ESTADÍSTICA

V.A. discreta parte I

Índice

- Variable aleatoria discreta
    * Función de probabilidad ó densidad
    * Función de distribución (acumulada)
- Esperanza matemática ó media de v.a. discreta
- Varianza y Desv. típica de v.a. discreta
- Transformaciones de v.a. discretas
- Variable aleatoria Bidimensional

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:

$$X : E \rightarrow \Re$$

$$w \rightarrow X(w)$$

Llamar variable a una función resulta algo confuso, por ello hay que insistir en que *es una función*.

La variable aleatoria puede ser discreta o continua. Veremos en este capítulo el caso discreto.

Función de probabilidad o densidad

Una vez definida una variable aleatoria *X*, podemos definir una **función de probabilidad** o **densidad** asociada a *X*, de la siguiente forma:

$$f : \Re \rightarrow [0,1]$$

$$x \rightarrow f(x) = P(X = x)$$

La función de probabilidad debe cumplir:

$$(i)\, 0 \leq f(x) \leq 1 \quad \forall x \in \Re$$

$$(ii) \sum_{x} f(x) = 1$$

(Suma sobre todos los posibles valores que puede tomar la variable aleatoria).

Requerimientos de una distribución de probabilidad

X	P(X)
-1	.1
0	.2
1	.4
2	.2
3	.1
	1.0

X	P(X)
-1	-.1
0	.3
1	.4
2	.3
3	.1
	1.0

X	P(X)
-1	.1
0	.3
1	.4
2	.3
3	.1
	1.2

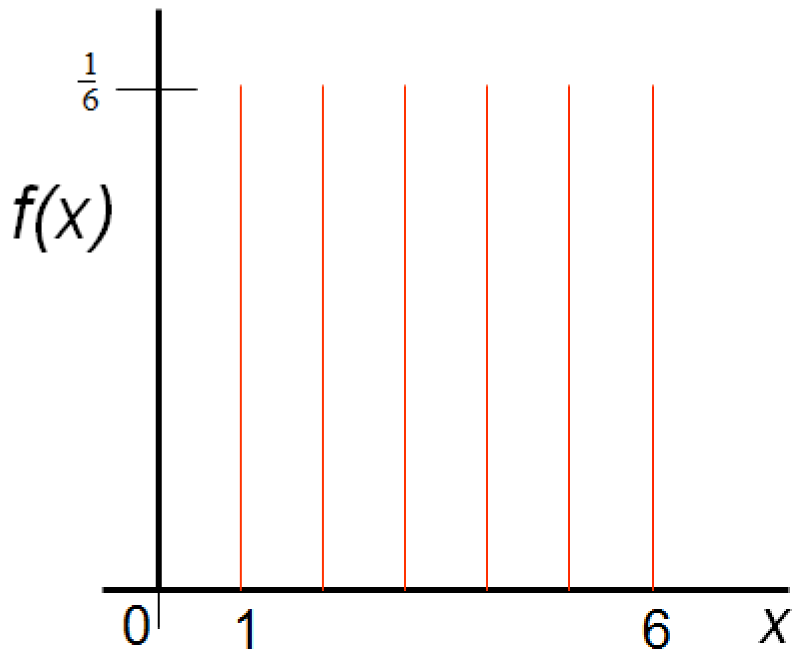
$$0 \leq f(x) \leq 1 \quad \text{para todo x}$$

$$\sum_{x} f(x) = 1$$

Dibuja la función de probabilidad f(x)

Realicemos el **experimento aleatorio** de lanzar un dado.

X tiene como posibles valores *x* = 1, 2, 3, 4, 5, 6 cada uno con probabilidad 1/6 (está **uniformemente distribuida**)



Función de probabilidad f(x)

Realicemos el siguiente **experimento aleatorio**:
lanzar una moneda dos veces

Espacio muestral

Ley de correspondencia

#caras

Valores	Probabilidad
0	1/4 = 0.25
1	2/4 = 0.50
2	1/4 = 0.25

$$X : E \rightarrow \Re$$
$$w \rightarrow X(w)$$

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

Realicemos el siguiente **experimento aleatorio**:

Lanzar 1 moneda, 3 veces.

Elementos del espacio muestral

+++ ++C +C+ C++ CC+ C+C +CC CCC

Ley de correspondencia

Nº reales (# de caras)

0 1 2 3 caras

$$X : E \rightarrow \Re$$

$$w \rightarrow X(w)$$

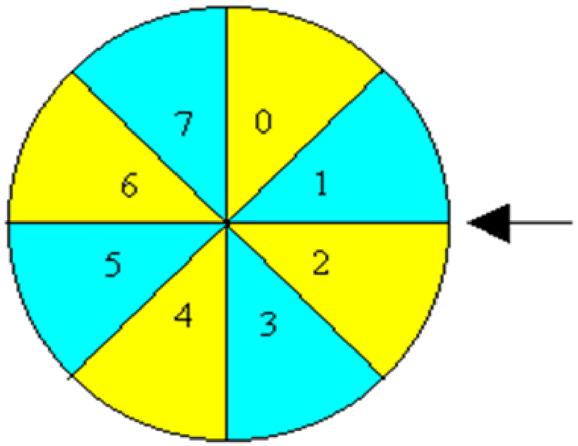
Probabilidad

1/8 3/8 3/8 1/8

$$f : \Re \rightarrow [0,1]$$

$$x \rightarrow f(x) = P(X = x)$$

Realicemos el siguiente **experimento aleatorio**: girar la ruleta de la imagen y apuntar el número del sector que coincide con la flecha.



La variable aleatoria *X* de este experimento asocia cada sector a un número entero, como podemos observar en la imagen.

Es una variable aleatoria discreta.
Los resultados posibles son: {0, 1, 2, 3, 4, 5, 6, 7}.
Por simetría podemos establecer una función de probabilidad: la probabilidad de cada resultado es 1/8.

Por tanto, la variable aleatoria *X* en el caso de esta ruleta está **uniformemente distribuida**, ya que todos los resultados tienen la misma probabilidad

Función de distribución (acumulada)

Dada una variable aleatoria discreta *X* se llama **función de distribución** a la función *F* definida como:

$$F:\Re \to [0,1]$$

$$x \to F(x) = P(X \leq x)$$

Además se puede obtener **la inversa de la función de distribución** que será F^{-1} definida como:

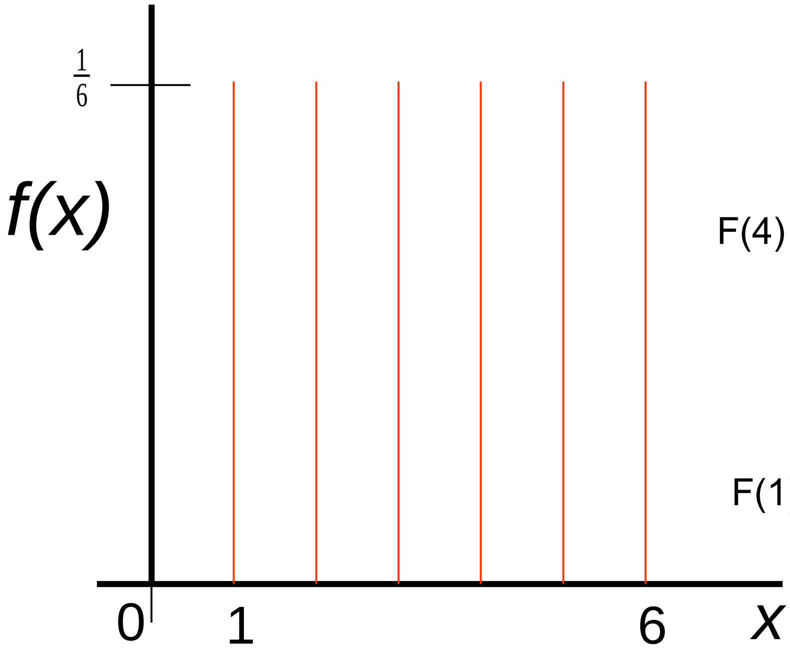
$$F^{-1}:\Re \leftarrow [0,1]$$

$$x \leftarrow F(x) = P(X \leq x)$$

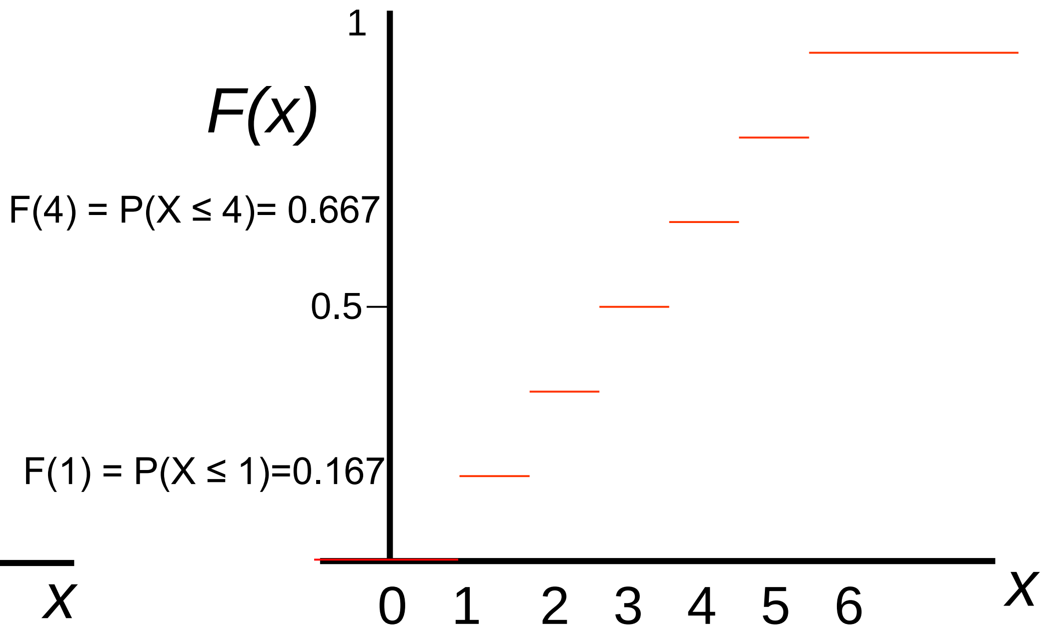
Dibuja la función de distribución F(x)

Experimento aleatorio de lanzar un dado (continuación)

X tiene como posibles valores *x* = 1, 2, 3, 4, 5, 6 cada uno con probabilidad 1/6 (está **uniformemente distribuida**)

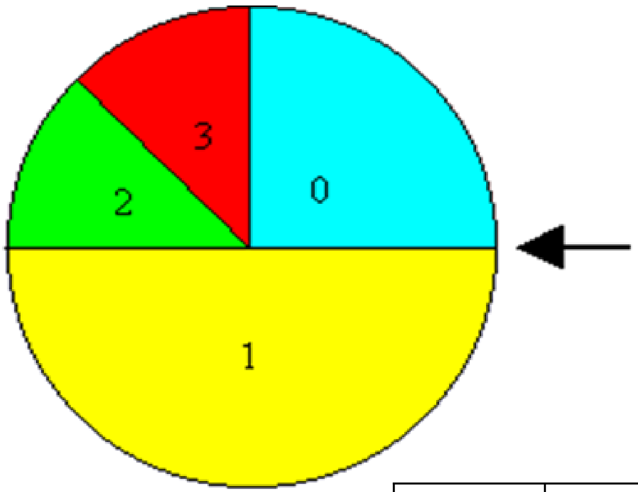


Función de probabilidad f(x)



Función de distribución F(x)

Realizamos un **experimento aleatorio** con esta nueva ruleta. Obtener su función de distribución F(x) y su inversa. Dibuja F(x).

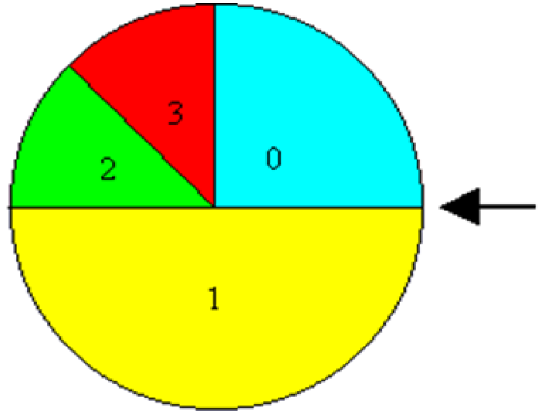


El espacio muestral está compuesto por 4 eventos.

La probabilidad es proporcional al ángulo del sector

	Función de probabilidad	**P. Acumulada ó Función de distribución**
x_i	$f(x) = P(X=x_i)$	$F(x) = P(X \le x_i)$
0	0.25	0.25
1	0.5	0.75
2	0.125	0.875
3	0.125	1

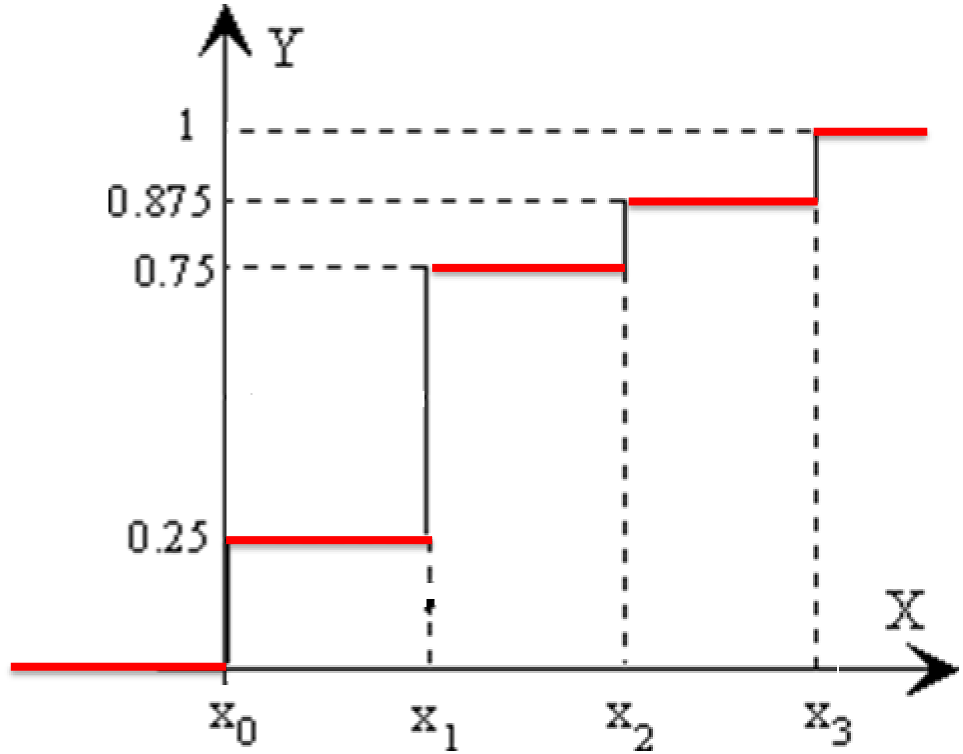
x_i	$F(x) = P(X \leq x_i)$
0	0.25
1	0.75
2	0.875
3	1



Función de distribución inversa

$\gamma = P(X \leq valor)$	$valor = x_i$
$0 \leq \gamma < 0.25$	0
$0.25 \leq \gamma < 0.75$	1
$0.75 \leq \gamma < 0.875$	2
$0.875 \leq \gamma < 1$	3

Gráfico de la función de distribución F(x)



Sea el **experimento aleatorio** “lanzar dos dados uno rojo y uno azul”. Dibuja la función de probabilidad y de distribución.

Luego, el espacio muestral E es como sigue:

E = {(1,1),(1,2),...(1,6),....,(5,6),(6,6)}

Definamos la variable aleatoria discreta *X, como*

X = “Suma de puntos al lanzar dos dados rojo y azul”

$$X : E \rightarrow S \in \Re$$ con ***S = { 2, 3, ...,12 }*** *la suma de puntos.*

Luego, su función de probabilidad es:

$$f : \Re \rightarrow [0,1]$$

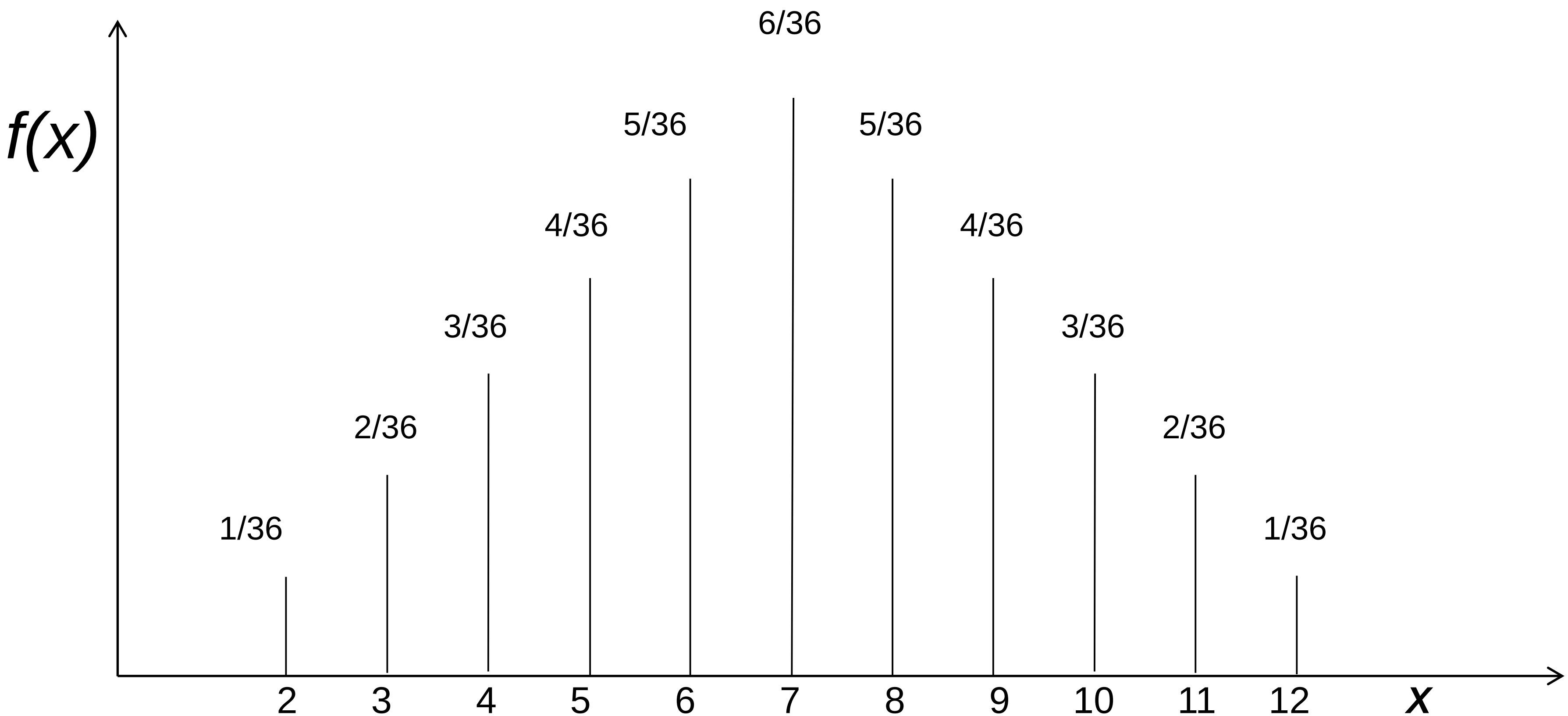
f(2) = P (X = 2) = P((1,1)) = 1/36

f(3) = P (X = 3) = P((1,2) U (2,1)) = 2/36

f(4) = P (X = 4) = P((1,3) U (3,1) U (2,2)) = 3/36

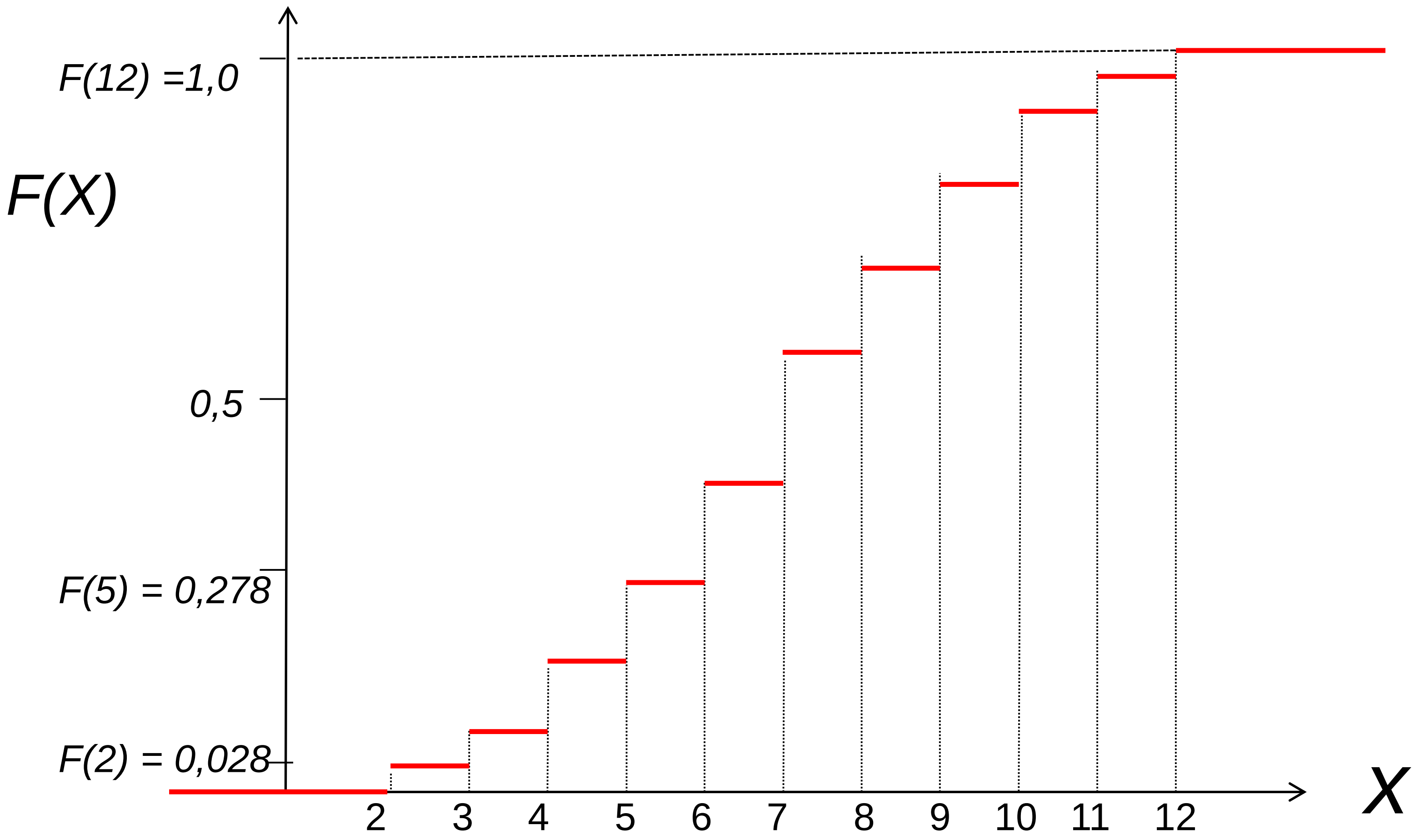
….

Función de probabilidad de la variable aleatoria X



Observa que cumple las dos condiciones: es siempre positiva (y menor o igual a 1) y está normalizada.

Función de distribución de la variable aleatoria X



$P(X \leq 5) = P(x = 2 \text{ o } x = 3 \text{ o } x = 4 \text{ o } x = 5) = 1/36 + 2/36 + 3/36 + 4/36 = 10/36 = 0.278 = F(5)$

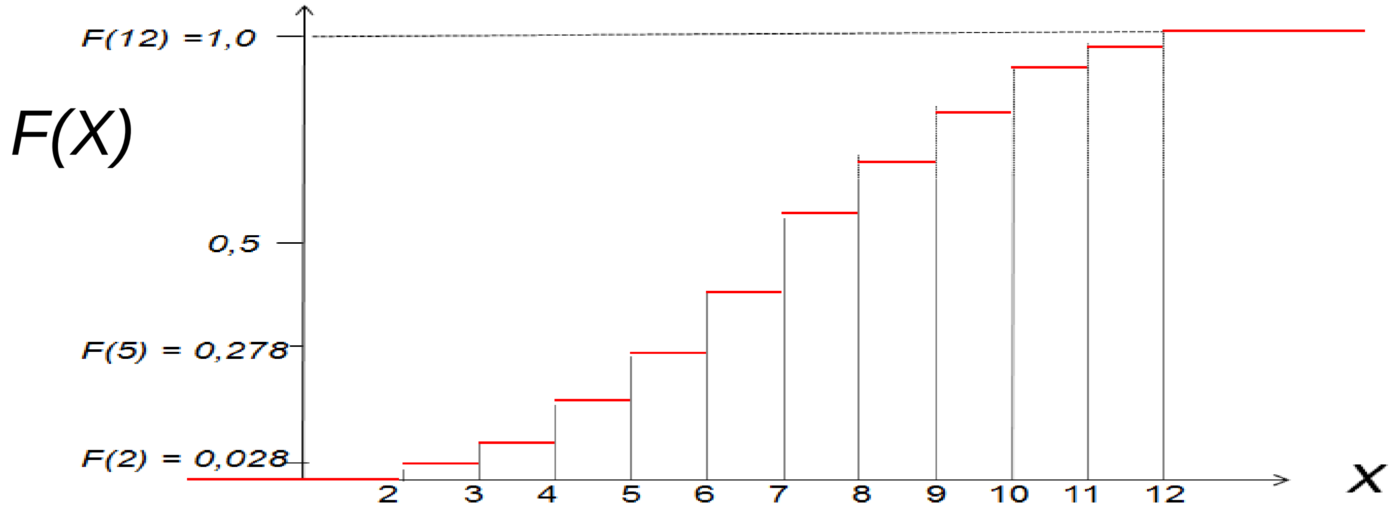
$P(X < 5) = P(X \leq 4) = P(x = 2 \text{ o } x = 3 \text{ o } x = 4) = 1/36 + 2/36 + 3/36 = 6/36 = 0.167 = F(4)$

$P(X > 5) = 1 - P(X \leq 5) = 1 - 0.278 = 0.722 = 1 - F(5)$

$P(X \geq 5) = 1 - P(X < 5) = 1 - P(X \leq 4) = 1 - 0.167 = 0.833 = 1 - F(4)$

Algunos problemas de probabilidad están relacionados con la probabilidad $P(a < X \leq b)$ de que X asuma algún valor en un intervalo $(a, b]$. Observa que:

$$P(a < X \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a)$$



En el ejemplo de los dos dados, calcula la probabilidad de que los dos dados sumen al menos 4 pero no más de 8:

$P(4 \leq X \leq 8) = P(3 < X \leq 8) = \mathrm{P}(\mathrm{X} \leq 8) - \mathrm{P}(\mathrm{X} \leq 3) = F(8) - F(3) = 26/36 - 3/36 = 23/36$

Esperanza matemática o **media** de una función de probabilidad discreta

$$\mu = E\left(X\right) = \sum_i x_i P(X = x_i) = \sum_i x_i f(x_i)$$

Momento de orden k

Los momentos de orden k <u>centrados en el origen</u>:

$$m_k = E(X^k) = \sum_i x_i^k f(x_i)$$

Obseva que: $m_1 = E(X) = \mu$

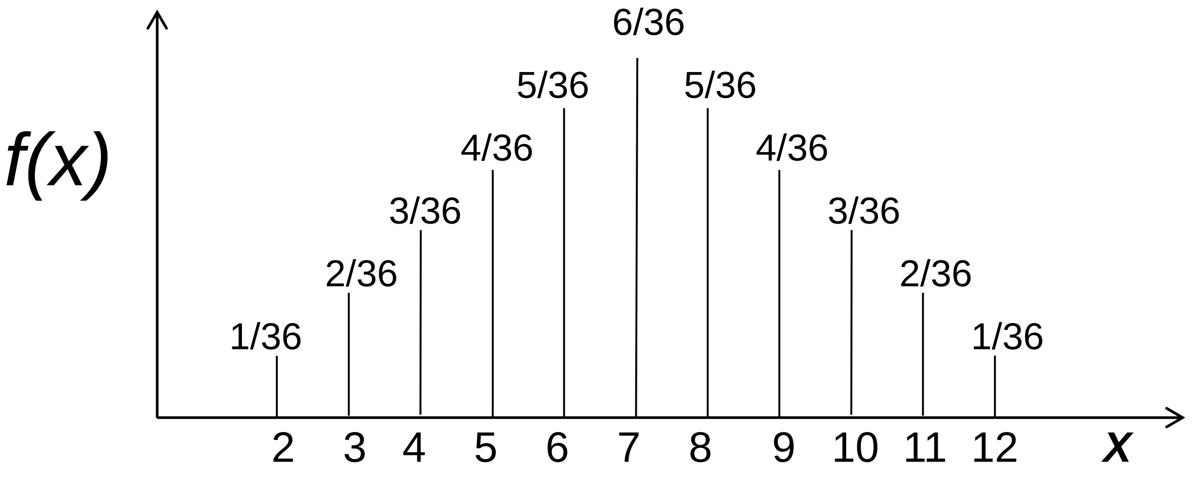
Ejemplo. Probabilidades en una tabla

x_i	$P(X=x_i)$	$x_i f(x_i)$
-1	0.1	-0.1
0	0.2	0.0
1	0.4	0.4
2	0.2	0.4
3	0.1	0.3

$$E(X) = \sum_i x_i f(x_i) = -0.1 + 0.0 + 0.4 + 0.4 + 0.3 = 1 = \mu$$

Experimento: lanzar dos dados, uno rojo y otro azul (continuación)

Calcular la esperanza de la variable aleatoria *X*



$$E(X)=\sum_{i=2}^{12} f(x_i)\cdot x_i =$$

$$\frac{1}{36}\cdot 2+\frac{2}{36}\cdot 3+\ldots+\frac{6}{36}\cdot 7+\ldots+\frac{1}{36}\cdot 12 = 7 = \mu$$

Ejercicio: juegos

Decía Albert Einstein (1879 – 1955): “La mejor forma de ganar dinero en un casino es asaltándolo”. Con humor, recor-

Sea el juego que consiste en sacar una bola de una urna que contiene 7 bolas rojas y 3 bolas negras. Ganamos 50 euros si la bola extraída es roja y pagamos 150 euros en el caso de que sea negra. ¿Qué podemos esperar si jugamos muchas veces?

Si la esperanza matemática es 0 se dice que el **juego** es **justo.**
Si es mayor que 0 se dice que el **juego** es **favorable** al jugador.
Si es menor que 0 se dice que perjudica al jugador y **no** es **favorable**.

Espacio muestral $E = \{R, N\}$. Consideramos las ganancias como positivas y las pérdidas negativas:

Variable aleatoria X		Función de probabilidad
$R \longrightarrow$	$50 \longrightarrow$	0,7
$N \longrightarrow$	$-150 \longrightarrow$	0,3

$$\mu = 50 \cdot 0{,}7 + (-150) \cdot 0{,}3 = -10$$

Ganancia media

Varianza y desviación estándar o típica de una función de probabilidad discreta

Varianza

$$\sigma^2 = Var(X) = E((X-\mu)^2) = \sum_i (x_i - \mu)^2 f(X = x_i)$$

Desviación estándar o típica

$$\sigma = \sqrt{Var(X)}$$

Ambas miden la "dispersión de los datos". Observa que la desviación típica lo hace con las mismas unidades que los propios datos.

Ejemplo. *Probabidades en una tabla (continuación)*

Sabemos que μ=1

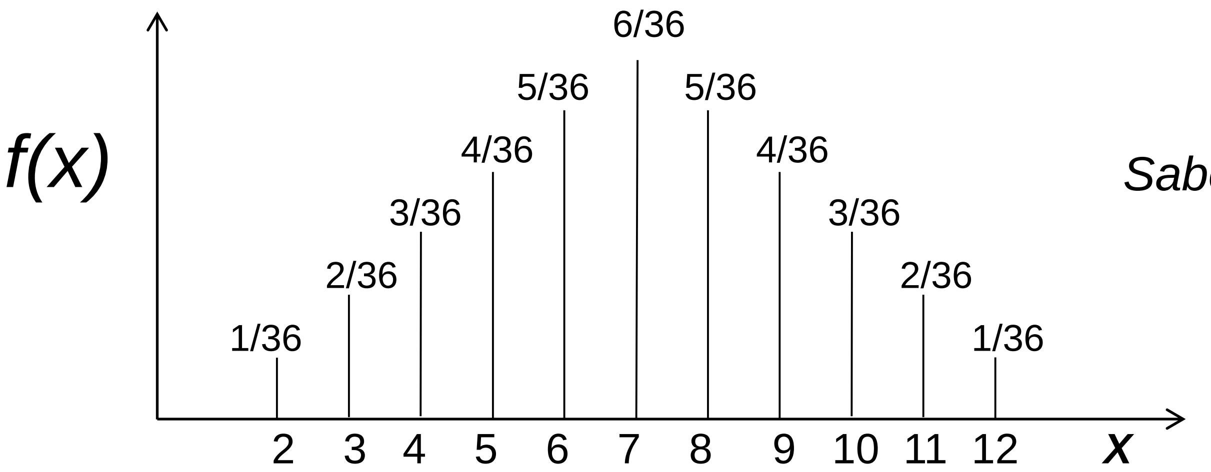
x_i	$P(X=x_i) = f(x)$	$x_i - \mu$	$(x_i - \mu)^2$	$(x_i - \mu)^2 \cdot f(x)$
-1	.1	-2	4	.4
0	.2	-1	1	.2
1	.4	0	0	.0
2	.2	1	1	.2
3	.1	2	4	.4
				1.2

$$\sigma^2 = \sum_i (x_i - \mu)^2 f(x_i) = 1{,}2$$

$$\sigma = \sqrt{Var(X)} = 1.10$$

Experimento: *lanzar dos dados, uno rojo y otro azul (continuacion)*

Calcula la varianza y desv. típica de la v.a. X



Sabemos que μ=7

$$\mathrm{Var}(X) = \sum_{i=2}^{12} f(x_i) \cdot (x_i - 7)^2 =$$

$$\frac{1}{36} \cdot (2-7)^2 + \frac{2}{36} \cdot (3-7)^2 + \ldots + \frac{1}{36} \cdot (12-7)^2 = 5{,}83$$

$$\sigma = \sqrt{Var(X)} = \sqrt{5{,}83} = 2{,}41$$

Algunas propiedades de la varianza

Momento de orden k

Los momentos de orden k centrados en la media de X:

$$\boxed{M_k = E((X-\mu)^k) = \sum_i (x_i - \mu)^k f(x_i)}$$

Obseva que: $M_1 = E(X-\mu) = E(X) - \mu = 0$

$$M_2 = E((X-\mu)^2) = Var(X)$$

$$\boxed{\sigma^2 = E(X^2) - (E(X))^2}$$

Ejemplo. *Probabidades en una tabla (continuación)*
Sabemos que E(X)=μ=1

x_i	$P(X=x_i) = f(x)$
-1	.1
0	.2
1	.4
2	.2
3	.1

$$E(X^2) = \sum_{i=1}^{5} f(x_i) \cdot x_i^2 =$$

$$0.1*(-1)^2 + 0.2*0^2 + 0.4*1^2 + 0.2*2^2 + 0.1*3^2 = 2.2$$

$$\boxed{\sigma^2 = E(X^2) - (E(X))^2 = 2.2 - 1^2 = 1.2}$$

Transformaciones de variables aleatorias

Ejemplo.

X número de hijos por familia

x_i	$P(X=x_i) = f(x)$
0	0.47
1	0.3
2	0.1
3	0.06
4	0.04
5	0.02
6	0.01

Y = 15X

Y ingreso por hijo: Y = 15X

Y_i	$P(Y=y_i) = f(y)$
0	0.47
15	0.3
30	0.1
45	0.06
60	0.04
75	0.02
90	0.01

$$\mu_x = E(X) = \Sigma\, xf(x) = 1$$

$$\sigma_x^2 = V[X] = E(X^2) - (E(X))^2 = 2.74 - 1^2 = 1.74$$

$$\mu_y = E(Y) = \Sigma\, yf(y) = 15 = 15 * E(X)$$

$$\sigma_y^2 = V[Y] = E(Y^2) - (E(Y))^2 = 616.5 - 15^2 = 391.5 = 15^2 * V[X]$$

Sean a y b constantes.

(1) $E(aX + b) = aE(X) + b$

(2) $V(aX + b) = a^2 V(X)$

(3) $E(X + W) = E(X) + E(W)$

(4) $\sigma^2 = V(X) = E(X^2) - (E(X))^2$

Variable aleatoria Bidimensional

- Vector Aleatorio: $\bar{A} = (X, Y)$

- **F. densidad Conjunta:** $f(x,y)$

- **F. densidad Marginales:** Para **g(x)** se suma la probabilidad de cada categoría en Y. Análogo para **h(y)**.
  y Condicionadas: Para **$Y/X=x_i$** sería $f(x_j,y_i)/\ g(X=x_i)$. Análogo para **$X/Y=y_i$**

- **Independencia de V.A.:** Las variables aleatorias X, Y son independientes si

$$\boxed{f(x,y) = g(x) \cdot h(y)}$$

$$\boxed{\sigma_X^2 = \Sigma\, x^2 g(x) - E[X]E[X]}$$

$$\boxed{\sigma_{XY} = \Sigma\, xyf(x,y) - E[X]E[Y]}$$

- **Vector de medias:** $\bar{\mu}_A = \begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix}$

- **Matriz de Varianzas, Covarianzas:** $\begin{pmatrix} \sigma_X^2 & \sigma_{XY} \\ \sigma_{YX} & \sigma_Y^2 \end{pmatrix}$

Coeficiente de correlación:

$$-1 < \rho = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} < 1$$

Ejemplo.

En un estudio sobre las preferencias de los habitantes de la ciudad de Pamplona. Sea X la variable que, en euros, expresa la renta familiar y sea Y otra variable que toma el valor 0 cuando el lugar de vacaciones elegido sea la playa y 1 si se elige la montaña. Como resultado del estudio se propone la siguiente distribución de probabilidad conjunta para las variables citadas:

Y\X	**6.000**	**18.000**	**36.000**	$P(Y=y_i) = h(y)$
0=playa	0'15	0'35	0'15	**0.65**
1=montaña	0'15	0'15	0'05	**0.35**
$P(X=x_i) = g(x)$	**0.3**	**0.5**	**0.2**	

a) ¿Son independientes X e Y?

b) Obtener la distribución condicionada de X/Y=0(playa)

c) Calcula la Covarianza(X,Y)

a) Independientes si f(x,y)= g(x)*h(y)

f(6000, montaña) = 0.15

g(6000) * h(montaña) = 0.3 * 0.35 = 0.105

Luego, NO independientes.

b) X/Y=0 (playa)

X_i	$P(X=x_i /Y=0)$
6000	0.15/0.65 = **0.2308**
18000	0.35/0.65 = **0.5384**
36000	0.15/0.65 = **0.2308**

c) Cov(X,Y) = σ_{XY}

Σ xy f(x,y) = 5400

E[X] = Σ x g(x) = 18000

E[Y] = Σ y h(y) = 0.35

Cov(X,Y) = Σ xy f(x,y) - E[X] E[Y]
= - 900