

Variables Aleatorias y Distribuciones de Probabilidad

Jhon Jairo Padilla A., PhD.

Ejemplo

- Suponga el caso de la prueba de tres componentes electrónicos tomados al azar de un lote de componentes. El resultado de la prueba para cada componente puede ser: D (Defectuoso), ó N (No defectuoso).
- El espacio muestral será $S=\{NNN, NND, NDN, DNN, NDD, DND, DDN, DDD\}$
- Si es de nuestro interés el número de componentes defectuosos que se presenten, a cada resultado del espacio muestral se podría asignar un valor numérico especificando esta cantidad.
- A la función (Y) que asigna estos valores numéricos a cada resultado se le llama una **Variable Aleatoria**

Resultado	Número componentes defectuosos (Y)
NNN	0
NND	1
NDN	1
NDD	2
DNN	1
DND	2
DDN	2
DDD	3

Variable Aleatoria

- Definición
 - Una **variable aleatoria** es una función que asocia un número real con cada elemento del espacio muestral de un experimento aleatorio
- Nomenclatura:
 - Se utilizarán letras mayúsculas (p.ej. X) para designar una variable aleatoria.
 - Se utilizarán letras minúsculas (p.ej. x) para denotar cada uno de los valores de la variable aleatoria (X).

Ejemplo

- Para el ejemplo anterior, Y denota la variable aleatoria (v.a.) que asigna valores numéricos a los resultados del experimento aleatorio en cuestión.
- En este caso, la v.a. toma el valor 2 para todos los elementos del subconjunto $E = \{DDN, DND, NDD\}$
- Por tanto,
 - Cada valor posible de la v.a. Y representa un evento que es un subconjunto del espacio muestral para el experimento dado.

Ejemplo

- Se sacan dos bolas de manera sucesiva sin reemplazo, de una urna que contiene 4 bolas rojas y 3 negras. Si X es la variable aleatoria que describe el número de bolas rojas que se sacaron en el experimento, los posibles resultados y los valores x de la variable aleatoria X son,

Resultado	Y
RR	2
RB	1
BR	1
BB	0

Espacio muestral discreto

- En los ejemplos anteriores, el espacio muestral contiene un número finito de elementos.
- Otro tipo de experimentos puede tener un número infinito de elementos, pero a cada uno de ellos se les puede asignar un número entero.
- **Ejemplo:** Lanzamiento de un dado hasta que se saque el número 5.
 - Espacio muestral: $S = \{O, NO, NNO, NNNO, \dots\}$
 - Donde O: ocurrencia, N: No ocurrencia.

- **Definición:**
 - Si un espacio muestral contiene un número finito de posibilidades, o una serie interminable con tantos elementos como números enteros existen, se llama **espacio muestral discreto**.

Espacio muestral continuo

- Justificación:
 - Los resultados de algunos experimentos estadísticos no pueden ser ni finitos ni contables
- Ejemplo:
 - Experimento para medir las distancias que recorre cierta marca de automóvil, en una ruta de prueba pre-establecida, con cinco litros de gasolina.
 - Al medir la distancia, se tiene un número infinito de distancias posibles en el espacio muestral, que no se pueden igualar a números enteros.
- Definición:
 - Si un espacio muestral contiene un número infinito de posibilidades similar al número de puntos en un segmento de línea, se le llama **espacio muestral continuo**.

Variables Aleatorias Discretas y Contínuas

- Una v.a. se llama **variable aleatoria discreta** si se puede contar su conjunto de resultados posibles
- Cuando una v.a. puede tomar valores en una escala contínua, se le denomina **variable aleatoria contínua**.

Ejemplos

v.a. discreta

- Experimento bolas rojas y negras

Resultado	Y
RR	2
RB	1
BR	1
BB	0

v.a. continúa

- Sea X una v.a. definida como el tiempo de espera entre conductores sucesivos que exceden los límites de velocidad, detectados por una unidad de radar en una carretera.
- La v.a. X toma todos los valores de x tales que $x \geq 0$

VARIABLES ALEATORIAS DISCRETAS

Distribución de probabilidad

- Necesidad:
 - En muchas ocasiones el interés se centra en la probabilidad de que una v.a. asuma un valor particular
 - La **distribución de probabilidad** de una v.a. X es una descripción de las probabilidades asociadas con los valores posibles de X . Puede expresarse como un listado o tabla, o como una fórmula.
- Ejemplo:
 - Existe la posibilidad de que un bit transmitido a través de un canal de comunicaciones se reciba con error. Sea X igual al número de bits con error cuando se transmiten cuatro bits

x	$P(x)$
0	0,6561
1	0,2916
2	0,0486
3	0,0036
4	0,0001

Cubren todos los Resultados posibles

La sumatoria da 1

Función de masa de probabilidad

- El conjunto de pares ordenados $(x, f(x))$ es una **función de probabilidades**, una **función de masa de probabilidad** o una **distribución de probabilidad**

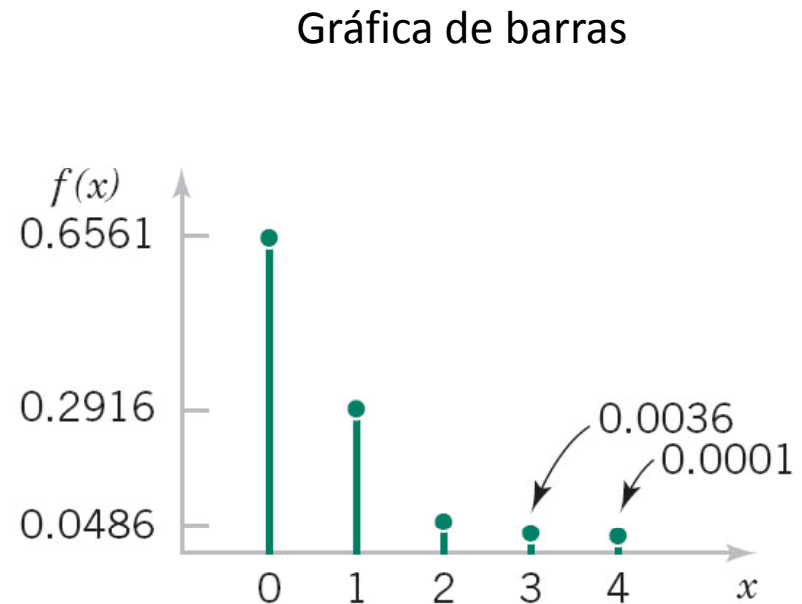


Figure 3-1 Probability distribution for bits in error.

Fuente: Montgomery, D. 2007

Ejemplo: Descripción de $f(x)$ como una fórmula

- Si una agencia automotriz vende 50% de su inventario de cierto vehículo extranjero equipado con bolsas de aire, encuentre la fórmula para la distribución de probabilidad del número de automóviles con bolsas de aire entre los siguientes 4 vehículos que venda la agencia.
- Solución:
 - La probabilidad de vender un vehículo con bolsas de aire es 0.5
 - El espacio muestral para los siguientes 4 vehículos tendría $2^4=16$ posibles resultados
 - Cada resultado sería $1/16$ del espacio muestral
 - El evento de vender x modelos con bolsas de aire y $4-x$ modelos sin bolsas de aire es una combinación con $\binom{4}{x}$ formas.
 - x puede ser 0, 1, 2, 3, 4
 - Entonces la distribución de probabilidad es $f(x) = \frac{\binom{4}{x}}{16}$

Función de distribución acumulada

- Necesidad:
 - En muchas situaciones se requiere calcular la probabilidad de que el valor observado de una variable aleatoria X sea menor o igual que algún número real x .
 - En otras palabras la probabilidad de este evento sería la suma de las probabilidades para todo valor de $X \leq x$

- Definición:
 - La función de distribución acumulada $F(x)$ de una variable aleatoria discreta X con distribución de probabilidad $f(x)$ es

$$F(x) = P(X \leq x) = \sum_{t \leq x} f(t)$$

- Para $-\infty < x < \infty$

Ejemplo

x	$P(x)$
0	0,6561
1	0,2916
2	0,0486
3	0,0036
4	0,0001

$$F(x) = P(X \leq 2) = 0,6561 + 0,2916 + 0,0486 = 0,9963$$

Función de distribución acumulada

- La distribución acumulada de probabilidad puede usarse para encontrar la función de masa de probabilidad de una variable aleatoria.
- Por tanto, utilizar la función de distribución acumulada es un método alternativo para describir la distribución de probabilidad de una variable aleatoria.

Ejemplo

- Suponga la distribución de probabilidad de la tabla. Calcule la función de distribución acumulada. Verifique que $f(2)=3/8$.

x	$f(x)$
0	1/16
1	1/4
2	3/8
3	1/4
4	1/16

$$F(0) = f(0) = 1/16$$

$$F(1) = f(0) + f(1) = 5/16$$

$$F(2) = f(0) + f(1) + f(2) = 11/16$$

$$F(3) = f(0) + f(1) + f(2) + f(3) = 15/16$$

$$F(4) = f(0) + f(1) + f(2) + f(3) + f(4) = 1$$

$$f(2) = F(2) - F(1) = \frac{11}{16} - \frac{5}{16} = \frac{3}{8}$$

Ejemplo

- Suponga que la función de distribución acumulada de la variable aleatoria X es la de la figura. Determine la función de masa de probabilidad de X .

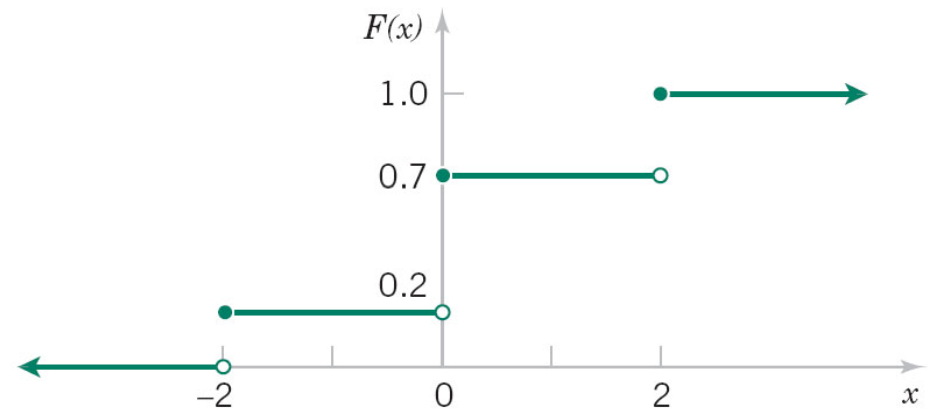


Figure 3-3 Cumulative distribution function for Example 3-7.

Fuente: Montgomery, D. 2007

$$f(-2) = 0,2 - 0$$

$$f(0) = 0,7 - 0,2 = 0,5$$

$$f(2) = 1,0 - 0,7 = 0,3$$

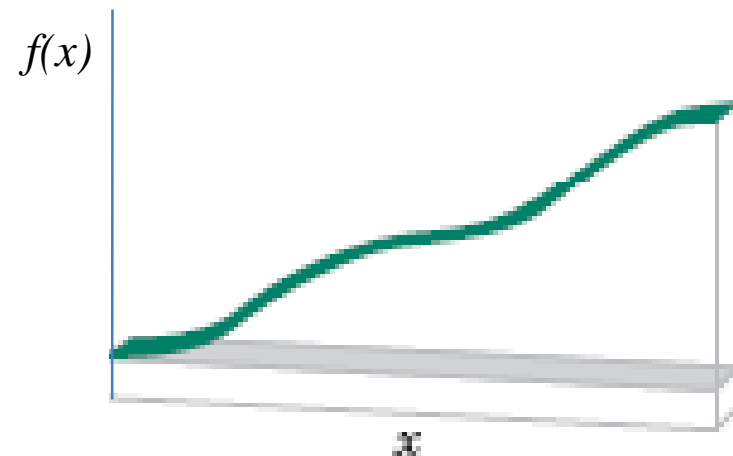
VARIABLES ALEATORIAS CONTÍNUAS

Para Recordar

- Una v.a. continua X toma valores en un intervalo de números reales, por lo que el rango de X puede considerarse como continuo.

Función de densidad de probabilidad

- La función de distribución de probabilidad no se puede representar de forma tabular, sino mediante una fórmula
- Cuando se tienen variables aleatorias continuas, se utiliza una **función de densidad de probabilidad**, $f(x)$, para describir la función de distribución de probabilidad.

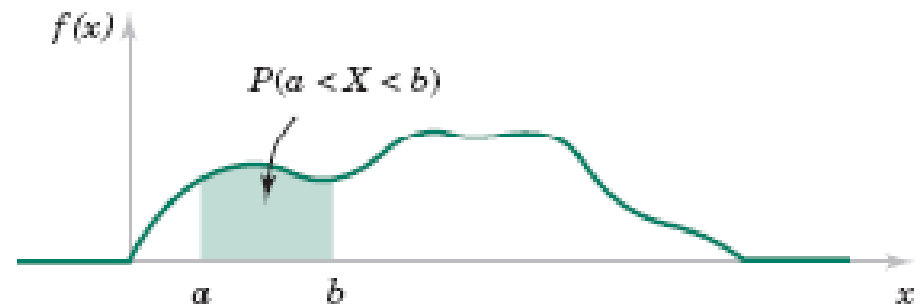


Fuente: Montgomery, D. 2007

Cálculo de la probabilidad

- La probabilidad de que x esté entre a y b se calcula como la integral de $f(x)$ de a hasta b .

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



Fuente: Montgomery, D. 2007

Función de densidad de probabilidad

- Definición:

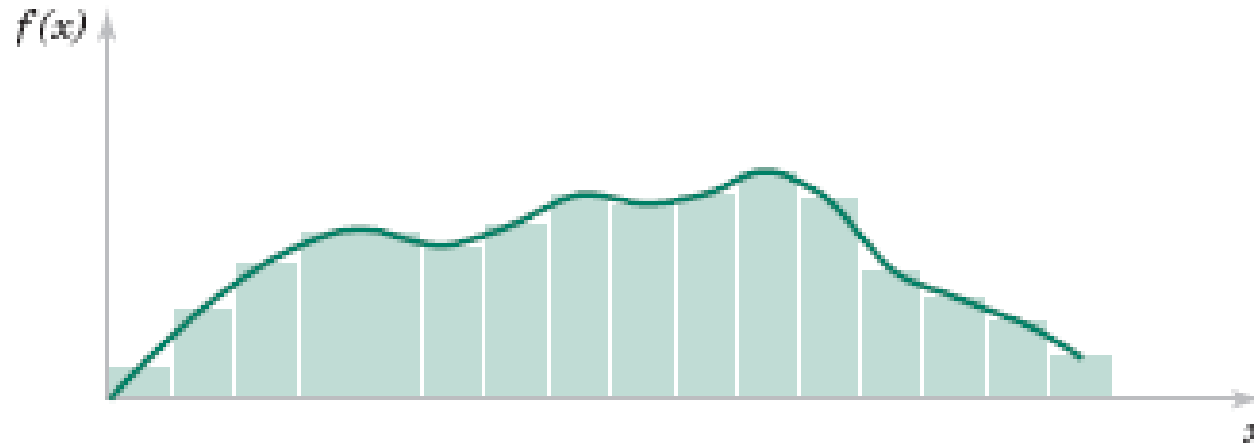
- La función $f(x)$ es una función de densidad de probabilidad (fdp) para la variable aleatoria continua X , definida en el conjunto de números reales R , si

$$f(x) \geq 0 \quad \text{para toda } x \in R$$

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

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

Significado del Histograma



Fuente: Montgomery, D. 2007

Un Histograma es una aproximación de una función de densidad de probabilidad

Ejemplo

- Sea que una v.a. continua denote la corriente medida en un alambre delgado de cobre en mA. Suponga que el rango de X es de 0 a 20mA, y suponga también que la fdp de X es $f(x)=0,05$ para $0 \leq x \leq 20$. Cuál es la probabilidad de que una medición de la corriente sea menor que 10mA? (se supone que $f(x)=0$ para $x < 0$)

$$P(X < 10) = \int_0^{10} 0,05 dx = 0,5$$

Ejemplo

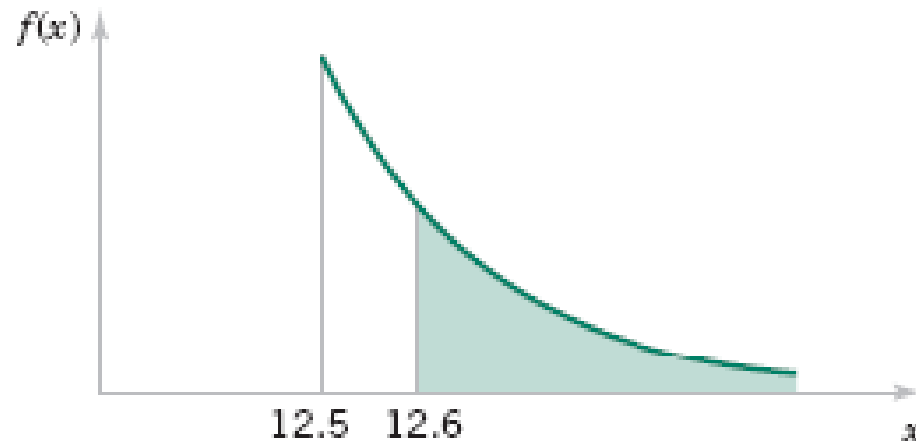
- Sea que la v.a. continua X denote el diámetro de un agujero taladrado en un componente metálico. El diámetro especificado es de 12.5mm. La mayoría de las perturbaciones aleatorias del proceso resultan en diámetros mayores. Datos históricos indican que la distribución de X puede modelarse con la función de densidad de probabilidad:

$$f(x) = 20e^{-20(x-12.5)} \quad \text{con } x \geq 12.5$$

- Si un componente con un diámetro mayor que 12.6mm se desecha, ¿qué proporción de componentes se desecha?

Ejemplo

Solución:



Fuente: Montgomery, D. 2007

$$P(x > 12.6) = \int_{12.6}^{\infty} f(x) = \int_{12.6}^{\infty} 20e^{-20(x-12.5)} dx = -e^{-20(x-12.5)} \Big|_{12.6}^{\infty} = 0.135$$

Función de distribución acumulada

- Al igual que con las v.a. discretas, las v.a. continuas pueden describirse mediante la función de distribución acumulada.

- Definición:

- La función de distribución acumulada $F(x)$ de una variable aleatoria continua X con función de densidad $f(x)$ es

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(t)dt \quad \text{para } -\infty < x < \infty$$

Función de distribución acumulada

- En consecuencia, de la definición:

$$P(a < x < b) = F(b) - F(a)$$

$$f(x) = \frac{dF(x)}{dx}$$

Ejemplo

$$f(x) = 20e^{-20(x-12.5)}$$

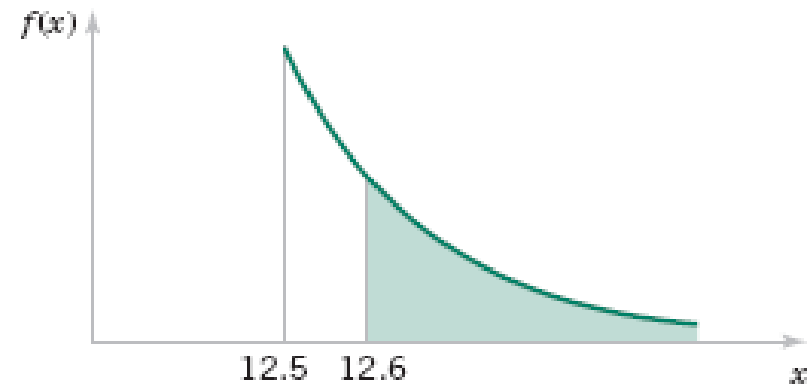
- Para el ejemplo anterior, hallar $F(x)$

- Solución:

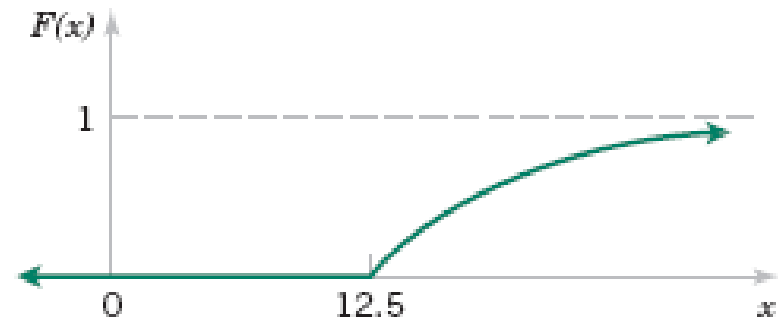
$$F(x) = 0 \text{ para } x < 12.5$$

$$F(x) = \int_{12.5}^x f(x) = \int_{12.5}^x 20e^{-20(u-12.5)} du = 1 - e^{-20(x-12.5)}$$

$$\text{para } x \geq 12.5$$



Fuente: Montgomery, D. 2007



Fuente: Montgomery, D. 2007

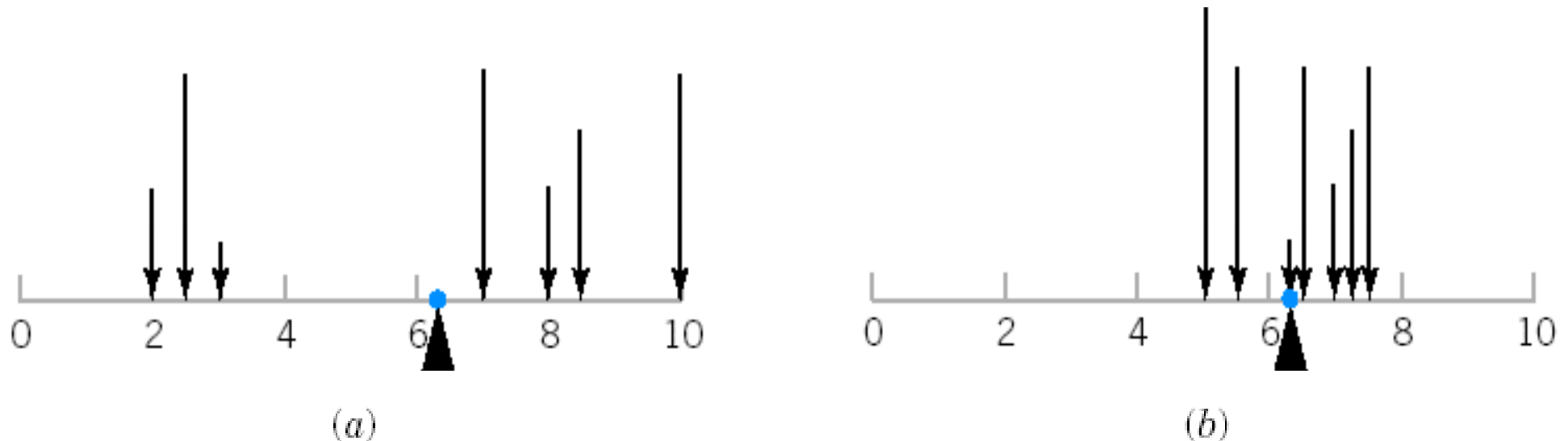
MEDIA Y VARIANZA DE UNA VARIABLE ALEATORIA

Parámetros de descripción de una variable aleatoria

- Una v.a. toma valores según los resultados que va arrojando un experimento aleatorio.
- Se requieren parámetros para describir la ubicación y variabilidad de la v.a.
- Principales Parámetros:
 - Ubicación: Media
 - Variabilidad: Varianza

Media y varianza de una variable aleatoria

- La **media, valor esperado o Esperanza matemática** de una variable aleatoria es el punto de equilibrio de los puntos muestrales.
- La **varianza** es un indicador de qué tan dispersos se encuentran los puntos muestrales.
- La v.a. del caso (a) tiene mayor varianza que la v.a. del caso (b).



Media de una Variable Aleatoria

- Definición:

- Sea X una v.a. con distribución de probabilidad $f(x)$. Si X es una v.a. discreta, la media o valor esperado de X es

$$\mu = E(X) = \sum_x x \cdot f(x)$$

- Si X es una v.a. continua, la media o valor esperado de X es

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

Ejemplo: significado de la media de una v.a.

- Supóngase el experimento aleatorio de lanzar una moneda balanceada dos veces. Los posibles resultados se ilustran en la tabla.
- Sea X una v.a. que representa el número de caras que ocurren en el resultado obtenido.
- Supónganse resultados equiprobables.
- Calcule la media de X

Resultado	X	$f(x)$
CC	2	$\frac{1}{4}$
CS	1	$\frac{1}{2}$
SC	1	$\frac{1}{4}$
SS	0	$\frac{1}{4}$

$$P(X = 0) = P(SS) = \frac{1}{4}$$

$$P(X = 1) = P(CS) + P(SC) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

$$P(X = 2) = P(CC) = \frac{1}{4}$$

$$\mu = E(X) = (0)\left(\frac{1}{4}\right) + (1)\left(\frac{1}{2}\right) + (2)\left(\frac{1}{4}\right) = 1$$

Significado: Una persona que realice el experimento, en el largo plazo (después de muchos experimentos) obtendrá, en media, 1 cara por cada lanzamiento.

Ejemplo: Media de una v.a. discreta

- Un inspector de calidad muestrea un lote que contiene 7 componentes; el lote contiene 4 componentes buenos y 3 defectuosos. El inspector toma una muestra de 3 componentes. Encuentre el valor esperado del número de componentes buenos en esta muestra. Suponga que todos los resultados tienen igual probabilidad.

x	f(x)
0	1/35
1	12/35
2	18/35
3	4/35

- Solución:

- Número de posibles resultados:

$$N = \binom{7}{3} = \frac{7!}{3!4!} = 35$$

- La probabilidad de cada resultado será entonces 1/35

- El número de resultados con x componentes buenos será

$$\binom{4}{x} \binom{3}{3-x}$$

- La probabilidad de que x componentes de la muestra estén buenos será:

$$f(x) = \frac{\binom{4}{x} \binom{3}{3-x}}{\binom{7}{3}}$$



$$\mu = E(X) = 0 \cdot \frac{1}{35} + 1 \cdot \frac{12}{35} + 2 \cdot \frac{18}{35} + 3 \cdot \frac{4}{35} = \frac{12}{7} = 1.7$$

Ejemplo: aplicación de la media

- En un juego de azar se pagarán \$5 a una persona si le salen puras caras o puros sellos cuando se lanzan tres monedas. La persona pagará \$3 si salen una o dos caras. ¿cuál es su ganancia esperada?

- Solución:
 - $S = \{CCC, CCS, CSC, CSS, SCC, SCS, SSC, SSS\}$
 - Probabilidad de cada resultado: $1/8$
 - Y : v.a. que describe el monto que el jugador puede ganar.
 - Entonces:
 - $Y = 5$ para $E_1 = \{CCC, SSS\}$
 - $Y = -3$ para $E_2 = \{CCS, CSC, CSS, SCC, SCS, SSC\}$

Luego:

$$P(E_1) = 2/8 = 1/4$$

$$P(E_2) = 6/8 = 3/4$$

Por tanto,

$$\mu = E(Y) = (5) \left(\frac{1}{4} \right) + (-3) \left(\frac{3}{4} \right) = -1$$

Conclusión: En media, el jugador perderá \$1



Ejemplo: media de una v.a. continua

- Sea X la v.a. que denota la vida en horas de cierto dispositivo electrónico. La f.d.p. es:

$$f(x) = \begin{cases} \frac{20,000}{x^3}, & x > 100 \\ 0; & x \leq 100 \end{cases}$$

- Encuentre la vida esperada para esta clase de dispositivo
- Solución:

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

$$\mu = E(X) = \int_{-\infty}^{100} x \cdot 0 dx + \int_{100}^{\infty} x \cdot \frac{20,000}{x^3}$$

$$\mu = E(X) = 20,000 \cdot \frac{x^{-1}}{-1} \Big|_{100}^{\infty} = -20,000 \left(\frac{1}{\infty} - \frac{1}{100} \right) = 200$$

Varianza de una variable aleatoria

- Definición:

Sea X una v.a. con distribución de probabilidad $f(x)$ y media μ . La varianza de X es,

$$\sigma^2 = E[(X - \mu)^2] = \sum_x (x - \mu)^2 f(x) \quad \text{si } X \text{ es discreta, y}$$

$$\sigma^2 = E[(X - \mu)^2] = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx \quad \text{si } X \text{ es continua.}$$

La raíz cuadrada positiva de la varianza se llama desviación estándar (σ).

Ejemplo

- Sea la v.a. X el número de automóviles que se utilizan con propósitos de negocios oficiales en un día de trabajo dado. Las distribuciones de probabilidad para las compañías A y B son:

x	0	1	2	3	4
f(x) var. A	0	0.3	0.4	0.3	0
f(x) var. B	0.2	0.1	0.3	0.3	0.1

$$\mu_A = E(X) = 1 \cdot 0,3 + 2 \cdot 0,4 + 3 \cdot 0,3 = 2,0$$

$$\sigma_A^2 = \sum_x (x-2)^2 f(x) = (1-2)^2 (0,3) + (2-2)^2 (0,4) + (3-2)^2 (0,3) = 0,6$$

$$\mu_B = E(X) = 0 \cdot 0,2 + 1 \cdot 0,1 + 2 \cdot 0,3 + 3 \cdot 0,3 + 4 \cdot 0,1 = 2,0$$

$$\sigma_B^2 = \sum_x (x-2)^2 f(x) = (0-2)^2 (0,2) + (1-2)^2 (0,1) + (2-2)^2 (0,3) + (3-2)^2 (0,3) + (4-2)^2 (0,1) = 1,6$$