Modelos de Distribución Discreta I

Introducción a la Estadística

DVM

Contenidos

1. Variable Aleatoria

2. Variable Aleatoria Discreta
   - Función de probabilidad (PMF)
   - Función de probabilidad acumulada (CMF)
   - Esperanza, Varianza y Desviación Típica

3. Variable Aleatoria Bidimensional Discreta
   - Función de probabilidad Conjunta (PMF)
   - Correlación y Covarianza

4. Teorema de la Probabilidad Total y Teorema de Bayes

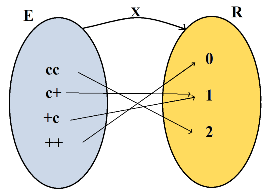
Variable Aleatoria

Variable Aleatoria

Una variable aleatoria X es una función que asocia a cada suceso del espacio muestral, E, de un experimento aleatorio un valor numérico real. Esta puede ser discreta o continua.

$$X : E \rightarrow \text{R}$$
$$s \rightarrow \text{X(s)}$$

Supongamos que hacemos el siguiente experimento: "Lanzamiento de dos monedas" $E = \{cc, c+, +c, ++\}$ y definimos la variable aleatoria X como "numero de caras"



Contenidos

1. Variable Aleatoria
2. Variable Aleatoria Discreta
   - Función de probabilidad (PMF)
   - Función de probabilidad acumulada (CMF)
   - Esperanza, Varianza y Desviación Típica
3. Variable Aleatoria Bidimensional Discreta
   - Función de probabilidad Conjunta (PMF)
   - Correlación y Covarianza
4. Teorema de la Probabilidad Total y Teorema de Bayes

Variable Aleatoria Discreta
Definición

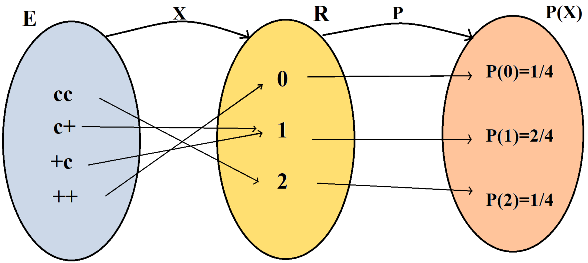
Es aquella variable aleatoria cuyo dominio supone un espacio muestral que se puede contar (es finito o infinito numerable).

- Numero de personas presentes en esta conferencia
- Numero de bolas rojas en una urna con bolas de distintos colores
- Numero de caras al lanzar tres monedas
- El nivel de notas de los alumnos (S, Ap, N, SB, MH)

Función probabilidad (PMF)

La distribución de probabilidad de una variable aleatoria refleja su comportamiento probabilistico. Nos referiremos a ella como PMF : función de masa de probabilidad.

$$f : R \rightarrow [0,1]$$
$$x \rightarrow f(x) \; (=P(X=x))$$



Como vemos, f es la función de cuantía o de probabilidad que resulta en el conjunto de los posibles valores de la v.a. y su probabilidad de ocurrencia asociada: $(x_i, f(x_i))$.

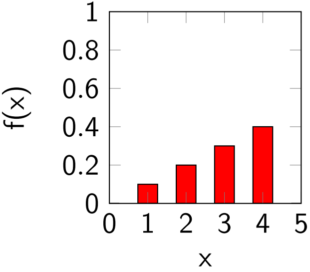
Función de probabilidad (PMF)

Ejemplo 1

Ejemplo Dada la función de probabilidad se pide calcular la distribución de probabilidad de X:

$$f(x) = \begin{cases} 0,1 * x & x = 1,2,3,4 \\ 0 & resto \end{cases}$$

x_i	$f(x_i)$
1	0,1
2	0,2
3	0,3
4	0,4



Función de probabilidad (PMF)

Ejemplo 2

Consideremos el experimento de "lanzar dos dados" y definamos la variable aleatoria discreta X como:

$$X : E \rightarrow \text{S} \in \text{R}$$

con S ="la suma de puntos"$(S = 2, 3, ..., 12)$

Podemos considerar la siguiente función de probabilidad,f :

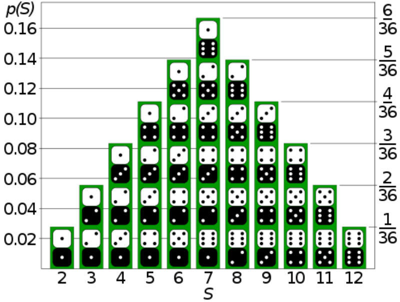
x	**f(x)**
2	P(1,1)
3	$P((1,2) \cup (2,1))$
4	$P((1,2) \cup (2,1) \cup (2,2))$
.	.
.	.
.	.
12	P(6,6)

Función de probabilidad (PMF)

Ejemplo 2

x	**f(x)**
2	$P(1,1) = 1/36$
3	$P((1,2) \cup (2,1)) = 2/36$
4	$P((1,2) \cup (2,1) \cup (2,2)) = 3/36$
.	.
.	.
.	.
12	$P(6,6)=1/36$

Si graficamos:



Función de probabilidad acumulada (CMF)

Describe la acumulación de probabilidades correspondiente a la variable aleatoria X. Describe la probabilidad de que X tenga un valor menor o igual que x: $F(x) = P(X \leq x)$

$$F : R \rightarrow [0,1]$$
$$x \rightarrow \text{F(x) (=P(X} \leq x))$$

Propiedades de la CMF en la v.a.:

1. No puede ser negativa: $0 \leq F(x) \leq 1$
2. F es monótona creciente por su carácter acumulativo (una función es monótona creciente cuando si $x_1 < x_2$ se ha de verificar que $F(X_1) \leq F(X_2)$)
3. Es continua por la derecha y discontinua por la izquierda así que la representación gráfica es discontinua en forma de escalera.

Función de probabilidad acumulada (CMF)

4 La probabilidad de un intervalo se resuelve en el caso de una v.a. discreta como: $P(a < x \leq b) = F(b) - F(a)$
$P(a < x < b) = F(b) - F(a) - P(X = b)$
$P(a \leq x < b) = F(b) - F(a) + P(X = a) - P(X = b)$
$P(a \leq x \leq b) = F(b) - F(a) + P(X = a)$

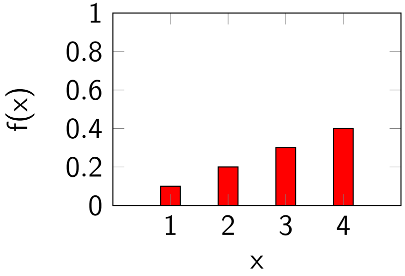
Función de probabilidad acumulada (CMF)

Ejemplo 1

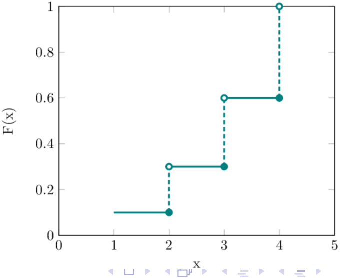
En el Ejemplo1 teníamos que:

x_i	$f(x_i)$	$F(x_i)$
1	0,1	$0,1$
2	0,2	$0,3$
3	0,3	$0,6$
4	0,4	1

Con gráfica de su PMF:



La gráfica de su CMF sería:



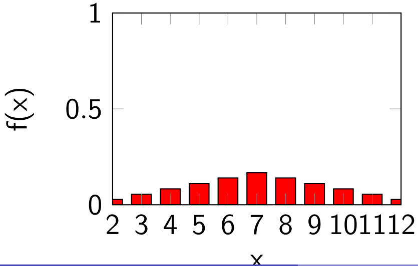
Función de probabilidad acumulada (CMF)

Ejemplo 2

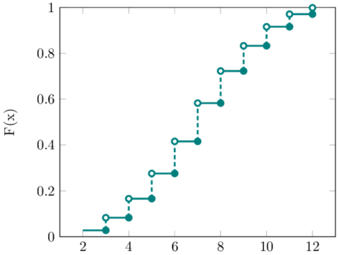
En el Ejemplo 2 teníamos que:

x_i	$f(x_i)$	$F(x_i)$
2	$1/36$	$1/36$
3	$2/36$	$3/36$
4	$3/36$	$6/36$
.	.	.
.	.	.
.	.	.
12	$1/36$	1

Gráfica de su PMF:



La gráfica de su CMF:



Media

La media es una **medida de centralidad** que presenta una variable aleatoria, un valor alrededor del cual “se espera” que se realice (de ahí lo de valor esperado o esperanza).
Por ello, la masa de probabilidad asignada a los diferentes valores de v.a., también se reordena alrededor de este valor medio.
Viene dado por:

$$\mu = E(X) = \sum_{i=1}^{n} x_i * p_i$$

Es decir, se tiene en cuenta de manera proporcional todos los posibles valores que puede tomar la v.a.

Media: Ejemplo 1

Dada la tabla con la distribución de probabilidad del Ejemplo 1 vamos a calcular la media:

x_i	$f(x_i)$	$F(x_i)$
1	0,1	$0,1$
2	0,2	$0,3$
3	0,3	$0,6$
4	0,4	1

- $\mu = E(X) = 1 * 0,1 + 2 * 0,2 + 3 * 0,3 + 4 * 0,4 = 3$

Media: Propiedades

Propiedades de la media de una v.a.

- Escalar: $E[c * X] = c * E[X] = c * \mu$
- Origen: $E[c + X] = c + E[X] = c + \mu$
- Esperanza de la suma: $E[aX + bY] = aE[X] + bE[Y]$

Varianza y Desviación Típica

Varianza

La varianza de una v.a. es una **medida de dispersion** que describe la variabilidad (la distancia más frecuente a la que se situan) de las observaciones respecto de su media.

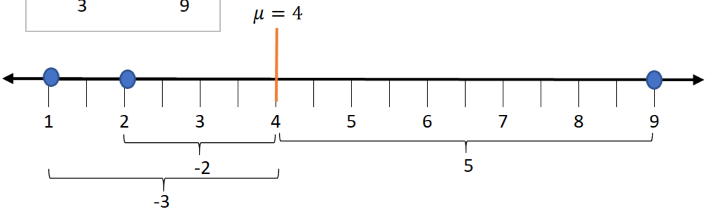
$$\sigma^2 = E[(x-\mu)^2] = \frac{\sum_{i=1}^{N}(x_i-\mu)^2}{N} = \sum_{i=1}^{N}(x_i-\mu)^2 * p_i$$

La varianza es la media del cuadrado de las desviaciones respecto a la media de una distribución estadística:

Varianza y Desviación Típica

¿Por qué el cuadrado?

Hamster id	Peso (g)
1	1
2	2
3	9



$$Media\ de\ la\ distancia = \frac{(-3)+(-2)+5}{3} = 0$$

$$Media\ de\ la\ distancia^2 = \frac{(-3)^2+(-2)^2+5^2}{3} \cong 12.67 \longrightarrow var(peso)\ [\ g^2\ ]$$

Varianza y Desviación Típica

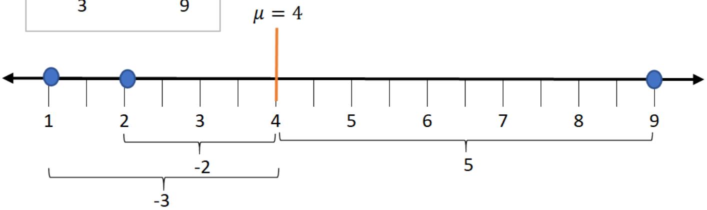
Propiedades de la varianza de una v.a.

- La varianza siempre es positiva: $\sigma^2 \geq 0$
- Escala: $Var(c + X) = Var(X)$
- Origen: $Var(cX) = c^2 * Var(X)$
- Escala y Origen: $Var(aX + b) = a^2 * Var(X)$
- Varianza de la suma de dos variables aleatorias X e Y independientes: $Var(X + Y) = Var(X) + Var(Y)$
- Varianza de la suma de 2 variables aleatorias no independientes, entonces: $Var(X + Y) = Var(X) + Var(Y) + 2cov(X, Y)$

Varianza y Desviación Típica

¿Por qué el cuadrado?

Hamster id	Peso (g)
1	1
2	2
3	9



$$Media\ de\ la\ distancia = \frac{(-3)+(-2)+5}{3} = 0$$

$$Media\ de\ la\ distancia^2 = \frac{(-3)^2+(-2)^2+5^2}{3} \cong 12.67$$

$var(peso)\ [\ g^2\ ]$

$\sqrt{12.67} \cong 3.6$

$std.desv(peso)\ [g]$

Varianza y Desviación Típica

Desviación Típica

El resultado de tomarle la raíz cuadrada al valor de la varianza para que la distancia quede expresada en las mismas unidades que el valor esperado (media) se le denomina Desviación típica o estándar.

$$\sigma = \sqrt{\sigma^2}$$

Contenidos

1. Variable Aleatoria

2. Variable Aleatoria Discreta
   - Función de probabilidad (PMF)
   - Función de probabilidad acumulada (CMF)
   - Esperanza, Varianza y Desviación Típica

3. Variable Aleatoria Bidimensional Discreta
   - Función de probabilidad Conjunta (PMF)
   - Correlación y Covarianza

4. Teorema de la Probabilidad Total y Teorema de Bayes

Variable Aleatoria Bidimensional Discreta

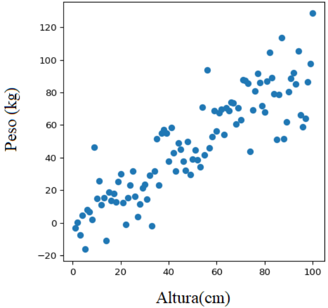
En algunos casos, tenemos dos variables aleatorias discretas de interes. Por ejemplo, es posible que tengamos:

- X: "numero de relaciones personales intimas que tiene z"
- Y: "numero de discusiones que tiene z por semana"

Entonces, nos gustaria saber si existe una relacion entre ambas o simplemente conocer la probabilidad de que X e Y tomen valores concretos : $P(X = x, Y = y)$

Variable Aleatoria Bidimensional Discreta

- Así, observar dos características X e Y de forma simultanea nos lleva a contemplar una **variable aleatoria bidimensional** o bivariante **(x,y)**.
- En términos matemáticos supone contemplar un vector en el plano. Este vector tendrá un campo de variación y una distribución de probabilidad que denominamos conjunta.



- ¿Cuál es la diferencia entre la variable (x,y) anterior y la que consideramos en el grafico anterior?

Función de probabilidad Conjunta (PMF)

Función de masa de probabilidad Conjunta (PMF)

Si tenemos dos v.a.'s discretas, X e Y y queremos estudiarlas conjuntamente, definimos la función de masa de probabilidad conjunta (pmf conjunta) así:

$$f_{xy} = P(X = x, Y = y)$$

Ejemplo: Considera la siguiente probilidad conjunta de X e Y discretas aportada por la siguiente tabla de contingencia:

f(x,y)	y=0	y=1	y=2
x=0	1/6	1/4	1/8
x=1	1/8	1/6	1/6

Función de Probabilidad Conjunta (PMF)

Funciones de probabilidad marginales

También llamaremos a f_x y f_y , las funciones de probabilidad marginales de X y de Y , respectivamente. Con:

$$f_x = \sum_y f(x,y) = P(X = x) \text{ y } f_y = \sum_x f(x,y) = P(Y = y)$$

Ejemplo:

f(x,y)	y=0	y=1	y=2	f_x
x=0	1/6	1/4	1/8	13/24
x=1	1/8	1/6	1/6	11/24
f_y	7/24	10/24	7/24	

Función de Probabilidad Conjunta (PMF)

Funciones de Probabilidad condicional e independencia

$$f(X = x|Y = y) = \frac{f((X=x, Y=y))}{f_y}$$

Si se da el caso que $f(x, y) = f_x * f_y$ entonces $f(X|Y) = f_x$, es decir , X e Y son **independientes**.

En el ejemplo anterior:

f(x,y)	y=0	y=1	y=2	f_x
x=0	1/6	1/4	1/8	13/24
x=1	1/8	1/6	1/6	11/24
f_y	7/24	10/24	7/24	

- $P(Y = 1|X = 0) ó f(Y = 1|X = 0) = \frac{1/4}{13/24} = 6/13$
- X e Y no son independientes porque $f(Y = 1|X = 0) = \frac{6}{13} \neq f(Y = 1) = \frac{10}{24}$

Función Probabilidad Conjunta (PMF)

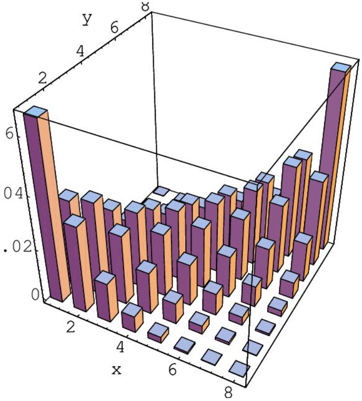


Figure: Gráfica de una función de probabilidad conjunta para X e Y discretas tomando valores de 1 a 8.

Función de Probabilidad Conjunta (PMF)

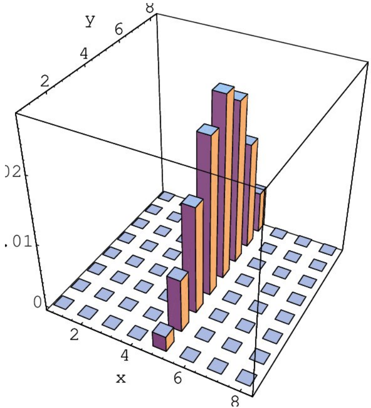


Figure: Gráfica de la distribución de probabilidadde Y condicionada a X=4 (de la función de la diapositiva anterior)

Función de Probabilidad Conjunta (PMF)
Ejercicio

Supongamos que lanzamos un par de dados de cuatro caras al aire. Uno de los dados es ROJO y otro es NEGRO. Vamos a definir las dos v.a.'s como:
$X = resultado\ dado\ ROJO = \{1, 2, 3, 4\}$
$Y = resultado\ dado\ NEGRO = \{1, 2, 3, 4\}$

a) Cual es la función de masa de probabilidad?
b) Calcular las probabilidades marginales y especificar el valor de $P(X = 1)$?
c) Cuál es la $P(X = 2|Y = 4)$?
d) Son X e Y sucesos inpendientes?

Función de Probabilidad Conjunta (PMF)

Ejercicio

a) Cual es la función de masa de probabilidad?

f(x,y)	y=1	y=2	y=3	y=4
x=1	1/16	1/16	1/16	1/16
x=2	1/16	1/16	1/16	1/16
x=3	1/16	1/16	1/16	1/16
x=4	1/16	1/16	1/16	1/16

b) Calcular las probabilidades marginales y especificar el valor de $P(X = 1)$?

f(x,y)	y=1	y=2	y=3	y=4	f_x
x=1	1/16	1/16	1/16	1/16	4/16
x=2	1/16	1/16	1/16	1/16	4/16
x=3	1/16	1/16	1/16	1/16	4/16
x=4	1/16	1/16	1/16	1/16	4/16
f_y	4/16	4/16	4/16	4/16	1

c) Cuál es la $P(X = 2|Y = 4)$?

$P(X = 2|Y = 4) = f(X = 2|Y = 4) = \frac{f(X=2,Y=4)}{f(Y=4)} = \frac{1/16}{4/16} = \frac{1}{4}$

d) Son X e Y sucesos inpendientes?

Sí, pues $P(X = 2|Y = 4) = \frac{1}{4} = f(X = 2) = \frac{4}{16}$

Covarianza

Covarianza

La covarianza es una medida de relacion lineal entre dos variables aleatorias. Muestra como ambas varian conjuntamente por lo que puede tener signo positivo (relación lineal positiva) o negativo(relacion lineal negativa). Viene dada por:

$cov_{x,y} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{n}$ si la v.a. (X,Y) toma valores (x_i, y_i) para $i = 1, ..., n$ equiprobables $(p_i = 1/n)$ sino :

$$cov_{x,y} = \sum_{i=1}^{n} p_i * (x_i - \bar{x})(y_i - \bar{y})$$

Covarianza

Propiedades

Se deducen de su propia definición y se tienen que tomar en cuenta a la hora de trabajar con variables.
Asumiendo que b y c son constantes, tenemos:

- $Cov(X, c) = 0$
- $Cov(X, X) = Var(X)$
- $Cov(X, Y) = Cov(Y, X)$
- $Cov(bX, cY) = cbCov(X, Y)$
- $Cov(b + X, c + Y) = Cov(X, Y)$
- $cov_{x,y} = 0$ cuando las variables X e Y son **independientes** pues : $E(XY) = E(X) * E(Y)$ pues tenemos $cov_{x,y} = E(XY) - E(X)E(Y)$ (no hay relación *lineal* entre las variables)

Covarianza

Ejemplo

- Veamos un ejemplo donde X: "Recuperarse (1) o no (0)" e Y: "Nº sesiones de una terapia (1, 2, 3)"

	y=1	y=2	y=3	f_x
x=0	0,07	0,20	0,08	0,35
x=1	0,13	0,15	0,37	0,65
f_y	0,20	0,35	0,45	1

- Datos:
  - $\bar{x} = 0,35 * 0 + 0,65 * 1 = 0,65$
  - $\bar{y} = 0,2 * 1 + 0,35 * 2 + 0,45 * 3 = 2,25$
  - Observese que los sucesos no son equiprobables.

$cov_{x,y} = \sum_{i=1}^{6} p_i * (x_i - 0,65)(y_i - 2,25)$
=0,07*(0-0,65)(1-2,25)+...+0,37*(1-0,65)(3-2,25)
=0,0775

¿Cómo interpretamos este valor?

Sabemos que tienen una relación lineal positiva, pero necesitamos más...

Correlación

Correlación

La versión normalizada de la covarianza es el coeficiente de correlación. Este indica la magnitud de la relación lineal. Viene dado por :

$$\rho(X, Y) = \frac{cov(X,Y)}{\sigma(X)\sigma(Y)}$$

Cumple que:

- $1 \leq \rho_{XY} \leq 1$
- Como para sucesos independientes cov_{XY} es nula, entonces temabién $\rho_{XY} = 0$

Revisemos el ejemplo anterior.

Correlación

Ejemplo

	y=1	y=2	y=3	f_x
x=0	0,07	0,20	0,08	0,35
x=1	0,13	0,15	0,37	0,65
f_y	0,20	0,35	0,45	1

Datos:

- $\bar{x} = 0,65$
- $\bar{y} = 2,25$
- $cov_{x,y} = 0,0775$
  ¿Cómo interpretamos este valor? Sabemos que tienen una relación lineal positiva, pero necesitamos más...

$$\rho_{XY} = \frac{0{,}0775}{\sigma(X)\sigma(Y)}$$

Correlación

Ejemplo

$$\sigma(X) = \sqrt{\sigma^2(X)} = \sqrt{\sum_{i=1}^{n} p_i * (x_i - \bar{x})} =$$
$$\sqrt{0,35 * (0 - 0.65)^2 + 0.65 * (1 - 0.65)^2} = \sqrt{0,2275} = 0,48$$

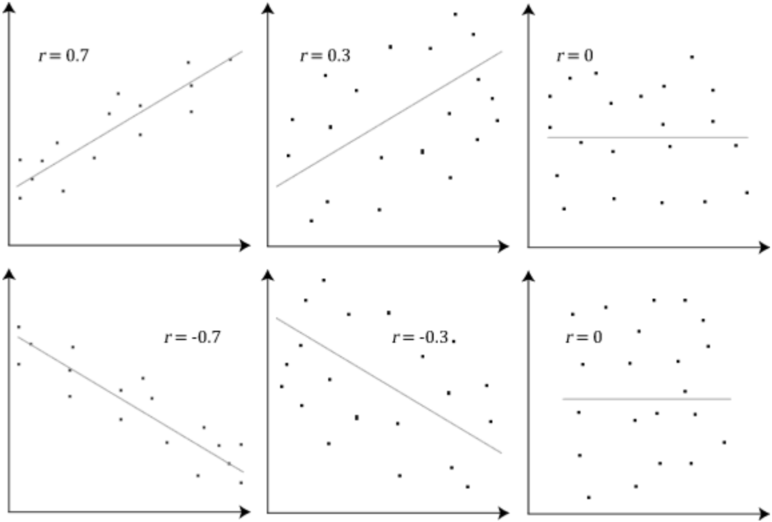
$$\sigma(Y) = \sqrt{\sigma^2(Y)} =$$
$$\sqrt{\sum_{i=1}^{n} p_i * (y_i - \bar{y})} = 0,2 * (1 - 2.25)^2 + 0.35 * (2 - 2.25)^2 + 0,45 * ($$
$$\sqrt{0,5875} = 0,77$$

Entonces,

$$\rho_{XY} = \frac{0{,}0775}{\sigma(X)\sigma(Y)} = \frac{0{,}0775}{0{,}48*0{,}77} \approx 0,21$$

Correlación

Ejemplo



Covarianza y Correlación

Ejercicio

Considera los siguientes datos de las v.a.'s X e Y:

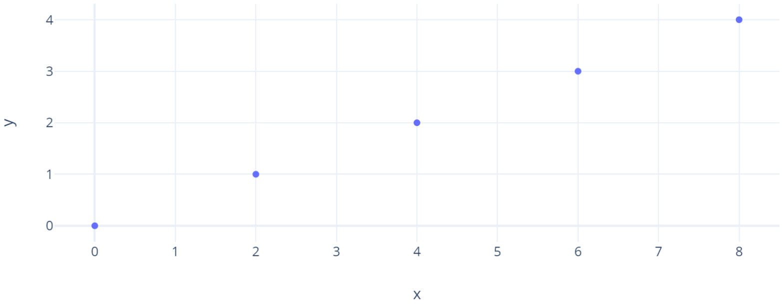
x	y
0	0
2	1
4	2
6	3
8	4

Calcula el coeficiente de correlacion entre X e Y e interpreta.

Covarianza y Correlación

Ejercicio

- ✓ Sabemos que $\rho_{XY} = \frac{cov_{XY}}{\sigma_X \sigma_Y}$
- ✓ Primeramente calculamos las medias, E(x) = 4 y E(y) = 2.
- ✓ Después calculamos las varianzas $\sigma(x) = 10$ y $\sigma(y) = 2.5$
- ✓ Y finalmente calculamos la covarianza, $Cov_{xy} = 5$ .Así, nuestro coeficiente de correlación es 1. Esto significa que ambas variables estan perfectamente corrrelacionadas (linealmente). De hecho: $y = \frac{x}{2}$



Contenidos

1. Variable Aleatoria
2. Variable Aleatoria Discreta
   - Función de probabilidad (PMF)
   - Función de probabilidad acumulada (CMF)
   - Esperanza, Varianza y Desviación Típica
3. Variable Aleatoria Bidimensional Discreta
   - Función de probabilidad Conjunta (PMF)
   - Correlación y Covarianza
4. Teorema de la Probabilidad Total y Teorema de Bayes

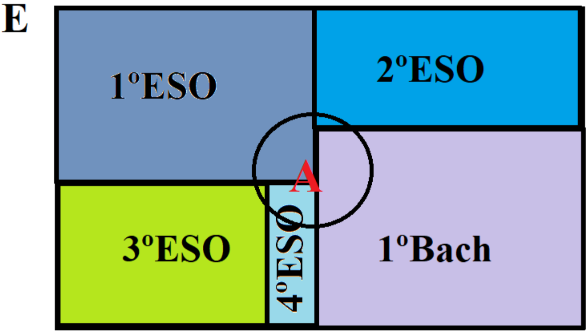
Teorema de la Probabilidad Total

Dados un suceso A y n sucesos $S_i disjuntosdosados$, $(S_i \cap S_j = \emptyset)$ tales que:

$$\cup_{i=1}^{n} S_i = E \text{ y } A \cap S_i \neq \emptyset,$$

entonces para cualquier suceso A se verifica que:

$$P(A) = \sum_{i=1}^{n} P(A|S_i) * P(S_i)$$



Teorema de la Probabilidad Total

Ejemplo

Tres máquinas M1, M2 y M3 fabrican 300, 450 y 600 piezas respectivamente. Los porcentajes de piezas defectuosas son 2%; 3,5% y 2,5% para cada una de las máquinas.De la producción total de las 3 máquinas al final de la jornada se toma una pieza al azar. Calcular la probabilidad de que sea defectuosa.

Teorema de la Probabilidad Total

Ejemplo

Tres máquinas M1, M2 y M3 fabrican 300, 450 y 600 piezas respectivamente. Los porcentajes de piezas defectuosas son 2%; 3,5% y 2,5% para cada una de las máquinas. De la producción total de las 3 máquinas al final de la jornada se toma una pieza al azar. Calcular la probabilidad de que sea defectuosa.

Según el Teorema de la Probabilidad Total,
$P(D) = P(D|M1) * P(M1) + P(D|M2) * P(M2) + P(D|M3) * P(M3) = ... = \frac{45.75}{1350}$

Resulta evidente al tabular:

Maquina	Piezas bien	Piezas defectuosas	Total
M1	294	6	300
M2	434.25	15.75	450
M3	585	15	600
Total	1304	45.75	1350

Teorema de Bayes

Fórmula del Teorema de Bayes

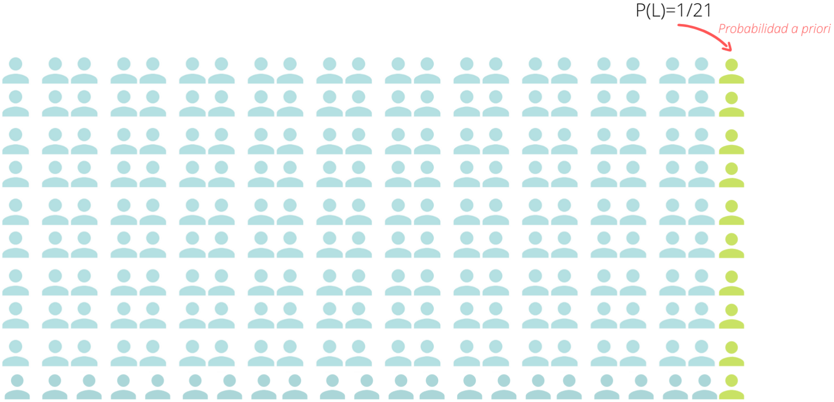
Sean n sucesos S_i con $P(S_i) > 0$ disjuntos dos a dos y tal que la unión de los n sucesos S_i constituyen el espacio muestral E .
Si se realiza otro suceso A con $P(A) > 0$ tal que $A \cap S_i \neq \emptyset$ y siendo conocidas las probabilidades de $(A|S_i)$, se verifica:

$$P(S_i|A) = \frac{P(A|S_i)*P(S_i)}{\sum_i^n P(A|S_i)*P(S_i)}$$

Teorema de Bayes

Intuición

Vamos a suponer que se nos presenta a un individuo, Steve, de una muestra que contiene a individuos Libreros y Granjeros en un ratio 1:21. Nuestra creencia (o hipotesis) es que Steve es efectivamente librero.



Teorema de Bayes

Intuición

Sin embargo, luego aparece una evidencia nueva:

Descripción de Steve

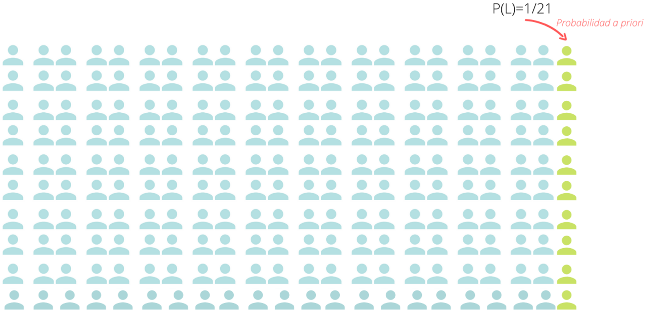
"Steve es muy tímido y retraído, siempre servicial, pero poco interesado por la gente o por el mundo real. De carácter disciplinado y metódico, necesita ordenarlo y organizarlo todo, y tiene obsesión por el detalle" . ¿Es probable que Steve sea un bibliotecario o un agricultor?

Para quien quiera leer más sobre Steve:

Kahneman y Tversky

Teorema de Bayes

Intuición



Teorema de Bayes

Ejercicio

Carlos se casa mañana en una ceremonia la aire libre en en el desierto. En los últimos años, solo ha llovido 5 días cada año. Desafortunadamente, el hombre del tiempo ha predicho lluvia para mañana. Cuando llueve, el hombre del tiempo predice correctamente el 90% del tiempo pero también predice lluvia cuando no llueve el 10% del tiempo. ¿Cuál es la probabilidad de que llueva en la boda de Carlos?

Teorema de Bayes

Ejercicio

- L = lluvia
- P = predicción de lluvia
  - $P(L) = 5/365 = 1/73$
  - $P(P|L) = 0.9$
  - $P(P|\bar{L}) = 0.1$

La probabilidad de lluvia incluyendo la nueva info (ha habido una predicción de lluvia):

$$P(L|P) = \frac{P(L)P(P|L)}{P(L)P(P|L)+P(P|\bar{L})P(\bar{L})}$$

$$= \frac{P(L)P(P|L)}{P(P)} = 1/9 \approx 11\%$$