} 4^1}{1!} + \frac{e^{-4} 4^2}{2!} + \dots + \frac{e^{-4} 4^5}{5!}$$

$$P(X > 5) = 1 - P(X \leq 5) = 1 - 0.7851 = 0.2149$$

Segunda deducción : Como límite de la distribución binomial.

La distribución de Poisson de una v.a. X es una distribución que se deriva de la distribución binomial; luego debe consistir en "r" pruebas de bernoulli.

Sea X es una v.a. discreta con distribución binomial:

$$P(X = x) = \binom{r}{x} p^x q^{r-x} ; x = 0, 1, 2, \dots, r$$

donde $r \rightarrow \infty$ ($r > 50$) y,

$$p \rightarrow 0 \text{ (} p < 0.01 \text{)}$$

Entonces x se distribuye según Poisson:

$$P(X = x) = \frac{e^{-\mu} \mu^x}{x!} , \text{ para } x = 0, 1, 2, \dots$$

$$\mu = rp = E[X]$$

Se usará la distribución de Poisson cuando $r \geq 50$ y $p < 0.01$

Ejemplo: Un almacén recibió 2,000 botellas de agua mineral. La probabilidad de que al transportar una botella resulte rota es igual a 0.003 . Hallar la probabilidad de que el almacén reciba rotas:

- Exactamente tres
- Por lo menos 2

Solución.

$r = 2,000$ botellas; $p = P(1 \text{ botella rota}) = 0.003$. Cumple las condiciones.

V.A. X : número de botellas rotas sigue una distribución de Poisson, con $\mu = 2000(0.003) = 6$ número promedio de botellas rotas.

$$a. P(X = 3) = \frac{e^{-6} \mu^3}{3!} = 0.089$$

$$b) P(X \geq 2) = P(X=2) + P(X=3) + \dots + P(X=2000)$$

$$P(X \geq 2) = 1 - P(X < 2) = 1 - (P(X=0) + P(X=1)) = 0.9826$$

Esperanza matemática y variancia de la v.a. X que se distribuye como Poisson.

1. Si se define un Proceso de Poisson:

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

$$V[X] = \lambda t$$

2. Si se define con el límite de la binomial:

$$E[X] = \mu = rp$$

$$V[X] = \sigma^2 = rp$$

3. DISTRIBUCION HIPERGEOMETRICA

Considere una población finita de N elementos, divididos en dos clases con " A " elementos en una clase (éxito) y $N-M$ elementos en otra clase (fracaso). Por ejemplo en el caso particular de N artículos fabricados, se puede considerar las clases: M elementos defectuosos y $(N-A)$ artículos no defectuosos.

Considere el siguiente experimento E : se extrae una muestra de tamaño " n " de la población de N elementos sin reemplazo.

Defina la variable aleatoria X como:

X : número de éxitos (de la clase de los " A " elementos) en la muestra.

Luego, $R_x = \{0, 1, 2, \dots, \min(n, A)\}$

La V.A. definida de esta forma se llama variable hipergeométrica y su función de probabilidad es:

$$P(X = x) = \frac{\binom{A}{x} \binom{N-A}{n-x}}{\binom{N}{n}}, \quad x = 0, 1, 2, \dots, \min(n, A)$$

La media de la variable aleatoria hipergeométrica X es:

$$\mu = E[x] = n \left[\frac{A}{N} \right]$$

La variancia de la variable aleatoria hipergeométrica X es:

$$\sigma^2 = V[x] = n \left[\frac{A}{N} \right] \left[\frac{N-A}{N} \right] \left[\frac{N-n}{N-1} \right]$$

Ejemplo. Se extraen al azar trece cartas sin reemplazo de una baraja de 52.

- Hallar la función de probabilidad para el número de cartas rojas en la muestra.
- ¿Cuál es la probabilidad de seleccionar 5 cartas rojas?
- ¿Cuál es la probabilidad de seleccionar todas las cartas del mismo color?
- Si las cartas fueron del mismo color, ¿cuál es la probabilidad de que estas sean negras.
- ¿Cuál es la media y la variancia del número de cartas rojas?

Solución. N = 52 cartas

A = 26 Cartas rojas

N-A = 26 Cartas negras

Se selecciona una muestra de tamaño n=13 sin reemplazo. Entonces la v.a. X: número de cartas rojas en la muestra es una variable hipergeométrica, con función:

$$P(X = x) = \frac{\binom{A}{x} \binom{N-A}{n-x}}{\binom{N}{n}} = \frac{\binom{26}{x} \binom{26}{13-x}}{\binom{52}{13}}; \quad x = 0, 1, \dots, 13$$

$$b. \quad P(X = 5) = \frac{\binom{26}{5} \binom{26}{8}}{\binom{52}{13}} = \frac{26!}{5!2!} \frac{26!}{8!1!}{52!} = \frac{52!}{13!39!}$$

$$c. \quad P((X=13) \cup (X=0)) = P(X=13) + P(X=0)$$

$$d. \quad P(x = 0 / (x = 13) \cup (x = 0)) = \frac{P(x = 0)}{P(x = 13) + P(x = 0)} = 0.5$$

$$e. \quad \mu = E[x] = n \left[\frac{A}{N} \right] = 13 \left[\frac{26}{52} \right] = 6.5$$

Número promedio de cartas rojas al extraer 13 cartas de la baraja es de 6.5

$$\sigma^2 = n \left[\frac{A}{N} \right] \left[\frac{N-A}{N} \right] \left[\frac{N-n}{N-1} \right] = 13 \left[\frac{26}{52} \right] \left[\frac{26}{52} \right] \left[\frac{39}{51} \right]$$

$$\sigma^2 = 2.4853$$

PRINCIPALES DISTRIBUCIONES DE VARIABLES ALEATORIAS CONTINUAS

1. DISTRIBUCION NORMAL

Definición .- Sea X una v.a. continua con media μ y variancia σ^2 , luego esta variable se distribuye como variable Normal si tiene la siguiente función de densidad:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-1/2 \left[\frac{x-\mu}{\sigma} \right]^2}; -\infty < x < +\infty$$

Graf 21

Notación $X \sim N(\mu, \sigma^2)$

Características:

1. La curva $f(x)$ es de forma acampanada y tiene como asíntota el eje de las abscisas.
2. Es simétrica respecto a la recta vertical $X=\mu$
3. Presenta una relación entre media " μ " y desviación estandar σ :

- $\int_{\mu-\sigma}^{\mu+\sigma} f(x)dx = 0.6826$
- $\int_{\mu-2\sigma}^{\mu+2\sigma} f(x)dx = 0.9545$
- $\int_{\mu-3\sigma}^{\mu+3\sigma} f(x)dx = 0.9973$

DISTRIBUCION NORMAL ESTANDAR : Z

Definición .- Si Z es una v.a. continua distribuida normalmente con media 0 y variancia 1, entonces Z tiene una distribución Normal Estándar con la siguiente función de densidad:

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-1/2(z^2)}; -\infty < z < +\infty$$

Graf 22

$$\mu_z = E[Z] = 0$$

$$\sigma_z^2 = V(Z) = 1$$

Notación $Z \sim N(0, 1)$

Características:

1. La curva $f(z)$ es de forma acampanada y tiene como asíntota el eje de las abscisas.
2. Es simétrica respecto a la recta vertical $z=\mu_z=0$
3. Presenta una relación entre media y desviación estándar.

- $\int_{\mu-1}^{\mu+1} f(z)dz = 0.6826$
- $\int_{\mu-2}^{\mu+2} f(z)dz = 0.9545$
- $\int_{\mu-3}^{\mu+3} f(z)dz = 0.9973$

Teorema.- Si X es una v.a. normal con media μ y variancia σ^2 , entonces la variable transformada $Z = (X-\mu)/\sigma$ se distribuye como normal estándar con media 0 y variancia 1.

Uso de la tabla de Z . La tabla de Z expresa la probabilidad acumulativa para un valor z_0 , es decir $P(-\infty < Z < z_0)$

Graf 23

Z_0	0.01	0.02	0.03	...	0.09
0.0	0.5000	0.5040	0.5080	...	0.5359
0.1	0.5398	0.5438	0.5478	...	0.5753
0.2	0.5793	0.5832	0.5871	...	0.6141
...				...	
1.0	0.8413	0.8438	0.8461	...	0.8621
...				...	
2.0	0.9772	0.9778	0.9783	...	0.9817
...				...	
3.5	0.9998	0.9998	0.9998	...	0.9998

Para z_0 positivo

(i) $P(-\infty < Z < z_0)$, directo de la tabla

Ej. $P(-\infty < Z < 0.12) = 0.5478$

(ii) $P(-\infty < Z < -z_0) = 1 - P(-\infty < Z < z_0)$

Ej. $P(-\infty < Z < -0.29) = 1 - P(-\infty < Z < 0.29)$
 $= 1 - 0.6141 = 0.3859$

(iii) $P(-z_0 < Z < z_0) = 2 P(-\infty < Z < z_0) - 1$

Ej. $P(-1 < Z < 1) = 2P(-\infty < Z < 1) - 1$
 $= 2(0.8413) - 1 = 0.6826$

Ejercicios.

1. Hallar $P(-0.11 < Z < 0.22)$

$$P(-0.11 < Z < 0.22) = P(Z < 0.22) - P(Z < -0.11)$$

$$= P(Z < 0.22) - (1 - P(Z < 0.11))$$

$$= P(Z < 0.22) + P(Z < 0.11) - 1$$

$$P(-0.11 < Z < 0.22) = 0.5871 + 0.5438 - 1 = 0.1309$$

2. Hallar $P(0.21 < Z < 0.29)$

$$P(0.21 < Z < 0.29) = P(Z < 0.29) - P(Z < 0.21) = 0.6141 - 0.5832 = 0.0309$$

3. Hallar el valor de z_0 , tal que $P(Z > z_0) = 0.4364$. Implica que $P(Z < z_0) = 1 - 0.4364 = 0.5636$

El valor que satisface es $z_0 = 0.16$

4. Hallar el valor de z_0 , tal que $P(Z > z_0) = 0.8461$. Como la probabilidad es mayor de 0.50, entonces el valor de z_0 es menor de cero.

$$P(Z > z_0) = 1 - P(Z < z_0) = 1 - (1 - P(Z < -(z_0))) = P(Z < -(z_0))$$

El valor $-(z_0)$ es positivo y se ve directo en la tabla, así, para la probabilidad de 0.8461, le corresponde a $-(z_0) = 0.12$, es decir $z_0 = -0.12$

5. Los tubos fabricados por cierta máquina tienen un diámetro medio de $\mu = 9.8$ mm, con desviación estándar $\sigma = 0.795$ mm. ¿Qué % de tubos será rechazado, sino se aceptan diámetros inferiores a 9.0 mm.?. Asumir que la v.a. medida del diametro del tubo sigue una distribución normal.

$$P(X < 9) = P\left(\frac{X - \mu}{\sigma} < \frac{9 - 9.8}{0.795}\right) = P(Z < -1.01)$$

$$P(Z < -1.01) = 1 - P(Z < 1.01) = 1 - 0.8438 = 0.1562$$

Entonces el % de tubos rechazados es de 15.6%

Problema. Una linterna grande es alimentada por cinco pilas, suponga que el tiempo de vida de una pila está normalmente distribuida con media $\mu = 120$ horas y una desviación estándar $\sigma = 10$ horas. La linterna dejará de funcionar si se agota una o más de sus pilas. suponiendo que el tiempo de vida de las pilas son independientes ¿Cuál es la probabilidad de que la linterna funcione más de 100 horas ?.

Solución.

V.A. X : tiempo de vida de una pila. $X \sim N(120 \text{ horas}, (10 \text{ horas})^2)$

$$P(\text{una pila dure más de 100 horas}) = P(X > 100)$$

$$P(X > 100) = P(Z > (100 - 120)/10) = P(Z > -2) = P(Z < 2) = 0.9772$$

v.a.Y: número de pilas que funcionan más de 100 horas de las 5 usadas.

$$P(Y = k) = \binom{5}{k} 0.9772^k 0.0228^{5-k}; \text{ para } k=0,1,2,3,4,5$$

$$P(Y=k) = 0 \quad \text{para otros valores de } k.$$

La probabilidad solicitada $P(Y=5)$ es:

$$P(Y = 5) = \binom{5}{5} 0.9772^5 0.0228^0 = 0.9772^5 = 0.8911$$

TEOREMA DEL LIMITE CENTRAL.

El teorema del límite central introducido por De Moivre a principios del siglo XVIII, ha sido expresado de muchas formas, una de ellas es la siguiente:

Sea $X_1, X_2, X_3, \dots, X_n$ una sucesión de variables aleatorias independientes distribuidas idénticamente con $E[X_i] = \mu$ y $V(X_i) = \sigma^2$ finito. Si S_n representa la suma de estas variables, entonces:

$\frac{S_n - n\mu}{\sigma\sqrt{n}}$, se aproxima a una variable normal con media 0 y variancia 1, cuando n tiende al infinito.

Otra forma de mencionar este teorema, considera lo siguiente: Sea $f(x)$ la función de densidad de una variable aleatoria X con media μ y variancia σ^2 , y \bar{X}_n el promedio de la muestra de tamaño " n " de $f(x)$. La variable Y_n definido como :

$Y_n = \frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}}$, se aproxima a una normal con media 0 y variancia 1

La importancia de este teorema es que nada se dice de la forma de distribución de $f(x)$. Cualquiera que sea la distribución con variancia finita, la media muestral tendrá una distribución cercana a la normal cuando la muestra sea grande.

DISTRIBUCION MUESTRAL.

Valor Estadístico.- Es una v.a. que depende de la muestra observada. La distribución del valor estadístico se llama distribución muestral y a la desviación estandar el error estándar.

Distribución muestral de la media (\bar{x})

Si x_1, x_2, \dots, x_n es una muestra aleatoria de tamaño " n " de una v.a. X con media μ y variancia σ^2 , entonces la distribución de la media muestral es aproximadamente normal con media μ y variancia σ^2/n , para $n \geq 30$.

$$\mu_{\bar{x}} = \mu; \quad \sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}$$

La variable transformada Z :

$$Z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \text{ sigue una normal con } \mu=0 \text{ y } \sigma^2=1$$

Tiene aproximadamente una distribución normal estándar.

Esta definición es válida para poblaciones finitas o infinitas, discretas o continuas cuando $n \geq 30$

Si la población tiene distribución normal, la distribución de la media de la muestra es normal para cualquier " n ".

Si la muestra es sin reemplazo de una población finita de tamaño " N ", entonces:

$$\mu_{\bar{x}} = \mu; \quad \sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} \left(\frac{N-n}{N-1} \right)$$

$\left(\frac{N-n}{N-1}\right)$, Factor de corrección para poblaciones finitas

El factor de corrección puede omitirse si $n \geq 30$ ó $n/N \leq 0.01$, en tal caso:

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}$$

Ejemplo 1 .- La fábrica de neumáticos "DURAMAS" produce un tipo de neumáticos que tiene un tiempo medio de vida útil de 80,000 Kms. y una desviación estándar de 8,000 Kms. En un día la fabrica produjo 64 neumáticos. ¿Cuál es la probabilidad de que el tiempo promedio de vida util de esta muestra sea menor de 82,000 Kms.

Solución : v.a. X : tiempo de duración de un neumático.

$$\mu_x = 80,000, \sigma_x^2 = (8,000)^2$$

La muestra de tamaño $n=64$ neumáticos.

La distribución de \bar{x} es normal

$\bar{x} \sim N(80,000, (8,000)^2/64)$, entonces:

$$Z = \frac{\bar{x} - 80,000}{\frac{8,000}{\sqrt{64}}}$$

Se pide $P(\bar{x} < 82,000) = P(Z < (82000 - 80000)/(8000/8))$

$$P(\bar{x} < 82,000) = P(Z < 2) = 0.9772$$

Ejemplo 2 .- El peso de los pescados atrapados por un barco es aproximadamente normal con una media de 4.5 kilos y una desviación estándar de 0.5 kilos. Si los pescados se embarcan en cajas que contienen 20 pescados. ¿Cuál es la probabilidad de que el peso total de los pescados contenidos en una caja sea mayor de 92 kilos ?.

Solución : v.a. X : peso de un pescado.

$$X \sim N(4.5, (0.5)^2)$$

$$P(x_1+x_2+\dots+x_{20} > 92) = P\left(\sum_{i=1}^{20} x_i > 92\right) = P\left(\sum_{i=1}^{20} x_i/20 > 92/20\right)$$

Es equivalente a $P(\bar{x} > 4.6)$.

Como $\bar{x} \sim N(4.5, (0.5)^2/20)$, entonces:

$$P(\bar{x} > 4.6) = P(Z > 0.89) = 1 - P(Z < 0.89) = 1 - 0.8133$$

$$P(\bar{x} > 4.6) = 0.1867$$

Distribución muestral de las diferencias de dos medias

Si \bar{x}_1 y \bar{x}_2 son las medias de dos muestras aleatorias de tamaño n_1 y n_2 de dos poblaciones con medias μ_1 y μ_2 y variancias σ_1^2 y σ_2^2 respectivamente, entonces la distribución muestral de las diferencias de medias:

$$(\bar{x}_1 - \bar{x}_2)$$

es aproximadamente normal con media $\mu_1 - \mu_2$ y variancia $\sigma_1^2/n_1 + \sigma_2^2/n_2$, luego la variable aleatoria Z:

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

Tiene aproximadamente una distribución normal estándar para $n \geq 30$.

Si las poblaciones son normales, la distribución de la

variable : $(\bar{x}_1 - \bar{x}_2)$ es normal

Ejemplo .- Suponga que en una oficina de correos (A) el peso (en grs.) de las cartas tiene una distribución normal con media 350 grs. y desviación estándar de 56.27 grs.

a) ¿Cuál debe ser el tamaño de muestra para que la probabilidad de que el peso promedio de las cartas difiera de su media en menos de 15 grs. sea igual a 0.9426 ?

b) En otra oficina de correos (B) se a encontrado que el peso (en grs.) de las cartas tiene una distribución normal con media de 320 grs. y desviación estándar de 50 grs. Si se extraen 20 cartas de la oficina de correos , ¿Cuál es la probabilidad de que el peso promedio de las cartas extraidas en la oficina A sea mayor al peso promedio de las cartas de la oficina B en no menos de 10 grs.

Solución: $X_A \sim N(350, (56.27)^2)$

$$a) n = ? \quad P(|\bar{x}_A - \mu| < 15) = 0.9426$$

$$\bar{x}_A \sim N(350, (56.27)^2/n)$$

$$P(|\bar{x}_A - \mu| / (\sigma / \sqrt{n}) < 15 / (56.27 / \sqrt{n})) = 0.9426$$

$$\text{es equivalente a: } P(|Z| < 0.2666 \sqrt{n}) = 0.9426$$

$$P(|Z| < z_0) = P(-z_0 < Z < z_0) = P(Z < z_0) - P(Z < -z_0)$$

$$P(|Z| < z_0) = P(Z < z_0) - [1 - P(Z < z_0)] = 2P(Z < z_0) - 1$$

$$2P(Z < z_0) - 1 = 0.9426$$

$$P(Z < z_0) = 0.9713 \quad (\text{ver tabla de Z acumulada})$$

el valor de z_0 correspondiente es 1.9, por lo tanto:

$$0.2666 \sqrt{n} = 1.9, \text{ resulta } n = 51$$

$$b) n_A = 20 \quad \bar{X}_A \sim N(350, (56.27)^2/20)$$

$$n_B = 20 \quad \bar{X}_B \sim N(320, (50)^2/20)$$

$$(\bar{X}_A - \bar{X}_B) \sim N(350-320, (56.27)^2/20 + (50)^2/20)$$

$$(\bar{X}_A - \bar{X}_B) \sim N(30, 283.31)$$

$$P(\bar{X}_A - \bar{X}_B \geq 10) = P(Z \geq (10-30)/16.8318) = P(Z \geq -1.19)$$

$$P(Z \geq -1.19) = P(Z < 1.19) = 0.8835 \text{ (tabla de Z)}$$

2. DISTRIBUCION CHI-CUADRADO : $X^2_{(v) gl}$.

Definición .- Sea $Z_1, Z_2, Z_3, \dots, Z_v$ variables aleatorias normales e independientes con media 0 y variancia 1.

La variable definida como : $w = \sum_{i=1}^v Z_i^2$ es una variable aleatoria que sigue una distribución Chi cuadrado con "v" gl. y la función de densidad dada por:

$$f(w) = \begin{cases} \frac{1}{\Gamma(v/2)} \frac{1}{2^{v/2}} w^{(v/2)-1} e^{-w/2} & ; \quad 0 < w < +\infty \\ 0 & ; \quad \text{Otros Valores} \end{cases}$$

v = grados de libertad

Graf 24

$\Gamma(a)$ es una función matemática definida como función gamma, la relación de esta función es como sigue:

$$\Gamma(a) = (a-1)\Gamma(a-1)$$

$$\Gamma(1/2) = \sqrt{\pi}$$

si a es un valor entero, $\Gamma(a) = (a-1)!$

La media y variancia de W, son respectivamente:

$$E(w) = v$$

$$V(w) = 2v$$

Ejemplo : Z_1, Z_2, Z_3 se distribuyen normalmente con media cero y variancia uno, entonces la variable:

$$w = Z_1^2 + Z_2^2 + Z_3^2$$

es Chi Cuadrado con 3 grados de libertad. Su función de densidad es:

$$f(w) = \frac{1}{\Gamma(3/2)} \frac{1}{2^{3/2}} w^{(3/2)-1} e^{-w/2}$$

$$f(w) = \frac{1}{\sqrt{2\pi}} w^{1/2} e^{-w/2}$$

$$E[w] = 3$$

$$V(w) = 6$$

Características :

1. Distribución continua con tendencia asimétrica hacia la derecha.
2. Los valores de W son positivos.
3. Para cada grado de libertad se tiene una distribución Chi Cuadrado.
4. A medida que aumenta el número de grados de libertad, la distribución tiende a la simetría.

Teorema . "Aditividad de la distribución Chi Cuadrado". Si W_1, W_2, \dots, W_n son variables aleatorias independientes distribuidas cada una como Chi Cuadrado con v_1, v_2, \dots, v_n grados de libertad respectivamente, entonces: $W = \sum W_i$ se distribuye como Chi Cuadrado con $v = \sum v_i$ grados de libertad.

Teorema . Si de una población normalmente distribuida con media desconocida y variancia σ^2 , se toman muestras de tamaño "n" entonces la variable W, definida por:

$$w = \sum_{i=1}^n \left[\frac{x_i - \bar{x}}{\sigma^2} \right]^2$$

se distribuye como Chi Cuadrado con (n-1) grados de libertad.

Donde \bar{x} : es el promedio de los "n" elementos muestrales.

Demostración. Por definición, $\sum_{i=1}^n z_i^2$ es una Chi Cuadrado con n gl.

$$z_i = \frac{x_i - \mu}{\sigma}; \quad \sum_{i=1}^n \left[\frac{x_i - \mu}{\sigma} \right]^2 \sim \chi^2 (n)gl$$

$$w = \sum_{i=1}^n \left[\frac{x_i - \mu}{\sigma} \right]^2 = \sum_{i=1}^n \left[\frac{(x_i - \bar{x}) + (\bar{x} - \mu)}{\sigma} \right]^2$$

$$w = \sum_{i=1}^n \left[\frac{x_i - \bar{x}}{\sigma} \right]^2 + \sum_{i=1}^n \left[\frac{(\bar{x} - \mu)}{\sigma} \right]^2 + 2 \sum_{i=1}^n \left[\frac{(x_i - \bar{x})(\bar{x} - \mu)}{\sigma} \right]$$

$$\sum_{i=1}^n \left[\frac{x_i - \mu}{\sigma} \right]^2 = \sum_{i=1}^n \left[\frac{x_i - \bar{x}}{\sigma} \right]^2 + \sum_{i=1}^n \left[\frac{(\bar{x} - \mu)}{\sigma} \right]^2 + 0$$

$$\chi_{(n)}^2 \quad \chi_{(n-1)}^2 \quad \chi_{(1)}^2$$

Por propiedad aditiva de Chi Cuadrado se tiene que W sigue como Chi Cuadrado con $(n-1)$ gl.

La variable W puede ser escrita como :

$$w = \frac{(n-1)S^2}{\sigma^2} \sim \chi_{(n-1)gl}^2$$

S^2 : variancia de la muestra; σ^2 : variancia poblacional

Uso de la tabla de Chi Cuadrado (probabilidad acumulativa)

La primera columna corresponde a los grados de libertad ($v=n-1$), la primera fila a la probabilidad de X^2 menor que X^2_0 y el contenido de la tabla son los valores de X^2_0 .

Cada fila de la tabla corresponde a una distribución de Chi Cuadrado.

Graf 26

Gl	Probabilidad = $P(X^2 < X^2_0)$								
$V=n-1$	0.01	0.02	0.05	0.10	0.20	...	0.90	...	0.99
1	0.000	0.001	0.004	0.016	0.064	...	2.706	...	6.635
2	0.020	0.040	0.103	0.211	0.446	...	4.605	...	9.210
11	3.053	3.609	4.575	5.578	6.989	...	17.275	...	24.725
15	5.229	5.985	7.261	8.547	10.307	...	22.307	...	30.578
20	8.260	9.237	10.851	12.443	14.578	...	28.412	...	37.566

EJERCICIOS

- Para $n=21$. $P(X^2 < 37.566) = 0.99$; $P(X^2 > 37.566) = 0.01$
- Hallar X^2_0 , para $n=21$, si $P(X^2_0 < X^2 < 10.851) = 0.04$
 Si $P(X^2 < 10.851) = 0.05$; entonces, $P(X^2 < X^2_0) = 0.01$
 Según la tabla, para 20 gl, $X^2_0 = 8.260$
- Hallar X^2_0 , para $n=16$, si $P(10.307 < X^2 < X^2_0) = 0.70$
 Si $P(X^2 < 10.307) = 0.20$; entonces, $P(X^2 > 10.307) = 0.80$
 implica que $P(X^2 > X^2_0) = 0.80 - 0.70 = 0.10$, y la probabilidad del complemento, dado por $P(X^2 < X^2_0) = 0.90$ en tabla, para 15 gl. el valor de X^2_0 es de 22.307
- De una población normal con media desconocida y variancia 4 se extrae una muestra de 14 observaciones. Hallar $P(S > 1.6299)$, donde S corresponde a la desviación estandar de la muestra.

$$\frac{(n-1)S^2}{\sigma^2} = \frac{(14-1)S^2}{4} = \frac{13S^2}{4} \sim \chi_{(13)gl}^2$$

$$P(S > 1.6299) = P(S^2 > 1.6299^2) = P\left[\frac{(n-1)S^2}{\sigma^2} > \frac{13(1.6299)^2}{4}\right]$$

$$P(\chi_{(13)gl}^2 > 8.634) = 1 - P(\chi_{(13)gl}^2 < 8.634) = 1 - 0.2$$

$$P(S > 1.6299) = 0.80$$

3. DISTRIBUCION "t" DE STUDENT

Definición .- Sea "Z" una v.a. normal estándar y "Y" otra variable distribuida como Chi Cuadrado con "v" grados de libertad. La variable "t" definida como:

$$t = \frac{Z}{\sqrt{\frac{Y}{v}}}; \text{ Tiene una distribución de student con } v \text{ gl.}$$

La función de densidad f(t) definida como:

$$f(t) = \begin{cases} \frac{\Gamma\left(\frac{v+1}{2}\right)}{\sqrt{v\pi}\Gamma\left(\frac{v}{2}\right)} \left(1 + \frac{t^2}{v}\right)^{-\frac{v+1}{2}} & ; -\infty < t < +\infty \\ 0 & ; \text{ Otros Valores} \end{cases}$$

Graf 27

Notación .- $t \sim t_{(v)gl}$

Características .

1. Es una curva simétrica y asintótica, de forma acampanada
2. Alcanza su máxima altura en $t=0$, donde: $\mu_t = Me_t = Mo_t = 0$
3. Para cada valor de "v" existe una curva de probabilidad.
4. A medida que "v" aumenta, la distribución de "t" se aproxima a la distribución normal estándar.
5. $\mu_t = 0$, $v > 1$
 $\sigma_t^2 = v/(v-2)$, $v > 2$

Teorema .- La distribución de probabilidades de la variable t, definida como:

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

es una "t" de student con (n-1) grados de libertad.

Uso de la tabla de "t" de student (probabilidad acumulativa)

La primera columna corresponde a los grados de libertad (v=n-1), la primera fila a la probabilidad de t menor que t₀ y el contenido de la tabla comprende los valores de t₀

Cada fila de la tabla corresponde a una distribución de t-student. Graf 27

Gl. v=n-1	Probabilidad = P(t < t ₀)						
	0.001 ...	0.2	0.8	0.9	...	0.995	0.999
1	-318.309	-1.376	1.376	3.078	...	63.656	318.294
2	-22.327	-1.061	1.061	1.886	...	9.965	22.327
3	-10.215	-0.978	0.978	1.658	...	5.841	10.214
...							
10	-4.144	-0.879	0.879	1.372	...	3.169	4.144
...							
15	-3.733	-0.866	0.866	1.341	...	2.947	3.733

EJERCICIOS

1. Para n=16.

$$P(t < 0.866) = 0.80 ; \quad P(t < -0.866) = 0.20$$

$$P(-0.866 < t < 0.866) = 0.80 - 0.20 = 0.60$$

$$P(t > 0.866) = 1 - 0.80 = 0.20$$

$$P(t > -0.866) = 1 - 0.20 = 0.80$$

2. Para v=3 gl. Hallar P(t > -0.978

De la tabla P(t < -0.978) = 0.20; entonces:

$$P(t > -0.978) = 1 - 0.20 = 0.80$$

3. Determinar para 10 gl. t₀, tal que P(t > t₀) = 0.005

De la tabla P(t < t₀) = 1 - 0.005 = 0.995

El valor que corresponde a t₀ es de 3.169

4. De una población normal con media μ se extrae una muestra de tamaño 16 y variancia muestral de 2.25 ¿cuál es la probabilidad de que la media muestral difiera de su media poblacional en una cantidad mayor a 0.3247 ?.

Solución: P(| $\bar{x} - \mu$ | > 0.3247) = ?

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{\bar{x} - \mu}{\frac{\sqrt{2.25}}{\sqrt{16}}} \sim t_{(15)gl}$$

$$P\left[\frac{|\bar{x} - \mu|}{\sqrt{\frac{S^2}{n}}} > \frac{0.3247}{\sqrt{\frac{2.25}{16}}}\right] = P(|t_{(15)gl}| > 0.866)$$

es equivalente a:

$$P(t_{(15)gl} > 0.866) + P(t_{(15)gl} < -0.866) = (1 - 0.8) + 0.2$$

$$P(|t_{(15)gl}| > 0.866) = 0.4$$

4. DISTRIBUCION "F" DE FISHER

Definición .- Sea " W_1 " una v.a. distribuida como Chi cuadrado con v_1 grados de libertad y " W_2 " otra variable aleatoria con distribución Chi cuadrado con v_2 grados de libertad; independientes, entonces la variable aleatoria "F" definida como:

$$F = \frac{\frac{W_1}{v_1}}{\frac{W_2}{v_2}} \sim F_{(v_1, v_2)gl}$$

significa que sigue una distribución F con gl. v_1 y v_2

Graf28

Características .

1. Continua y asimétrica hacia la derecha
2. Los valores de F son valores positivos, $F > 0$
3. Para cada par de grados de libertad v_1, v_2 se tiene una distribución de F.
4. A medida que aumentan los valores de v_1 y v_2 , la curva se hace menos asimétrica.
5. $\mu_F = v_1 / (v_2 - 2)$ para $v_2 > 2$

$$\sigma_F^2 = \frac{2v_2^2(v_1 + v_2 - 2)}{v_1(v_2 - 2)^2(v_2 - 4)}; \text{ para } v_2 > 4$$

6. Tiene la propiedad recíproca:

$$F_{(\alpha)(v_1, v_2)gl} = \frac{1}{F_{(1-\alpha)(v_2, v_1)gl}}$$

Teorema .- Si de una población normal con media desconocida y variancia σ_1^2 se extrae una muestra de tamaño n_1 cuya variancia es de S_1^2 , y de otra población normal también con media desconocida y variancia σ_2^2 se extrae una muestra de tamaño n_2 cuya variancia es de S_2^2 , entonces la variable aleatoria F definido como:

$$F = \frac{\frac{S_1^2}{\sigma_1^2}}{\frac{S_2^2}{\sigma_2^2}}; \text{ se distribuye como F con } v_1=n_1-1 \text{ y } v_2=n_2-1 \text{ gl.}$$

Uso de la tabla de "F" de Fisher (probabilidad acumulativa)

La primera fila corresponde a los grados de libertad del numerador ($v_1=n_1-1$), la primera columna a los grados de libertad del denominador ($v_2=n_2-1$).

En el cruce de ambos grados de libertad se tiene 2 valores de F_0 , que corresponden a la probabilidad acumulativa, el superior $P(F < F_0) = 0.95$ y el inferior a $P(F < F_0) = 0.99$. En cada cruce de los grados de libertad se definen una distribución de F-Fisher.

Graf 29

Gl. n_2-1	0.95= $P(F < F_0)$ $v_1= n_1-1$ (numerador)			0.99= $P(F < F_0)$			
	1	2	...	10	11	12	...
1	161	200		242	243	244	
	4052	4999		6056	6082	6106	
2	18.51	19		19.39	19.4	19.41	
	98.49	99		99.4	99.41	99.42	
...							
11	4.84	3.98		2.86	2.82	2.79	
	9.65	7.2		4.54	4.46	4.4	
14	4.6	3.74		2.6	2.56	2.53	
	8.86	6.51		3.94	3.86	3.8	

EJERCICIOS

1. Para $n_1=11$ y $n_2=12$; $P(F < 2.86) = 0.95$.

$$P(F > 2.86) = 1 - P(F < 2.86) = 0.05$$

$$P(F < 4.54) = 0.99$$

$$P(F > 4.54) = 1 - P(F < 4.54) = 0.01$$

$$P(2.86 < F < 4.54) = 0.99 - 0.95 = 0.04$$

2. Para $v_1= 12$ y $v_2= 2$ hallar F_0 si $P(F > F_0) = 0.01$

$$\text{implica } P(F < F_0) = 1 - 0.01 = 0.99, \text{ de la tabla, resulta } F_0 = 99.42$$

3. Para $n_1=3$, $n_2=12$ hallar F_0 tal que $P(F < F_0) = 0.05$

Por la relación recíproca, resulta:

$$F_0 = F_{0.05(2,11)} = 1/F_{0.95(11,2)} = 1/19.40 = 0.0515$$

4. Para $n_1=8$ y $n_2=15$ hallar F_1 y F_0 tal que: $P(F_0 < F < F_1) = 0.98$ y $P(F > F_1) = 0.01$

Resulta que $P(F < F_1) = 1 - 0.01 = 0.99$, implica $F_1 = 4.28$

Si $P(F_0 < F < F_1) = 0.98$ y $P(F < F_1) = 0.99$ implica $P(F < F_0) = 0.01$

$$F_0 = 1/F_{0.99(14,7)} = 1/6.35 = 0.15748$$

5. De dos poblaciones normales con medias desconocidas y variancias $\sigma^2_1 = 4$, $\sigma^2_2 = 5$ respectivamente, se extraen dos muestras de tamaño $n_1 = 12$ y $n_2 = 11$. Si:

$P(S^2_2/S^2_1 > K) = 0.05$, hallar el valor de K

Solución: $F = (S^2_2/S^2_1)(\sigma^2_1/\sigma^2_2) = (4/5)(S^2_2/S^2_1)$ se distribuye como F con 10 y 11 grados de libertad.

entonces, $P(S^2_2/S^2_1 > K) = P((4/5) S^2_2/S^2_1 > (4/5) K)$

resulta $P(F_{(10,11)gl.} > (4/5) K) = 0.05$, por complemento

$P(F_{(10,11)gl.} < (4/5) K) = 0.95$. En la tabla se tiene

el valor 2.86 para el valor de $(4/5)K$ igualando:

$$(4/5)K = 2.86 \implies K = 3.575$$