II. Variables Aleatorias Discretas

M.A. Fernando Jesús Martínez Eissa

02H2019

ITAM
División Académica de Actuaría, Estadística y Matemáticas

Introducción

- Cuando se realiza un experimento, normalmente vamos a estar interesados en la distribución de los resultados más que en el resultado mismo.
- Por ejemplo, al realizar el experimento de aventar dos dados podemos estar más interesados en conocer la suma del valor de ambos dados más que en conocer qué valores tomo cada dado.
- Es decir , no interesaría conocer que la suma sea 5 en lugar de conocer si ésta se dio como (3,2)(2,3)(1,4)o (4,1)
- De igual forma al tirar una moneda nos interesaría más conocer el número total de soles más que la secuencia en que se fue presentando.
- A estas cantidades de interés, más formalmente, a estas funciones reales definidas sobre Ω las conocemos como variables aleatorias.
- Dado que los valores que toman las variables aleatorias están en función Ω, podemos asignarles probabilidades.

Variables Aleatorias

- Se dice que X es una Variable Aleatoria si $X: \Omega \to \mathbb{R}$, es decir una función cuyo dominio es el espacio muestral y su contradominio la recta real.

- A la imagen o rango de la función X se le denomina soporte de la variable aleatoria X.
- Con base en su soporte las v.a. se clasifican en:
  a) Discretas: Se asocian a un espacio muestral discreto, su soporte es finito o infinito numerable y provienen de un proceso de conteo.
  b) Continuas: Se asocian a un espacio muestral continuo, su soporte es infinito no numerable y provienen de un proceso de medición.

Ejemplo 1

Considere el experimento aleatorio en que se realizan tres volados consecutivos con una moneda no cargada. Sean X, la variable aleatoria que indica el número de águilas que aparecen durante el experimento; y Y, la variable aleatoria que toma el valor 1 si aparece al menos un sol y 0 en otro caso.

a) Defina el espacio muestral del experimento por extensión.

b) Para cada punto del espacio muestral determine su probabilidad y los valores que toman las variables aleatorias X y Y.

c) Calcule $P[X = 2]$ y $P[Y = 0]$

Ejemplo 1, (*Continuación*)

Sea X la variable que cuenta el número de águilas observadas, entonces X es una variable aleatoria que puede tomar los valores 0,1,2,3 con las siguientes probabilidades:

$$P[X=0]=P[(s,s,s)]=\left(\frac{1}{2}\right)^3=\frac{1}{8}$$

$$P[X=1]=P[(a,s,s),(s,a,s),(s,s,a)]=3\left(\frac{1}{2}\right)^3=\frac{3}{8}$$

$$P[X=2]=P[(s,a,a),(s,a,a),(aa,a,s)]=3\left(\frac{1}{2}\right)^3=\frac{3}{8}$$

$$P[X=3]=P[(a,a,a)]=\left(\frac{1}{2}\right)^3=\frac{1}{8}$$

Dado que X puede tomar los valores 0,1,2 y 3, tenemos:

$$1=\sum_{i=0}^{3}P[X=i]=P\left(\bigcup_{i=0}^{3}(X=i)\right)$$

Variables aleatorias

Ejemplo 2

Suponga que un banco presta servicios de banca múltiple de 9:00 a 17:00 hrs. El banco cuenta con 3 cajas y una ventanilla para realizar trámites administrativos de los cuentahabientes. En las operaciones de un día en particular de este banco se han identificado algunas variables aleatorias que se definen a continuación. Indique en cada caso si se trata de una variable discreta o continua e indique sus posibles valores.

a) X = Número de personas que visitan el banco durante el día.

b) Y = Caja en la que será atendida una persona que desea realizar un depósito.

c) Z = Número de trámites administrativos que realiza un cuentahabiente.

d) T = Tiempo que tarda en ser atendida una persona que llega a las 15:00 hrs.

e) S = Monto total de los depósitos que recibe el banco en este día.

f) Ri = El tiempo que tarda en atender a una persona el cajero i, i = 1, 2, 3.

Ejemplo 3

Se seleccionan 4 pelotas sin reemplazo de una urna que contiene 20 pelotas numeradas del 1 al 20. Si X corresponde al valor más grande de la muestra, entonces X es una variable aleatoria que puede tomar los valores 4,5,6,…,20.

Dado que cada una de las $\binom{20}{4}$ muestras tienen la misma probabilidad de ser seleccionadas, la probabilidad de que X tome cada uno de sus posibles valores es:

$$P[X = i] = \frac{\binom{i-1}{3}}{\binom{20}{4}} \; con\; i = 4,5,6, \dots 20$$

Entonces si se quiere calcular $P[X > 10]$ tenemos

$$\sum_{i=11}^{20} P[X = i] = \frac{\sum_{i=11}^{20} \binom{i-1}{3}}{\binom{20}{4}}$$

Ejemplo 3 (*Continuación*)

Otra forma alterna de verlo es:.

$$P[X > 10] = 1 - P[X \leq 10]$$

$$P[X > 10] = 1 - \frac{\binom{10}{4}}{\binom{20}{4}}$$

Ejemplo 4

Considere el experimento en el que se realizan volados independientes hasta que salga un sol o se realicen n volados, en donde p es la probabilidad de que salga sol.

Sea W es la variable aleatoria que cuenta el número de volados, entonces puede tomar los valores: $1,2,3,\ldots,n$ con las siguientes probabilidades:

$$P[W=1]=P[(s)]=p$$
$$P[W=2]=P[(a,s)]=(1-p)p$$
$$P[W=3]=P[(a,a,s)]=(1-p)^2p$$

$$P[W=n-1]=P[(a,a,\dots,a,s)]=(1-p)^{n-2}p$$
$$P[W=n]=P[(a,a,\dots,a),(a,a,\dots,a,s)]=(1-p)^{n-1}$$

Comprobando lo anterior:

$$1=\sum_{i=1}^{n}P[W=i]=P\left(\bigcup_{i=1}^{n-1}(w=i)\right)$$

Ejemplo 5

Considere a un agente de seguros que tiene dos clientes de edad avanzada, cada uno de los cuales tiene una póliza de seguro de vida que paga USD$100,000 en caso de muerte. Sea X el evento de que muera el más joven en el próximo año y Y el evento de que muera el más viejo el próximo año.

Considere que X y Y son independientes con probabilidades de muerte P(X)=0.05 y P(Y)=0.10 Si W denota el monto total de dinero que será pagado (en múltiplos de 100,000) a cualesquiera de los beneficiarios, entonces W es una v.a. que toma los valores 0,1 y 2, con las siguientes probabilidades:

$$P[W=0]=P[(X^cY^C)]=P[X^c]P[Y^C]=(.95)(.9)=.855$$
$$P[W=1]=P[(X^cY)]+P[(XY^C)]=(.95)(0.1)+(.05)(.9)=.140$$
$$P[W=2]=P[X]P[Y]=(.05)(.1)=.005$$

Ejemplo 6

Suponga que hay N tipos diferentes de cupones y que cada vez que uno adquiere un cupón es independiente de las selecciones anteriores y que los N cupones tienen la misma probabilidad de ser seleccionados.

Una variable aleatoria de interés puede ser T, el número de cupones que se necesita tener hasta tener al menos uno de cada tipo diferente, es decir $P[T = n]$.

Empecemos por calcular $P[T > n]$ fijamos n y definimos el evento A_j como el evento que no contiene el j-*ésimo* cupón dentro de los primeros n cupones. Entonces

$$P[T > n] = P\left(\bigcup_{j=1}^{n} A_j\right)$$

Ejemplo 6 (*Continuación*)

$$P[T > n] = P\left(\bigcup_{j=1}^{N} A_j\right)$$

$$= \sum_j P(A_j) - \sum_{j_1<j_2}\sum P(A_{j_1}A_{j_2}) + \cdots + (-1)^{k+1} \sum_{j_1<j_2<\cdots<j_k}\sum\sum P(A_{j_1}A_{j_2}\cdots A_{j_k})$$

$$+(-1)^{N+1}P(A_1A_2\cdots A_N)$$

A_j ocurrirá si para los primeros n cupones escogidos no se tiene el tipo j. Dado que cada uno de los cupones no será j con probabilidad $\left(\frac{N-1}{N}\right)$ y bajo el supuesto de independencia, tenemos:

$$P(A_j) = \left(\frac{N-1}{N}\right)^n$$

Ejemplo 6 (*Continuación*)

Asimismo, el evento $A_{j1}A_{j2}$ ocurrirá si ninguno de los primeros n cupones escogidos son tipo j_1 o j_2, nuevamente bajo el supuesto de independencia, tenemos:

$$P\left(A_{j_1}A_{j_2}\right) = \left(\frac{N-2}{N}\right)^n$$

De manera similar

$$P\left(A_{j_1}A_{j_2}\cdots A_{j_k}\right) = \left(\frac{N-k}{N}\right)^n$$

Para n>0

$$P[T > n] = N\left(\frac{N-1}{N}\right)^n - \binom{N}{2}\left(\frac{N-2}{N}\right)^n - \binom{N}{3}\left(\frac{N-3}{N}\right)^n - (-1)^N\binom{N}{N-1}\left(\frac{1}{N}\right)^n$$
$$= \sum_{i=1}^{N-1}\binom{N}{i}\left(\frac{N-i}{N}\right)^n(-1)^{i+1}$$

Ejemplo 6 (*Continuación*)

Entonces, la probabilidad de que T sea igual a n, la podemos obtener de:

$$P[T > n-1] = P[T = n] + P[T > n]$$

$$P[T = n] = P[T > n-1] - P[T > n]$$

Función de masa de probabilidad

Una variable aleatoria que puede tomar un valor que sea numerable se dice que es discreta. Para una variable aleatoria discreta X, se define la función de masa de probabilidades de f como:

$$f(x) = \begin{cases} P[X = x] > 0 \; si \; x = x_1, x_2, \ldots \\ 0 \qquad en \; otro \; caso \; (e.o.c) \end{cases}$$

La función de densidad $f(x)$ cumple:

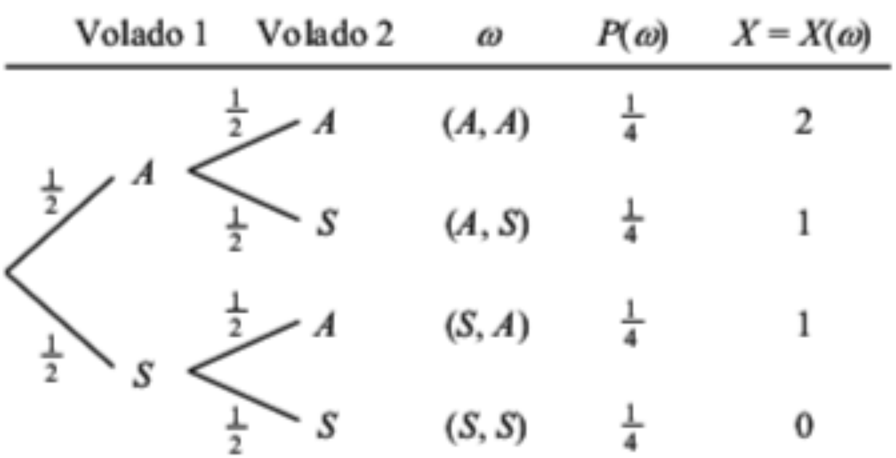
i. $f(x_i) \geq 0 \; para \; i = 1,2,3, \ldots$

ii. $\sum_{i=1}^{\infty} f(x_i) = 1$

Variables aleatorias

Ejemplo 7

Considere el experimento de lanzar dos volados con una moneda no cargada y X es el número de águilas, entonces $X\varepsilon\{0,1,2\}$, el espacio muestral es:



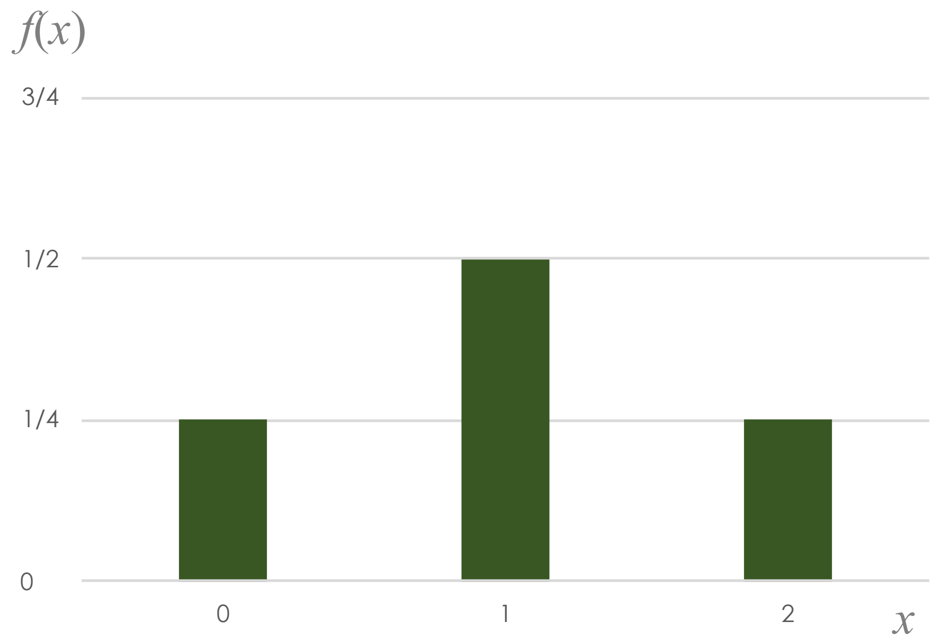
$$f_X(0) = P[X=0] = P((S,S)) = \frac{1}{4}$$

$$f_X(1) = P[X=1] = P((A,S)\cup(S,A)) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

$$f_X(2) = P[X=2] = P((A,A)) = \frac{1}{4}$$

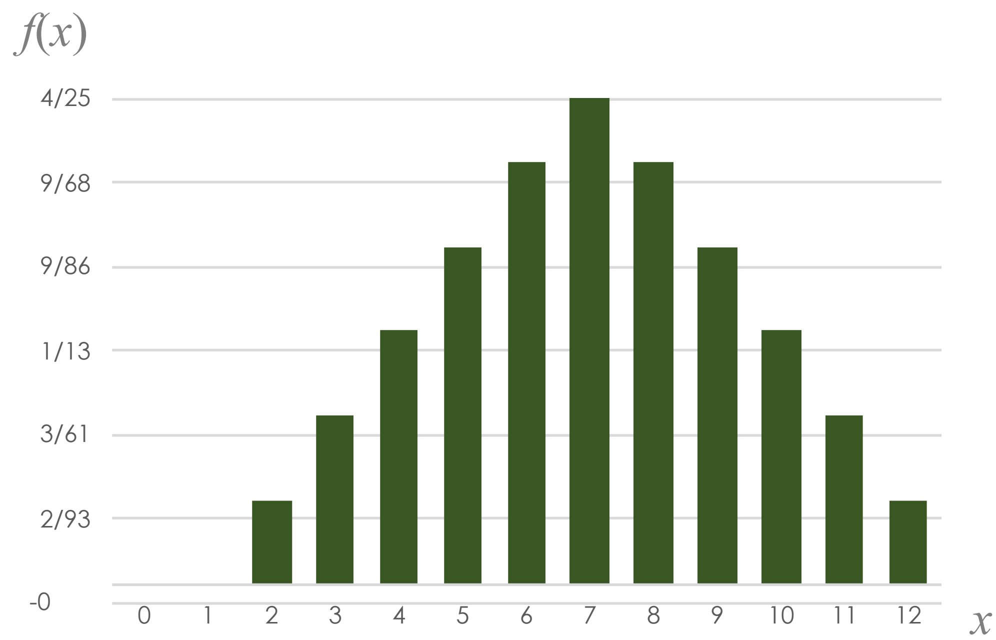
Función de masa de probabilidad

x	$f(x)$
0	1/4
1	1/2
2	1/4



Función de masa de probabilidad

x	$f(x)$
0	
1	
2	1/36
3	1/18
4	1/12
5	1/9
6	5/36
7	1/6
8	5/36
9	1/9
10	1/12
11	1/18
12	1/36



Función de masa de probabilidad

Ejemplo 8

Una persona tiene la siguiente estrategia de juego en un casino de Las Vegas. Apuesta 100 dólares a que la ruleta caerá en número rojo y si gana, se retira. Si pierde, entonces hace la misma apuesta pero con 200 dólares e independientemente del resultado se retira.

Suponga que la ruleta tiene 36 números, que la mitad de los números son rojos y que X es la variable aleatoria de la ganancia o pérdida de esta persona, determine la función de masa de probabilidad de X:

Función de masa de probabilidad

Ejemplo 9

Se tira un dado consecutivamente hasta que aparece un 6:

a) Si se define a la variable aleatoria Y como el número de tiros necesarios hasta que aparezca un 6, obtenga su función de masa de probabilidad.
b) ¿Cuál es la probabilidad de que a lo más el dado se tire 3 veces?
c) ¿Cuántos tiros se necesitan para asegurar que la probabilidad de obtener un 6 sea al menos 0.5?

Función de masa de probabilidad

Al conjunto de probabilidades asociadas al v.a. X se le llama distribución de probabilidades o distribución de la v.a. X.

Una forma común de representar a la distribución de probabilidades de una v.a. discreta, es a través de la representación tabular de su fmp.

Nota: Para el caso discreto el contradominio de la f.m.p. es [0, 1], ya que $f_X(x)$ es una probabilidad.

Función de masa de probabilidad

Proposición. Cálculo de probabilidades vía f.m.p.

Si X es variable aleatoria discreta y $A \subset \mathbf{R}$, entonces $P[X \in A] = \sum_{x \in A} f_X(x)$

Corolario. En particular:

Si X es variable aleatoria discreta $P[a \leq X \leq b] = \sum_{\{x : a \leq x \leq b\}} f_X(x)$

Función Indicadora

La función indicadora del conjunto A se define por $I_A(x)=\begin{cases}1 & si\;\; x\in A\\ 0 & si\;\; x\notin A\end{cases}$

Ejemplo 10

Si A es el conjunto de los dígitos, entonces:

$$I_A(x)=\begin{cases}1 & si\;\; x\in\{0,1,\ldots,9\}\\ 0 & e.o.c.\end{cases}$$

De modo que $I_A(4)=1$ per $I_A(11)=0$, $I_A(3.5)=0$ e $I_A(-3)=0$

Propiedades de la Función Indicadora

i) $I_A(x)=1-I_{A^c}(x)$

ii) $I_{A_1\cap A_2\cap\cdots\cap A_n}(x)=I_{A_1}(x)I_{A_2}(x)\cdots I_{A_n}(x)$

iii) $I_{A_1\cup A_2\cup\cdots\cup A_n}(x)=\max\{I_{A_1}(x),I_{A_2}(x),\cdots,I_{A_n}(x)\}$

iv) $I_A^2(x)=I_A(x)$ (idempotencia)

Función Indicadora

Las funciones indicadoras ayudan a simplificar la notación de las funciones utilizadas en Probabilidad ya que los distintos casos convierten en sumandos y el "cero en otro caso" queda implícito.

Por ejemplo, considerando nuevamente el lanzamiento de 2 volados con una moneda honesta, si X es el número de águilas obtenidas, entonces:

$$f_X(x)=\begin{cases} \frac{1}{4} & si \;\; x=0,2 \\ \frac{1}{2} & si \;\; x=1 \\ 0 & e.o.c. \end{cases} \quad \text{se expresaría como} \quad f_X(x)=\frac{1}{4}I_{\{0,2\}}(x)+\frac{1}{2}I_{\{1\}}(x).$$

Ejemplo 10

Reescribir la siguiente $f.m.p$ utilizando funciones indicadoras:

$$f_X(x) = \begin{cases} \dfrac{x}{10} & si\ x = 2,3 \\ \dfrac{7-x}{10} & si\ x = 4,5 \\ 0 & e.o.c. \end{cases}$$

Solución

$$f_X(x) = \frac{x}{10} I_{[2.3]}(x) + \frac{7-x}{10} I_{[4.5]}(x)$$

Modelos Paramétricos

Un modelo paramétrico es una función de masa de probabilidad que involucra constantes llamadas parámetros y cuyos valores determinan por completo el comportamiento probabilístico de un fenómeno aleatorio (distribución de probabilidad).

Al conjunto de valores posibles de los parámetros se le denomina espacio paramétrico. Por ejemplo, un modelo paramétrico utilizado frecuentemente en Probabilidad es la Distribución Geométrica.

$$X \sim Ge(\underset{\substack{\uparrow \\ \text{Parámetro}}}{p}) \quad \Rightarrow \quad f_X(x) = p(1-p)^{x-1} I_{\{1,2,\ldots\}}(x), \quad \underbrace{0 < p < 1}_{\text{Espacio paramétrico}}$$

Modelos Paramétricos

Las aplicaciones de las Distribución Geométrica aplican si un experimento aleatorio:

i) presenta un éxito con probabilidad p o fracaso con probabilidad $1\text{-}p$;

ii) el resultado de cada experimento es independiente de los demás;

iii) X es el número de experimentos necesarios hasta obtener el primer éxito, entonces X tiene una Distribución Geométrica con parámetro p.

Este modelo puede ser utilizado para modelar el número de lanzamientos de un dado necesarios hasta obtener un 6 (en este caso $p = \frac{1}{6}$), el número de hijos que tendrá una pareja que decide tener hijos hasta tener una niña (en este caso el parámetro p es la probabilidad de tener una niña en cierta población), etc.

Función de Acumulación Discreta

Ejemplo 11

La función de probabilidad de una variable aleatoria X está dada por:

$$p(i) = \frac{c\lambda^i}{i!}\ I_{[0,1,2,\dots)}(i)\ y\ \lambda > 0$$

Encontrar

a) $P[X = 0]$

$$\sum_{i=1}^{\infty} p(i) = 1 \Rightarrow \sum_{i=1}^{\infty} \frac{c\lambda^i}{i!} = 1 \Rightarrow c\sum_{i=1}^{\infty} \frac{\lambda^i}{i!} = 1 \Rightarrow ce^x = 1 \therefore c = e^{-x}$$

$$P[X = 0] = p(0) = \frac{c\lambda^0}{0!} = c = e^{-x}$$

b) $P[X > 2]$

$$P[X > 2] = 1 - P[X = 0] - P[X = 1] - P[X = 2]$$

$$= 1 - e^{-x} - \lambda e^{-\lambda} - \frac{\lambda^2 e^{-\lambda}}{2}$$

Ejemplo 12

La Distribución Poisson con parámetro λ, $\lambda > 0$, se caracteriza por la siguiente función de masa de probabilidad:

$$f_X(x) = \frac{\lambda^x e^{-\lambda}}{x!} I_{\{0,1,2,\dots\}}(x)$$

a) Verifique que, efectivamente, $f_X(X)$ es función de masa de probabilidad.

b) Calcule $P[X \geq 2]$

Función de Acumulación Discreta

Se define para una variable aleatoria discreta X, la función acumulación $F(x)$ como:

$$F(x) = P(X \leq x) \; para \; -\infty < x < \infty$$

Esta función especifica, para todos lo reales, la probabilidad de que la variable aleatoria sea menor o igual a x.

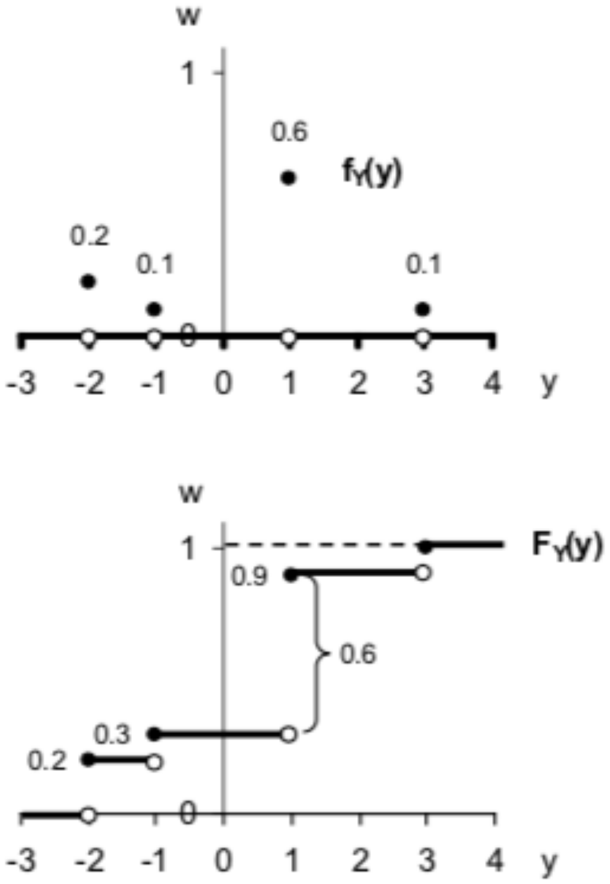
Considere ahora que $a \leq b$. Entonces se sigue de $[X \leq a] \epsilon [X \leq b]$, por lo tanto se sigue que $F(a) \leq F(b)$.

Nota: Tiene la misma definición para variables discretas y continuas.

Función de Acumulación Discreta

Sea X una v.a. Discreta con soporte $\{x_1, x_2, \ldots, x_n\}$ entonces:

$$F_X(x) = P[X \leq x] = \sum_x f_X(x) \;\; para\ toda\ x \varepsilon \mathbb{R}\ (función\ escalonada)$$



Función de Acumulación Discreta

1. F_X es una función no decreciente, es decir si $a < b$, entonces $F_X(a) \leq F_X\,(b)$

2. $\lim_{x \longrightarrow +\infty} F_X(x) = 1$

3. $\lim_{x \longrightarrow -\infty} F_X(x) = 0$

4. F_X es una función continua por la derecha. Es decir,

$$\lim_{h \longrightarrow 0} F_X(x+h) = F_X(x)\ para\ h > 0$$

Función de Acumulación Discreta

Ejemplo 13

Se lanza un dado y se define la v.a. X como el valor que resulte de dicho lanzamiento.

a) Determine la función de masa de probabilidad de X y grafíquela.

b) Construya la función de distribución acumulada de X y grafíquela

c) Compruebe que $F_X(x)$ satisface las condiciones de una función de acumulación.

d) A partir de la función de distribución acumulada de X calcule:

$P[X < 3]; P[X \leq 2]; P[X > 2]; P[3 < X < 6]; P[3 \leq X < 6]$

Función de Acumulación Discreta

Proposición

Si X es una v.a. Discreta con $a < b \implies P[a \leq X \leq b] = F_X(b) - F_X(a)$

Construcción de la $f.m.p$ a partir de la $F_X(x)$.

$$f_X(x) = F_X(x) - \lim_{h \longrightarrow 0} F_X(a-h) = F_X(x) - \lim_{x \longrightarrow a^-} F_X(x) \; para \; h > 0$$

Función de Acumulación Discreta

Ejemplo 14

El número de solicitudes aperturas de crédito de un banco es una v.a. X con función de distribución acumulada:

$$F_X(x) = \begin{cases} 0 & si\ x < 0 \\ 0.1 & si\ 0 \leq x < 1 \\ 0.3 & si\ 1 \leq x < 2 \\ 0.7 & si\ 2 \leq x < 4 \\ 1 & si\ x \geq 4 \end{cases}$$

a) Encuentre la función de masa de probabilidad de la v.a. X

b) Si en un día en particular se ha recibido una solicitud ¿cuál es la probabilidad de que se reciban más de 3 solicitudes?.

c) Calcule la probabilidad de que se reciban 3 solicitudes en un día en particular.

Esperanza de una v.a. discreta

Uno de los conceptos más importantes en la teoría de probabilidad es el de valor esperado de una variable aleatoria.

Si X es una variable aleatoria discreta con función de probabilidad $f(x)$, entonces su esperanza o valor esperado $E[X]$ es:

$$E[X] = \sum_{\forall x} x f(x)$$

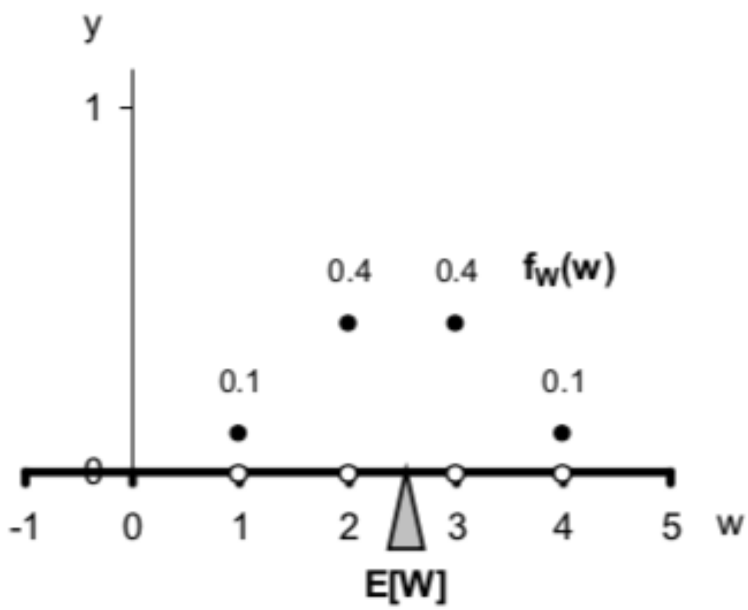
Es decir, corresponde al promedio ponderado de todos los posibles valores que puede tomar X multiplicados por su probabilidad.

Geométricamente representa el centro de masa de la función, es decir en donde se equilibran.

Esperanza de una v.a. discreta

Si $E[X]$ existe y $a \leq X \leq b \Rightarrow a \leq E[X] \leq b \Rightarrow$.

Si X es una variable aleatoria discreta, $E[X]$ puede tomar un valor acotado por su soporte pero que no necesariamente pertenece al soporte, como se observa a continuación:



Esperanza de una v.a. discreta

- Considere que se tienen dos posibles valores: 0 y 1. Si $f(0)=1/3$ y $f(1)=2/3$ entonces:

$$E[X] = (0)\left(\frac{1}{3}\right) + (1)\left(\frac{2}{3}\right) = \frac{2}{3}$$

Es decir, corresponde al promedio ponderado de todos los posibles valores que puede tomar X multiplicados por su probabilidad.

- Sea X una v.a. que puede tomar los valores x_1, x_2, x_3,…, x_n con probabilidades $p(x_1)$, $p(x_2)$,…, $p(x_n)$ respectivamente, y considere que X representa el número de éxitos en un juego de azar. Esto es, con probabilidad $p(x_i)$ se ganan x_i unidades i=1,2,3,…,n. Bajo el supuesto de interpretación de frecuencias, si se juega de manera continua, se estima que promedio de juegos ganados sea:

$$\sum_{i=1}^{n} x_i p(x_i) = E[X]$$

Esperanza de una v.a. discreta

Ejemplo 15

Encontrar la $E[X]$ si x corresponde al resultado de tirar un dado sin cargar.

Solución

$$E[X] = (1)\left(\frac{1}{6}\right) + (2)\left(\frac{1}{6}\right) + (3)\left(\frac{1}{6}\right) + (4)\left(\frac{1}{6}\right) + (5)\left(\frac{1}{6}\right) + (6)\left(\frac{1}{6}\right) = \frac{21}{6} = \frac{7}{2}$$

Esperanza de una v.a. discreta

Ejemplo 16

Se dice que I es la función indicadora para un evento A si: .

$$I = \begin{cases} 0 & si\ no\ ocurre\ A \\ 1 & si\ ocurre\ A \end{cases}$$

Encontrar la $E[I]$

Solución

$$E[I] = (1)(P(A)) + (0)(1 - P(A)) = P(A)$$

Esperanza de una v.a. discreta

Ejemplo 17

Suponga que X es la v.a. que mide el número de veces que llueve por día en cierta región y que su función de masa de probabilidad está dada por:

X	0	1	2
$P[X=x]$	0.1	0.6	0.3

a) Determine la expresión de $f_X(x)$ y construya su gráfica

b) ¿Cuántas veces se espera que llueva por día en esta región?. Grafique este valor en la gráfica del inciso a).

Esperanza de una v.a. discreta

Ejemplo 18

La función de masa de probabilidad de la variable aleatoria W está definida por:

$$f_W(w) = \begin{cases} \dfrac{2w}{n(n+1)} & si\ w = 1,2,\dots,n \quad ,n \in \mathbb{Z}^+ \\ 0 & e.o.c \end{cases}$$

Calcule la esperanza de W.

Esperanza de una v.a. discreta

Ejemplo 19

- En 100 ITAMitas dijeron se responden dos preguntas, las cuales se pueden ordenar en el orden que prefiera el concursante. Si se decide por contestar la *i-ésima* pregunta primero y ésta es correcta, pasa a la *j-ésima*, En caso contrario no pasa.

  El concursante recibe X_i dólares por contestar la *i-ésima* pregunta correcta, por ejemplo si contesta ambas preguntas correctas gana x_1+x_2.

  Si la probabilidad de que el concursante conozca la *i-ésima* respuesta es P_i i=1,2, ¿qué pregunta debe contestar primero para maximizar sus ganancias esperadas?

Esperanza de una v.a. discreta

Ejemplo 19 (*Continuación*)

Solución

Si decide contestar primero la primer pregunta, tenemos:

$$\begin{array}{cc} 0 & con\ probabilidad\ P_1 \\ X_1 & con\ probabilidad\ P_1(1-P_2) \\ X_1+X_2 & con\ probabilidad\ P_1P_2 \end{array}$$

En este caso sus ganancias esperadas son:

$$E(*) = (0)(P_1) + (X_1)(P_1(1-P_2)) + (X_1+X_2)(P_1P_2)$$

Si contesta primero la segunda pregunta, entonces:

$$E(*) = (0)(P_2) + (X_2)(P_2(1-P_1)) + (X_1+X_2)(P_1P_2)$$

Entonces para que su mejor opción sea contestar primero 1:

$$(X_1)(P_1(1-P_2)) \geq (X_2)\big(P_2(1-P_1)\big)\ i.e. \frac{X_1P_1}{1-P_1} \geq \frac{X_2P_2}{1-P_2}$$

Si tiene 60% de chance de contestar correctamente la 1 y gana 200 y 80% de la 2 ganando 100, tenemos:

$$400 = \frac{(100).8)}{0.2} \geq \frac{(200)(.6)}{0.4} = 300$$

Valor Esperado (*discreto*)

El concepto de esperanza se puede generalizar a través del valor esperado.

Considere que se tiene una v.a. X y su correspondiente función de densidad $g(x)$ y que se busca calcular su valor esperado.

Dado que $g(x)$ es en sí un variable aleatoria discreta, entonces buscamos:

$$E[g(x)] = \sum_{\forall x} g(x)f(x)$$

Ejemplos de transformaciones pueden ser: $g(x) = x^2; g(x) = ln(x); g(x) = (x-c)^2$. El valor esperado es la generalización de la esperanza ya que $Y = g(X)$ es una nueva v.a.

Valor Esperado (*discreto*)

Propiedades del Valor Esperado de una función discreta

- $E[c] = c$ con c constante
- $E[cX] = cE[X]$
- $E[g(x) + h(x)] = E[g(x)] + E[h(x)]$

Si a y b son constantes:

$$E[aX + b] = aE[X] + b$$

El valor esperado de la variable aleatoria X, $E[X]$ también llamado la media o primer momento de X.

Valor Esperado (*discreto*)

Otra propiedad importante del valor esperado es que el valor esperado para la suma de variables aleatorias corresponde a la suma de las esperanzas.

Para las variables aleatorias $X_1, X_2, X_3, \ldots, X_n$

$$E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i]$$

Valor Esperado (*discreto*)

Por la forma en que se define, al cálculo del valor esperado, también se le conoce como ley del estadístico inconsciente (LEI) pues para calcular el valor esperado de la nueva variable aleatoria Y=g(x) basta aplicar la transformación $g(\cdot)$ a X (lado izquierdo) y a sus posibles valores (lado derecho).

Cuando $E[g(x)] = \infty \; o \; \infty - \infty$ se dice que el valor esperado no existe. El valor esperado solo existe si $E[g(x)] < \infty$, es decir, cuando converge

Valor Esperado (*discreto*)

Aunque X es la v.a. Para la cual se conoce $f_X(x)$, la definición de valor esperado permite calcular $E[Y] = E[g(x)]$ sin necesidad de conocer $f_Y(y)$.

Por ejemplo, si X denota la cantidad producida de algún bien con $f_X(x)$ conocida y $C(X)$ es su función de costos, entonces el costo esperado es $E[C(x)]$ del mismo modo es posible encontrar la utilidad esperada o el ingreso esperado, etc.

Si $Y = g(x)$ y $f_X(X)$ es conocida pero $f_Y(Y)$ no, es posible hacer el cálculo de probabilidades asociadas a la v.a. Y a partir de $f_X(X)$ "despejando adecuadamente" a X de la ecuación $Y = g(x)$

Valor Esperado (*discreto*)

Ejemplo 20

Un puesto de periódicos vende un diario francés poco demandado. Si W es el número de diarios vendidos en un día, su función de masa de probabilidad está dada por:

W	5	6	7	8
$P[W=w]$	0.2	0.4	0.3	0.1

Suponga que la función de utilidad (en MXN$) para el puesto de periódicos por la venta del diario francés está dada por

$$U(W) = 60\sqrt{w-5}$$

a) Calcular la utilidad esperada.

b) Calcular la probabilidad de que la utilidad sea mayor a MXN$ 60

Valor Esperado (*discreto*)

Ejemplo 21

Sea X la v.a. que toma los valores de -1,0 y 1 con las siguientes probabilidades:

$P[X=-1]$	0.2
$P[X=0]$	0.5
$P[X=1]$	0.3

Calcular $E[X^2]$

Solución

$$E[X^2] = (-1^2)(0.2) + (0^2)(0.5) + (1^2)(0.3) = 0.5$$

Valor Esperado (*discreto*)

Ejemplo 22

Un producto que se vende por temporada bajo una utilidad neta de b dólares por unidad vendida y una pérdida neta de l dólares por cada unidad no vendida al final de temporada.

El número de unidades del producto que son ordenadas en una determinada tienda departamental durante cualquier temporada es una v.a. que tiene una función de densidad de probabilidades $f(i)$, $i \geq 0$. Si la tienda debe mantener un stock de su producto. Calcule el número de unidades que debe tener la tienda para maximizar su utilidad esperada.

Ejemplo 22 (*Continuación*)

Sea X la v.a. Del número de unidades ordenadas. Si las unidades se almacenan, entonces la utilidad ($U(x)$) es:

$$U(s) = \begin{matrix} bX - (s - X)l & si\ X \leq s \\ sb & si\ X > s \end{matrix}$$

Entonces el valor esperado de la utilidad es:

$$E\big(U(s)\big) = \textstyle\sum_{i=1}^{5}[bi - (s - i)l]f(i) + \sum_{i=s+1}^{\infty} sbf(i)$$

$$= (b + l)\sum_{i=1}^{5} if(i) - sl\sum_{i=0}^{5} f(i) + sb\left[1 - \sum_{i=0}^{5} f(i)\right]$$

$$= (b + l)\sum_{i=1}^{5} if(i) - (b + l)s\sum_{i=0}^{5} f(i) + sb$$

$$= sb + (b + l)\sum_{i=1}^{5}(i - s)f(i)$$

Valor Esperado (*discreto*)

Ejemplo 22 (*Continuación*)

Para determinar el valor óptimo de s, analicemos que pasa con la utilidad si s se incrementa por una unidad. Por sustitución podemos observar que la utilidad esperada es:

$$E[U(s+1)] = b(s+1) + (b+l)\sum_{i=0}^{s+1}(i-s-1)f(i)$$

$$= b(s+1) + (b+l)\sum_{i=0}^{s}(i-s-1)f(i)$$

Entonces

$$E[U(s+1)] - E[U(s)] = b - (b+l)\sum_{i=0}^{s} f(i)$$

Por lo tanto tener en inventario s+1 es mejor que tener s siempre que $\sum_{i=0}^{s} f(i) < \frac{b}{b+l}$

A la cantidad $E[X^n], n\epsilon\, \mathbb{Z}^+$ se le llama el *n-ésimo* momento de X

$$\mu'_n = E[X^n] = \sum_{x, f(x)>0} x^n f(x)$$

y su *n-ésimo* momento central se define como:

$$\mu_n = E[(X - E[X])^n]$$

El cálculo de los momentos y de los momentos centrales se realiza aplicando la ley del estadístico inconsciente considerando las transformaciones $g(X) = X^n \; y \; g(X) = (X - E[X])^n$

Nótese que cuando n=1 $\mu'_1 = E[X^1] = \sum_{x,f(x)>0} x^1 f(x) = E[X]$, es decir el primer momento de una variable aleatoria es su esperanza.

Corolario

Si X es una v.a. y $E[X] = \mu$, entonces:

I. $\mu_2 = \mu_2' - (\mu_1')^2, i.e. E[(X-\mu)^2] = E[X^2] - (E[X])^2$

II. $\mu_3 = \mu_3' - 3\mu_1'\mu_2' + 2(\mu_1')^3, i.e.$
$E[(X-\mu)^3] = E[X^3] - 3E[X]E[X^2] + 2(E[X])^3$

III. $\mu_4 = \mu_4' - 4\mu_1'\mu_3' + 6(\mu_1')^2\mu_2' - 3(\mu_1')^4 i.e.$
$E[(X-\mu)^4] = E[X^4] - 4E[X]E[X^3] + 6(E[X])^2E[X^2] - 3(E[X])^4$

Varianza (*discreto*)

Si X es una v.a. Con media μ, entonces su varianza denotada por Var(X), se define como:

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

Entonces

$$Var[X] = E[X^2 - 2X\mu + \mu^2]$$

$$Var[X] = E[X^2] - 2\mu E[X] + \mu^2$$

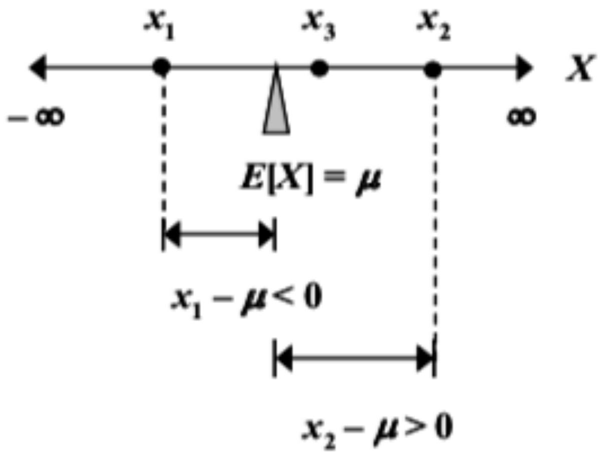
$$Var[X] = E[X^2] - 2E[X]E[X] + E^2[X]$$

$$Var[X] = E[X^2] - E^2[X]$$

Nota: observe que la varianza corresponde al segundo momento central.

Varianza (*discreto*)

La varianza es una medida de dispersión de la distribución de la v.a. X respecto a su esperanza como se muestra a continuación:



Observe

$X - E[X]$ es una distancia aleatoria con signo (+ o -)

$g(X) = (X - E[X])^2 \geq 0$ es una distancia aleatoria positiva o cero, pero con unidades al cuadrado.

Proposición $Var[X] \geq 0$

Ejemplo 23

Calcule la varianza de X, si X es una v.a. que corresponde a los posibles valores de tirar un dado no cargado.

Solución

$$E[X] = (1)\left(\frac{1}{6}\right) + (2)\left(\frac{1}{6}\right) + (3)\left(\frac{1}{6}\right) + (4)\left(\frac{1}{6}\right) + (5)\left(\frac{1}{6}\right) + (6)\left(\frac{1}{6}\right) = \frac{21}{6}$$
$$= \frac{7}{2}$$

$$E[X^2] = (1^2)\left(\frac{1}{6}\right) + (2^2)\left(\frac{1}{6}\right) + (3^2)\left(\frac{1}{6}\right) + (4^2)\left(\frac{1}{6}\right) + (5^2)\left(\frac{1}{6}\right) + (6^2)\left(\frac{1}{6}\right)$$
$$= \frac{91}{6}$$

$$\therefore Var[X] = \frac{91}{6} - \frac{49}{4} = \frac{182 - 147}{12} = \frac{35}{12}$$

Varianza (*discreto*)

Una identidad útil para a y b que son constantes:

$$Var[aX+b] = a^2Var[X]$$

$$Var[aX+b] = E[(aX+b)^2] - E^2[(aX+b)]$$

$$Var[aX+b] = E[a^2X^2 + 2abX + b^2] - E[(aX+b)]E[(aX+b)]$$

$$Var[aX+b] = a^2E[X^2] + 2abE[X] + b^2 - (a^2E^2[X] + 2abE[X] + b^2)$$

$$Var[aX+b] = a^2E[X^2] - a^2E^2[X]$$

$$Var[aX+b] = a^2\{E[X^2] - E^2[X]\}$$

$$Var[aX+b] = a^2Var\{X\}$$

Varianza (*discreto*)

Propiedades de la Varianza de una función discreta

- $Var[c] = 0$ con c constante
- $Var[cX] = c^2 Var[X]$
- $Var[X + c] = Var[X]$

Para resolver el problema de las unidades al cuadrado de la varianza se define la desviación estándar como:

$$\sigma_X = \sigma[X] = \sqrt{Var[X]}$$

La desviación estándar es una medida de dispersión de la distribución de la v.a. X respecto a su esperanza, pero medida en las mismas unidades que la variable.

Varianza (*discreto*)

Ejemplo 24

Encuentre el número total de éxitos esperados que resultan de n ensayos cuando el *i-ésimo* ensayo tiene una probabilidad de éxito p. i=1,2,3,…,n. Encuentre el valor esperado y la varianza

Solución

Sea

$$X_i = \begin{cases} 1, & si\ el\ i-\acute{e}simo\ ensayo\ es\ \acute{e}xito \\ 0, & si\ el\ i-\acute{e}simo\ ensayo\ es\ fracaso \end{cases}$$

Sea $Z = \sum_{i=1}^{n} X_i$

Entonces

$$E[Z] = E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i] = p + p + p + \cdots + p = np$$

Ejemplo 24, (*Continuación*)

Encuentre la varianza de Z

Solución

Sea $V[Z]=E[Z^2]-E^2[Z]$

$$E[Z^2] = E\left[\sum_{i=1}^{n} X_i^2\right] = \sum_{i=1}^{n} E\left[X_i^2\right] = \sum_{k\geq 0}^{n} k^2 \binom{n}{k} p^k (1-p)^{n-k}$$

Se conoce que $k\binom{n}{k} = k\frac{n!}{(n-k)!k!} = k\frac{n(n-1)!}{(n-1-k+1)!k(k-1)!} = n\binom{n-1}{k-1}$

$$= \sum_{k\geq 1}^{n} kn\binom{n-1}{k-1} p^k (1-p)^{n-k} = np\sum_{k\geq 1}^{n} k\binom{n-1}{k-1} p^{k-1}(1-p)^{n-1-k+1}$$

Sea j=k-1 y m=n-1

$$= np\sum_{k\geq 1}^{n} j+1\binom{m}{j} p^j (1-p)^{m-j} = np\left[\sum_{k\geq 1}^{n} j\binom{m}{j} p^j(1-p)^{m-j} + \sum_{k\geq 1}^{n}\binom{m}{j} p^j (1-p)^{m-j}\right]$$

Ejemplo 24, (*Continuación*)

$$= np\left[\sum_{k\geq 1}^{n} j\binom{m}{j}p^{j}(1-p)^{m-j}+\sum_{k\geq 1}^{n}\binom{m}{j}p^{j}(1-p)^{m-j}\right]$$

$$E[Z^2] = np[(n-1)p+1]$$

$$V[Z] = E[Z^2] - E^2[Z] = np[(n-1)p+1] - (np)^2 = np[np-p+1] - (np)^2$$

$$V[Z] = (np)^2 - np^2 + np - (np)^2 = np(1-p)$$

Esperanza y Varianza (*discreto*)

Ejemplo 25

Un inversionista realiza dos inversiones. La inversión 1 tendrá una ganancia de \$1,000 con probabilidad 0.6 o una pérdida de \$400 con probabilidad de 0.4. La inversión 2 tendrá una ganancia de \$2,000 con probabilidad de 0.5 o una pérdida de \$500 con probabilidad de 0.5.

a. Grafique y compare las funciones de masa de probabilidad de la ganancia/pérdida de estas inversiones
b. Calcule la esperanza y desviación estándar de cada inversión
c. Si usted fuera el analista financiero y considera solo la ganancia/pérdida ¿qué inversión recomendaría?
d. En Finanzas, la desviación estándar de una inversión es considerada como su riesgo financiero. Si usted fuera analista financiero y sólo considera el riesgo de las inversiones ¿qué inversión recomendaría?

Medidas Poblacionales

Medidas poblacionales

Medidas que caracterizan y cuantifican aspectos relevantes de una distribución

Grupo	Descripción	Medidas
Tendencia central	Ubican la parte central de una distribución de probabilidades	• Media • Mediana • Moda
Localización	Segmentan el soporte de una distribución mediante valores que cuantifican la probabilidad acumulada	• Percentiles • Cuartiles • Cuantiles
Dispersión	Miden el grado de variabilidad de la distribución, es decir, su grado de concentración o dispersión	• Varianza • Desviación estándar • Coeficiente de variación • Amplitud intercuartílica • Desviación media
Simetría y forma	Permiten inferir el perfil de la distribución de probabilidades	• Coeficiente de asimetría • Coeficiente de curtosis

v.a.'s Discretas | v.a.'s Continuas

Medidas estudiadas previamente

Media

Media aritmética para datos no agrupados: Corresponde a la suma de todos los valores divida entre el total.

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N}, \text{ Poblacional}$$

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}, \text{ Muestral}$$

Ejemplo 25:

Calcular la edad promedio del grupo.

Alumno	edad	Alumno	edad
1	20	13	22
2	22	14	23
3	20	15	24
4	22	16	20
5	22	17	20
6	22	18	21
7	22	19	20
8	21	20	21
9	22	21	20
10	21	22	21
11	25	23	23
12	19	24	21

$$\mu = \frac{20+22+20+\cdots+21}{24} = 21 \text{ años } 5 \text{ meses}$$

Media Geométrica

Media geomética para datos no agrupados: Corresponde a la raíz n-ésisma de producto de los valores

$$\mu_{GEO} = \sqrt[N]{\prod_{i=1}^{N} x_i}, \text{ Poblacional}$$

$$\overline{x_{GEO}} = \sqrt[n]{\prod_{i=1}^{n} x_i}, \text{ Muestral}$$

Ejemplo 26:
Calcular la media geométrica del grupo.

Alumno	edad	Alumno	edad
1	20	13	22
2	22	14	23
3	20	15	24
4	22	16	20
5	22	17	20
6	22	18	21
7	22	19	20
8	21	20	21
9	22	21	20
10	21	22	21
11	25	23	23
12	19	24	21

$$\mu = \sqrt[24]{20 * 22 * 20 \cdots 23 * 21} = \text{21 años 4 meses}$$

Media (datos agrupados)

Media ponderada para datos agrupados: Cuando los datos están agrupados, se les asigna una ponderación con base en la importancia relativa.

$$\mu = \bar{x} = \frac{\sum_{i=1}^{N} w_i x_i}{\sum_{i=1}^{N} w_i},$$

$$\mu = \bar{x} = \frac{Marca\ de\ clase * Frecuencia\ Absoluta}{Total} = Marca\ de\ clase * Frecuencia\ Relativa$$

Ejemplo 27:

Calcular el precio promedio ponderado del siguiente inventario:

Producto	Existencias	Precio
Camisas	4,000	MX$ 575.00
Pantalones	3,500	MX$ 585.00
Zapatos	2,000	MX$ 595.00
Cinturones	500	MX$ 600.00

$$\mu = \frac{(4000)(MX\$575)+(3500)(MX\$585)+(2000)(MX\$595)+500(MX\$600)}{10,000} = \text{MX\$583.75}$$

Media (*datos agrupados*)

Ejemplo 27:

Intervalos	Marca de Clase	Frecuencias ABSOLUTA (fi)	RELATIVA (pi)
(75000,125000]	100000	3	0.08
(125000,175000]	150000	8	0.20
(175000,225000]	200000	10	0.25
(225000,275000]	250000	8	0.20
(275000,325000]	300000	5	0.13
(325000, 375000]	350000	6	0.15
		40	**1**

$$\mu = \frac{100000(3)+150000(8)+200000(10)+250000(8)+300000(5)+350000(6)}{40} = 227{,}500$$

Mediana

Mediana para datos no agrupados: Corresponde al valor del elemento central del conjunto. Es el valor de X por debajo (y por arriba) del cual se acumula el 50% de la distribución de probabilidades.

Para el caso discreto, la mediana es el valor mínimo de x tal que:

$$\frac{1}{2} \leq F_X(x)$$

Ventajas de la mediana sobre la media:

i. La mediana siempre existe
ii. La mediana es mejor medida de tendencia central en distribuciones no simétricas

Distribución Simétrica

Para el caso discreto, Si X es una v.a. con $f.m.p$ y $c \in \mathbb{R}$, se dice que X tiene una distribución simétrica respecto a c si:

$$f_X(c-x) = f_X(c+x)$$

Ejemplo 28:
Calcular la mediana del grupo.

Alumno	edad	Alumno	edad
1	20	13	22
2	22	14	23
3	20	15	24
4	22	16	20
5	22	17	20
6	22	18	21
7	22	19	20
8	21	20	21
9	22	21	20
10	21	22	21
11	25	23	23
12	19	24	21

Alumno	edad	Alumno	edad
1	19	13	$\boxed{21}$
2	20	14	22
3	20	15	22
4	20	16	22
5	20	17	22
6	20	18	22
7	20	19	22
8	21	20	22
9	21	21	23
10	21	22	23
11	21	23	24
12	$\boxed{21}$	24	25

Notas

1. Cualitativos
2. Al ser un número par, se considera el promedio de los dos números centrales.

Me=21

Mediana (*datos agrupados*)

Mediana para datos agrupados: La mediana se encuentra en el intervalo donde la frecuencia acumulada llega hasta la mitad de la suma de las frecuencias absolutas. $i.e.$ Se busca el intervalo en donde está $\frac{N}{2}$ y posteriormente se calcula como:

$$Me = L_{inf} + \frac{\frac{N}{2} - F_{MED-1}}{f_i} t_i$$

donde:

$L_{inf} = Límite\ inferior\ del\ intervalo\ en\ donde\ se\ encuentra\ la\ mediana$

$\frac{N}{2} = Semisuma\ de\ frecuencias\ absolutas$

$F_{MED-1} = Frecuencia\ acumulada\ anterior\ a\ la\ clase\ mediana$

$f_i = Frecuencia\ absoluta\ del\ intervalo\ mediano$

$t_i = amplitud\ de\ los\ intervalos$

Mediana (*datos agrupados*)

Ejemplo 29

Intervalos	Marca de Clase	Frecuencias ABSOLUTA (fi)	RELATIVA (pi)	Frecuencias Acumuladas ABSOLUTA (Fi)	RELATIVA (Pi)
(75000,125000]	100000	3	0.08	3	0.08
(125000,175000]	150000	8	0.20	11.00	0.28
(175000,225000]	200000	10	0.25	21.00	0.53
(225000,275000]	250000	8	0.20	29.00	0.73
(275000,325000]	300000	5	0.13	34.00	0.85
(325000, 375000]	350000	6	0.15	40.00	1.00
		40	**1**		

1. $\frac{N}{2} = \frac{40}{2}$= 20
2. Buscar intervalo donde la Frecuencia acumulada contenga 20
3. 175000+ $\frac{20-11}{10}$ 50000=220000

$L_{inf} = Límite\ inferior\ del\ intervalo\ en\ donde\ se\ encuentra\ la\ mediana\ 175000$

$\frac{N}{2} = Semisuma\ de\ frecuencias\ absolutas.$ 20

$F_{MED-1} = Frecuencia\ acumulada\ anterior\ a\ la\ clase\ mediana, 11$

$f_i = Frecuencia\ absoluta\ del\ intervalo\ mediano, 10$

$t_i = amplitud\ de\ los\ intervalos,$ 50000

Moda

Moda para datos no agrupados: Corresponde al valor que ocurre con mayor frecuencia, es decir, el valor más probable.

La Moda (Mo) de una v.a. X, es el valor del soporte de X que maximiza a $f_X(x)$.

Ejemplo 30:
Calcular la moda del grupo.

Alumno	edad	Alumno	edad
1	20	13	22
2	22	14	23
3	20	15	24
4	22	16	20
5	22	17	20
6	22	18	21
7	22	19	20
8	21	20	21
9	22	21	20
10	21	22	21
11	25	23	23
12	19	24	21

Alumno	edad	Alumno	edad
1	19	13	21
2	20	14	22
3	20	15	22
4	20	16	22
5	20	17	22
6	20	18	22
7	20	19	22
8	21	20	22
9	21	21	23
10	21	22	23
11	21	23	24
12	21	24	25

Moda=22

Moda (*datos agrupados*)

Moda para datos agrupados: Es el valor que representa la mayor frecuencia absoluta se calcula como:

$$Moda = L_{inf} + \frac{f_i - f_{i-1}}{(f_i - f_{i-1})(f_i - f_{i+1})} t_i$$

donde:

$L_{inf} = Límite\ inferior\ del\ intervalo\ modal\ (mayor\ frecuencia\ absoluta)$

$f_i = Frecuencias\ absoluta\ del\ intervalo\ modal$

$f_{i-1} = Frecuencias\ absoluta\ del\ intervalo\ anterior\ al\ modal$

$f_{i+1} = Frecuencias\ absoluta\ del\ intervalo\ posterior\ al\ modal$

$t_i = amplitud\ de\ los\ intervalos$

Moda (*datos agrupados*)

Ejemplo 31

Intervalos	Marca de Clase	Frecuencias ABSOLUTA (fi)	Frecuencias RELATIVA (pi)	Frecuencias Acumuladas ABSOLUTA (Fi)	Frecuencias Acumuladas RELATIVA (Pi)
(75000,125000]	100000	3	0.08	3	0.08
(125000,175000]	150000	8	0.20	11.00	0.28
(175000,225000]	200000	10	0.25	21.00	0.53
(225000,275000]	250000	8	0.20	29.00	0.73
(275000,325000]	300000	5	0.13	34.00	0.85
(325000, 375000]	350000	6	0.15	40.00	1.00
		40	**1**		

1. Buscar intervalo donde la Frecuencia absoluta es mayor
2. Sustituya en la Formula

$$Moda = 175000 + \frac{10-8}{(10-8)(10-8)} 50000 = 200000$$

$L_{inf} = Límite\ inferior\ del\ intervalo\ modal\ (mayor\ frecuencia\ absoluta)$ 175000

$f_i = Frecuencias\ absoluta\ del\ intervalo\ modal$ 10

$f_{i-1} = Frecuencias\ absoluta\ del\ intervalo\ anterior\ al\ modal$ 8

$f_{i+1} = Frecuencias\ absoluta\ del\ intervalo\ posterior\ al\ modal$ 8

$t_i = amplitud\ de\ los\ intervalos$ 50000

Ejemplo 32

Suponga que Y es el número de veces que falla la impresora diariamente en una oficina. La probabilidad de que la impresora no falle es 0.1, pero la probabilidad de que falle 1,2, o 3 veces es de 0.3, 0.4 y 0.2 respectivamente. Calcule la media, mediana y moda de la v.a. Y

Medidas de posición

Los percentiles o medidas de posición. Una vez que se han ordenado los datos de menor a mayor, los percentiles indican el valor por debajo del cual se encuentra un porcentaje dado de observaciones de un grupo.

Por ejemplo, el percentil 10° es el valor bajo el cual se encuentra el 10% de las observaciones.

Cuartiles: Son valores que dividen al conjunto de observaciones ordenadas en cuatro partes. i.e. Son abscisas que tienen por debajo al 25%, 50% (mediana) y 75% de los valores de la distribución de frecuencias. Los muestrales se denotan q y los poblacionales Q.

Cuartil inferior: o primer cuartil tiene por debajo de su valor al 25% de los valores de la distribución de frecuencias y se obtiene:

1. Ordenando de menor a mayor
2. Aplicando $l(q) = \dfrac{[Med] + 1}{2}$

Medidas de posición

Cuartil superior: o tercer cuartil tiene por debajo de su valor al 75% de los valores de la distribución de frecuencias y se obtiene:

1. Ordenando de mayor a menor
2. Aplicando $l(q) = \frac{[Med] + 1}{2}$

Ejemplo 33:

Se tiene la muestra de ventas diarias de un producto a lo largo de 7 semanas

Día/Semana	1	2	3	4	5	6	7
L	0	2838	413	5592	0	465	2119
M	515	590	47	673	80	703	
M	746	331	340	561	159	462	
J	1237	450	265	548	183	175	
V	879	570	1083	216	113	422	

Medidas de posición

Ejemplo 33 (*Continúa*):

Sin Orden		A > Z			Z > A		
1	0	1	0		1	5592	
2	515	2	0		2	2838	
3	746	3	47		3	2119	
4	1237	4	80		4	1237	
5	879	5	113		5	1083	
6	2838	6	159		6	879	
7	590	7	175		7	746	
8	331	8	183	199.5	8	703	688
9	450	9	216		9	673	
10	570	10	265		10	590	
11	413	11	331		11	570	
12	47	12	340		12	561	
13	340	13	413		13	548	
14	265	14	422		14	515	
15	1083	15	450		15	465	
16	5592	16	462		16	462	
17	673	17	465		17	450	
18	561	18	515		18	422	
19	548	19	548		19	413	
20	216	20	561		20	340	
21	0	21	570		21	331	
22	80	22	590		22	265	
23	159	23	673		23	216	
24	183	24	703		24	183	
25	113	25	746		25	175	
26	465	26	879		26	159	
27	703	27	1083		27	113	
28	462	28	1237		28	80	
29	175	29	2119		29	47	
30	422	30	2838		30	0	
31	2119	31	5592		31	0	

Medidas de posición

Un método de cálculo para obtener el *p-ésimo* percentil es:

1. Ordenar ascendentemente los datos

2. Calcular $i = \frac{p}{100}n$

   Donde

   p = percentil de interés

   n = tamaño de muestra

3. A)Si i no es entero, entonces se redondea y el valor entero inmediato indicará la posición del *p-ésimo* percentil

   B) Si i es entero, el *p-ésimo* percentil es el promedio de los valores ubicado en los lugares i e $i+1$.

Porcentaje	i	**Valor**
10%	3	47
20%	6	159
30%	9	216
40%	12	340
50%	16	462
60%	19	548
70%	22	590
80%	25	746
90%	28	1237

Medidas de dispersión

Amplitud (R): Mide la distancia entre la observación de mayor valor y la de menor valor en el conjunto de observaciones:

$$R = Amplitud = Máx - min$$

Ejemplo 34:

$$R_{edad} = 25 - 19 = 6$$

$$R_{valor\ catastral} = 370325 - 79928 = 290397$$

Es una medida muy sencilla de calcular y se utiliza para control estadístico de calidad.

Medidas de dispersión

Amplitud Intercuartílica (A.I.): Mide la distancia entre el cuartil superior y el inferior

$$A.I. = q_3 - q_1$$

Ejemplo 35:

A Partir del Ejemplo 33, tenemos:

$$A.I. = 688 - 199.5 = 488.5$$

La $A.I.$ Es una estadística resistente ya que su valor no se ve afectado por observaciones atípicas.

Bibliografía

1. Wackerly, D.D., Mendenhall, W., and Scheaffer, R.L. (2008), *Mathematical Statistics with Applications* 7th Edition, Duxbury, Thompson, Brooks/Cole.

2. Wackerly, D.D., Mendenhall, W., and Scheaffer, R.L. (2010), *Estadística Matemática con aplicaciiones* 7ma Edición, CENGAGE Learning.

3. Pitman, J. (1993), *Probability*. Springer 6a. Ed.

4. Ross, S. (1993), *A First Course in Probability*. Pearson 9th. Ed.

5. Canavos, G.C. (1987), *Probabilidad y Estadística*, McGraw Hill.

6. Mittelhammer (2013),*Mathematical Statistics for Economics and Business*, 2nd Ed. Springer.