# Probabilidad y Estadística (m) / Probabilidad (d)

## Esperanza

Departamento de Matemática / Instituto de Cálculo

15 de mayo de 2024

# Outline

1. Esperanza
   - Esperanza de variables aleatorias discretas
   - Esperanza de variables aleatorias continuas
   - Propiedades de la esperanza
   - La esperanza como mejor predictor

2. Varianza

3. Covarianza
   - El espacio $\mathcal{L}^2$

# Outline

# Ejemplo: Album de figuritas

100 niños tienen un álbum de figuritas de 3 figuritas cada uno. Los paquetes que se venden en los kioscos tienen 2 figuritas. La cantidad de paquetes que tuvieron que comprar para llenar el album fue de entre 2 y 7. La cantidad de niños que tuvo que comprar cada cantidad de paquetes viene dado por la siguiente tabla:

| $i$ | 2 | 3 | 4 | 5 | 7 |
|---|---|---|---|---|---|
| $n_i$ | 57 | 25 | 15 | 2 | 1 |

La cantidad de paquetes promedio que los niños tuvieron que comprar fue

$$M = \frac{2 \cdot 57 + 3 \cdot 25 + 4 \cdot 15 + 5 \cdot 2 + 7 \cdot 1}{100} = 2\frac{57}{100} + 3\frac{25}{100} + 4\frac{15}{100} + 5\frac{2}{100} + 7\frac{1}{100}$$

$$= 2\frac{n_2}{n} + 3\frac{n_3}{n} + 4\frac{n_4}{n} + 5\frac{n_5}{n} + 6\frac{n_6}{n} + 7\frac{n_7}{n}$$

donde $n = 100$ y $n_i$ es la cantidad de niños que compró $i$ paquetes.

# Ejemplo: Album de figuritas

100 niños tienen un álbum de figuritas de 3 figuritas cada uno. Los paquetes que se venden en los kioscos tienen 2 figuritas. La cantidad de paquetes que tuvieron que comprar para llenar el album fue de entre 2 y 7. La cantidad de niños que tuvo que comprar cada cantidad de paquetes viene dado por la siguiente tabla:

| $i$ | 2 | 3 | 4 | 5 | 7 |
|---|---|---|---|---|---|
| $n_i$ | 57 | 25 | 15 | 2 | 1 |

La cantidad de paquetes promedio que los niños tuvieron que comprar fue

$$M = \frac{2\cdot 57 + 3\cdot 25 + 4\cdot 15 + 5\cdot 2 + 7\cdot 1}{100} = 2\frac{57}{100} + 3\frac{25}{100} + 4\frac{15}{100} + 5\frac{2}{100} + 7\frac{1}{100}$$

$$= 2\frac{n_2}{n} + 3\frac{n_3}{n} + 4\frac{n_4}{n} + 5\frac{n_5}{n} + 6\frac{n_6}{n} + 7\frac{n_7}{n}$$

donde $n = 100$ y $n_i$ es la cantidad de niños que compró $i$ paquetes.

# Ejemplo: Album de figuritas

100 niños tienen un álbum de figuritas de 3 figuritas cada uno. Los paquetes que se venden en los kioscos tienen 2 figuritas. La cantidad de paquetes que tuvieron que comprar para llenar el album fue de entre 2 y 7. La cantidad de niños que tuvo que comprar cada cantidad de paquetes viene dado por la siguiente tabla:

| $i$ | 2 | 3 | 4 | 5 | 7 |
|---|---|---|---|---|---|
| $n_i$ | 57 | 25 | 15 | 2 | 1 |

La cantidad de paquetes promedio que los niños tuvieron que comprar fue

$$M = \frac{2\cdot 57 + 3\cdot 25 + 4\cdot 15 + 5\cdot 2 + 7\cdot 1}{100} = 2\frac{57}{100} + 3\frac{25}{100} + 4\frac{15}{100} + 5\frac{2}{100} + 7\frac{1}{100}$$

$$= 2\frac{n_2}{n} + 3\frac{n_3}{n} + 4\frac{n_4}{n} + 5\frac{n_5}{n} + 6\frac{n_6}{n} + 7\frac{n_7}{n}$$

donde $n = 100$ y $n_i$ es la cantidad de niños que compró $i$ paquetes.

# Ejemplo: Album de figuritas

| $i$ | 2 | 3 | 4 | 5 | 7 |
|---|---|---|---|---|---|
| $n_i$ | 57 | 25 | 15 | 2 | 1 |

Sea $X =$ cantidad de figuritas que compra un niño hasta llenar el album y $p_X$ su función de probabilidad puntual.

Recordemos que $p_X(i)$ se interpreta como el límite cuando $n \to \infty$ de la frecuencia relativa del evento $\{X = i\}$ en $n$ repeticiones del experimento, es decir que si $n$ es grande

$$n_i/n \approx p_X(i)$$

Así, la cantidad promedio de paquetes que compraron los niños es aprox

$$M \approx 2p_X(1) + 3p_X(2) + 4p_X(4) + 5p_X(5) + 7p_X(6)$$

$$\approx \sum_{i \in R_X} i p_X(i).$$

La última aproximación se debe a que $n_i = 0$ para algunos $i$ pero $p_X(i) \neq 0$.

# Ejemplo: Album de figuritas

| $i$ | 2 | 3 | 4 | 5 | 7 |
|---|---|---|---|---|---|
| $n_i$ | 57 | 25 | 15 | 2 | 1 |

Sea $X =$ cantidad de figuritas que compra un niño hasta llenar el album y $p_X$ su función de probabilidad puntual.
Recordemos que $p_X(i)$ se interpreta como el límite cuando $n \to \infty$ de la frecuencia relativa del evento $\{X = i\}$ en $n$ repeticiones del experimento, es decir que si $n$ es grande

$$n_i/n \approx p_X(i)$$

Así, la cantidad promedio de paquetes que compraron los niños es aprox

$$M \approx 2p_X(1) + 3p_X(2) + 4p_X(4) + 5p_X(5) + 7p_X(6)$$

$$\approx \sum_{i \in R_X} ip_X(i).$$

La última aproximación se debe a que $n_i = 0$ para algunos $i$ pero $p_X(i) \neq 0$.

# Ejemplo: Album de figuritas

| $i$ | 2 | 3 | 4 | 5 | 7 |
|---|---|---|---|---|---|
| $n_i$ | 57 | 25 | 15 | 2 | 1 |

Sea $X =$ cantidad de figuritas que compra un niño hasta llenar el album y $p_X$ su función de probabilidad puntual.
Recordemos que $p_X(i)$ se interpreta como el límite cuando $n \to \infty$ de la frecuencia relativa del evento $\{X = i\}$ en $n$ repeticiones del experimento, es decir que si $n$ es grande

$$n_i/n \approx p_X(i)$$

Así, la cantidad promedio de paquetes que compraron los niños es aprox

$$M \approx 2p_X(1) + 3p_X(2) + 4p_X(4) + 5p_X(5) + 7p_X(6)$$

$$\approx \sum_{i \in R_X} i p_X(i).$$

La última aproximación se debe a que $n_i = 0$ para algunos $i$ pero $p_X(i) \neq 0$.

# Definición de esperanza de una variable discreta

## Definición 1

Dada una variable aleatoria discreta $X$ con $R_X = \{x_1, x_2, \cdots\}$ y función de probabilidad puntual $p_X(x_i)$, definimos la esperanza de $X$ mediante la fórmula,

$$E[X] = \sum_{i=1}^{\infty} x_i \, p_X(x_i) \,, \tag{1}$$

siempre que $\sum_{i=i}^{\infty} |x_i| \, p(x_i) < \infty$.

# Esperanza infinita

Denotemos con

$$R_X^+ = \{x \in R_X : x > 0\}$$
$$R_X^- = \{x \in R_X : x < 0\}\,.$$

Si $\sum_{x \in R_X} |x| p_X(x) < \infty$,

$$\sum_{x \in R_X} |x| p_X(x) = \sum_{x \in R_X^+} x p_X(x) - \sum_{x \in R_X^-} x p_X(x)$$

Si $\sum_{x \in R_X} |x| p_X(x) = +\infty$. Entonces pueden ocurrir tres casos distintos.

1. $\sum_{x \in R_X^+} x p_X(x) = +\infty$ y $\sum_{x \in R_X^-} x p_X(x) = -\infty$.
2. $\sum_{x \in R_X^+} x p_X(x) = +\infty$ y $\sum_{x \in R_X^-} x p_X(x) > -\infty$.
3. $\sum_{x \in R_X^+} x p_X(x) < +\infty$ y $\sum_{x \in R_X^-} x p_X(x) = -\infty$.

En el caso 1 no se puede definir la esperanza de $X$.
En el caso 2 se define $E(X) = +\infty$.
En el 3 $E(X) = -\infty$.

# Esperanza infinita

Denotemos con

$$R_X^+ = \{x \in R_X : x > 0\}$$
$$R_X^- = \{x \in R_X : x < 0\}.$$

Si $\sum_{x \in R_X} |x| p_X(x) < \infty$,

$$\sum_{x \in R_X} |x| p_X(x) = \sum_{x \in R_X^+} x p_X(x) - \sum_{x \in R_X^-} x p_X(x)$$

Si $\sum_{x \in R_X} |x| p_X(x) = +\infty$. Entonces pueden ocurrir tres casos distintos.

1. $\sum_{x \in R_X^+} x p_X(x) = +\infty$ y $\sum_{x \in R_X^-} x p_X(x) = -\infty$.
2. $\sum_{x \in R_X^+} x p_X(x) = +\infty$ y $\sum_{x \in R_X^-} x p_X(x) > -\infty$.
3. $\sum_{x \in R_X^+} x p_X(x) < +\infty$ y $\sum_{x \in R_X^-} x p_X(x) = -\infty$.

En el caso 1 no se puede definir la esperanza de $X$.
En el caso 2 se define $E(X) = +\infty$.
En el 3 $E(X) = -\infty$.

# Esperanza infinita

Denotemos con

$$R_X^+ = \{x \in R_X : x > 0\}$$
$$R_X^- = \{x \in R_X : x < 0\}.$$

Si $\sum_{x \in R_X} |x| p_X(x) < \infty$,

$$\sum_{x \in R_X} |x| p_X(x) = \sum_{x \in R_X^+} x p_X(x) - \sum_{x \in R_X^-} x p_X(x)$$

Si $\sum_{x \in R_X} |x| p_X(x) = +\infty$. Entonces pueden ocurrir tres casos distintos.

1. $\sum_{x \in R_X^+} x p_X(x) = +\infty$ y $\sum_{x \in R_X^-} x p_X(x) = -\infty$.
2. $\sum_{x \in R_X^+} x p_X(x) = +\infty$ y $\sum_{x \in R_X^-} x p_X(x) > -\infty$.
3. $\sum_{x \in R_X^+} x p_X(x) < +\infty$ y $\sum_{x \in R_X^-} x p_X(x) = -\infty$.

En el caso 1 no se puede definir la esperanza de $X$.
En el caso 2 se define $E(X) = +\infty$.
En el 3 $E(X) = -\infty$.

# Esperanza infinita

Denotemos con

$$R_X^+ = \{x \in R_X : x > 0\}$$
$$R_X^- = \{x \in R_X : x < 0\}.$$

Si $\sum_{x \in R_X} |x| p_X(x) < \infty$,

$$\sum_{x \in R_X} |x| p_X(x) = \sum_{x \in R_X^+} x p_X(x) - \sum_{x \in R_X^-} x p_X(x)$$

Si $\sum_{x \in R_X} |x| p_X(x) = +\infty$. Entonces pueden ocurrir tres casos distintos.

1. $\sum_{x \in R_X^+} x p_X(x) = +\infty$ y $\sum_{x \in R_X^-} x p_X(x) = -\infty$.
2. $\sum_{x \in R_X^+} x p_X(x) = +\infty$ y $\sum_{x \in R_X^-} x p_X(x) > -\infty$.
3. $\sum_{x \in R_X^+} x p_X(x) < +\infty$ y $\sum_{x \in R_X^-} x p_X(x) = -\infty$.

En el caso 1 no se puede definir la esperanza de $X$.
En el caso 2 se define $E(X) = +\infty$.
En el 3 $E(X) = -\infty$.

# Esperanza infinita

Denotemos con

$$R_X^+ = \{x \in R_X : x > 0\}$$
$$R_X^- = \{x \in R_X : x < 0\}\,.$$

Si $\sum_{x \in R_X} |x| p_X(x) < \infty$,

$$\sum_{x \in R_X} |x| p_X(x) = \sum_{x \in R_X^+} x p_X(x) - \sum_{x \in R_X^-} x p_X(x)$$

Si $\sum_{x \in R_X} |x| p_X(x) = +\infty$. Entonces pueden ocurrir tres casos distintos.

1. $\sum_{x \in R_X^+} x p_X(x) = +\infty$ y $\sum_{x \in R_X^-} x p_X(x) = -\infty$.
2. $\sum_{x \in R_X^+} x p_X(x) = +\infty$ y $\sum_{x \in R_X^-} x p_X(x) > -\infty$.
3. $\sum_{x \in R_X^+} x p_X(x) < +\infty$ y $\sum_{x \in R_X^-} x p_X(x) = -\infty$.

En el caso 1 no se puede definir la esperanza de $X$.
En el caso 2 se define $E(X) = +\infty$.
En el 3 $E(X) = -\infty$.

# Ejemplo

## Ejemplo 2

Se tira un dado. Se ganan tantos puntos como indica el dado.

$X$ = puntos obtenidos en el juego

$R_X = \{1, 2, 3, 4, 5, 6\}$

$$E(X) = \frac{1}{6} + 2\frac{1}{6} + 3\frac{1}{6} + 4\frac{1}{6} + 5\frac{1}{6} + 6\frac{1}{6} = 3{,}5$$

# Ejemplo

## Ejemplo 3

Se tira un dado. Si sale par se ganan 2 pesos. Si sale 1,3 o 5 se pierden 1,3 o 5 pesos respectivamente. Hallar la ganancia esperada.

$X =$ ganancia obtenida en el juego

$R_X = \{-1, -3, -5, 2\}$

Función de probabilidad puntual:

| $i$ | -1 | -3 | -5 | 2 |
|---|---|---|---|---|
| $p_X(i)$ | 1/6 | 1/6 | 1/6 | 1/2 |

$$E(X) = -\frac{1}{6} - 3\frac{1}{6} - 5\frac{1}{6} + 2\frac{1}{2} = -\frac{1}{2}$$

# Ejemplo

## Ejemplo 3

Se tira un dado. Si sale par se ganan 2 pesos. Si sale 1,3 o 5 se pierden 1,3 o 5 pesos respectivamente. Hallar la ganancia esperada.

$X =$ ganancia obtenida en el juego
$R_X = \{-1, -3, -5, 2\}$
Función de probabilidad puntual:

| $i$ | -1 | -3 | -5 | 2 |
|---|---|---|---|---|
| $p_X(i)$ | 1/6 | 1/6 | 1/6 | 1/2 |

$$E(X) = -\frac{1}{6} - 3\frac{1}{6} - 5\frac{1}{6} + 2\frac{1}{2} = -\frac{1}{2}$$

# Ejemplo

## Ejemplo 3

Se tira un dado. Si sale par se ganan 2 pesos. Si sale 1,3 o 5 se pierden 1,3 o 5 pesos respectivamente. Hallar la ganancia esperada.

$X =$ ganancia obtenida en el juego
$R_X = \{-1, -3, -5, 2\}$
Función de probabilidad puntual:

| $i$ | -1 | -3 | -5 | 2 |
|---|---|---|---|---|
| $p_X(i)$ | 1/6 | 1/6 | 1/6 | 1/2 |

$$E(X) = -\frac{1}{6} - 3\frac{1}{6} - 5\frac{1}{6} + 2\frac{1}{2} = -\frac{1}{2}$$

# Ejemplo

## Ejemplo 3

Se tira un dado. Si sale par se ganan 2 pesos. Si sale 1,3 o 5 se pierden 1,3 o 5 pesos respectivamente. Hallar la ganancia esperada.

$X$ = ganancia obtenida en el juego
$R_X = \{-1, -3, -5, 2\}$
Función de probabilidad puntual:

| $i$ | -1 | -3 | -5 | 2 |
|---|---|---|---|---|
| $p_X(i)$ | 1/6 | 1/6 | 1/6 | 1/2 |

$$E(X) = -\frac{1}{6} - 3\frac{1}{6} - 5\frac{1}{6} + 2\frac{1}{2} = -\frac{1}{2}$$

# Ejemplo

## Ejemplo 4

Sea $X$ que toma los valores $-1, 0, 1$ con probabilidades $\frac{1}{3}$ cada uno, tenemos

$$E[X] = 0$$

Sea $Y = X^2 = g(X)$, toma valores 0 y 1 , con probabilidades $\frac{1}{3}$ y $\frac{2}{3}$ respectivamente.
Entonces

$$E[Y] = 0 \cdot \frac{1}{3} + 1\frac{2}{3} = \frac{2}{3}$$

# Esperanza de variables discretas famosas

**Proposición 1.1**

1. *Si* $X \sim B(1,p)$ *entonces* $E(X) = p$
2. *Si* $X \sim B(n,p)$ *entonces* $E(X) = np$
3. *Si* $X \sim \mathcal{G}(p)$*, es decir,* $p_X(k) = p(1-p)^{k-1}$ *para* $k \in \mathbb{N}$*, entonces* $E(X) = 1/p$.
4. *Si* $X \sim \mathcal{P}(\lambda)$ *entonces* $E(X) = \lambda$

# Ejemplo:Poisson

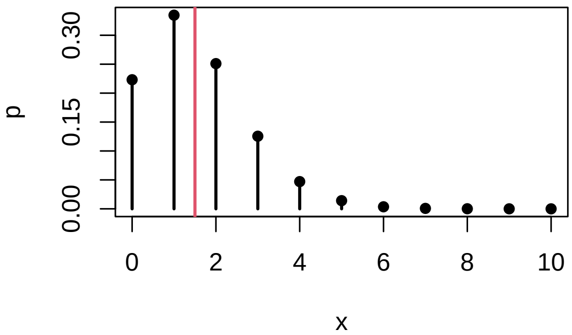



Figura: Función de probabilidad puntual y esperanza de una v.a. $\mathcal{P}(1{,}5)$

# La esperanza como centro de masa

El concepto de esperanza es análogo al concepto físico del centro de gravedad de una distribución de masa (o centro de masa). Para la variable $X$ que toma los valores $k_i$ con probabilidad $p_X(k_i)$, imaginamos el eje $x$ como una barra que no tiene masa y sobre el ubicamos en los puntos $k_i$ un peso con masa $p_X(k_i)$ (masa total 1), entonces el punto en el cual al apoyar la barra esta estaría en equilibrio e el centro de gravedad y coincide con $E(X)$.

## Teorema 5

*Sea $X$ una variable aleatoria. Sea $g : \mathbb{R} \to \mathbb{R}$ una función boreliana. Consideremos la variable aleatoria $Y = g(X)$. Entonces, si $X$ es discreto y*

$$\sum_{x \in R_X} |g(x)|\, p_X(x) < +\infty$$

*entonces*

$$E[g(X)] = \sum_{x \in R_X} g(x)\, p_X(x)$$

## Demostración.

Pizarrón $\square$

## Teorema 6

*Sea* $(X_1, X_2, \ldots, X_k)$ *un vector aleatorio k-dimensional. Sea* $g : \mathbb{R}^k \to \mathbb{R}$ *una función boreliana. Consideremos la variable aleatoria* $Y = g(X_1, X_2, \ldots, X_k)$*. Entonces, si* $(X_1, X_2, \ldots, X_k)$ *es discreto y*

$$\sum_{x_1, x_2, \cdots, x_k} |g(x_1, x_2, \ldots, x_k)| \, p_{X_1 X_2 \ldots X_k}(x_1, x_2, \ldots, x_k) < +\infty$$

*entonces*

$$E[g(X_1, \ldots, X_k)] = \sum_{x_1, \cdots, x_k} g(x_1, \ldots, x_k) \, p_{X_1 X_2 \ldots X_k}(x_1, \ldots, x_k)$$

## Demostración.

Análoga al caso $k = 1$ □

## Ejemplo 7

Sea $X$ que toma los valores $-1, 0, 1$ con probabilidades $\frac{1}{3}$ cada uno, tenemos

$$E[X] = 0$$

Sea $Y = X^2 = g(X)$, toma valores 0 y 1 , con probabilidades $\frac{1}{3}$ y $\frac{2}{3}$ respectivamente, Habíamos calculado la $E[Y]$ calculando primero la $p_Y$. Hagámoslo ahora directamente usando la fórmula del Teorema tenemos

$$E[Y] = g(-1) \cdot \frac{1}{3} + g(0) \cdot \frac{1}{3} + g(1) \cdot \frac{1}{3} = \frac{1}{3} + \frac{1}{3} = \frac{2}{3}$$

# Ejemplo

## Ejemplo 8

Me invitan a jugar el siguiente juego de azar. Se tira un dado. Se gana tantos pesos como el cuadrado de los puntos que indica el dado. Se pagan 14 pesos para participar. ¿Cuál es la ganancia esperada en este juego?

## Resolución.

Pizarrón □

# Outline

# Definición de esperanza de una variable continua

## Definición 9

Dada una variable aleatoria continua $X$ con función de densidad $f_X(x)$, definimos la esperanza de $X$ mediante la fórmula,

$$E[X] = \int_{-\infty}^{\infty} x \, f_X(x) \, dx \, , \tag{2}$$

siempre que $\int_{-\infty}^{\infty} |x| \, f(x) \, dx < \infty$.

# Esperanza infinita

Denotemos con Si $\int_{-\infty}^{+\infty} |x| f_X(x) dx < \infty$,

$$\int_{-\infty}^{+\infty} |x| f_X(x) dx = \int_{-\infty}^{0} x f_X(x) dx - \int_{0}^{+\infty} x f_X(x) dx$$

Si $\int_{-\infty}^{+\infty} |x| f_X(x) dx = \infty$. Entonces pueden ocurrir tres casos distintos.

1. $\int_0^{\infty} x f_X(x) dx = +\infty$ y $\int_{-\infty}^{0} x f_X(x) dx = -\infty$.
2. $\int_0^{\infty} x f_X(x) dx = +\infty$ y $\int_{-\infty}^{0} x f_X(x) dx > -\infty$.
3. $\int_0^{\infty} x f_X(x) dx < +\infty$ y $\int_{-\infty}^{0} x f_X(x) dx = -\infty$.

En el caso 1 no se puede definir la esperanza de $X$.

En el caso 2 se define $E(X) = +\infty$.

En el 3 $E(X) = -\infty$.

# Esperanza infinita

Denotemos con Si $\int_{-\infty}^{+\infty} |x| f_X(x)dx < \infty$,

$$\int_{-\infty}^{+\infty} |x| f_X(x)dx = \int_{-\infty}^{0} x f_X(x)dx - \int_{0}^{+\infty} x f_X(x)dx$$

Si $\int_{-\infty}^{+\infty} |x| f_X(x)dx = \infty$. Entonces pueden ocurrir tres casos distintos.

1. $\int_0^{\infty} x f_X(x)dx = +\infty$ y $\int_{-\infty}^{0} x f_X(x)dx = -\infty$.
2. $\int_0^{\infty} x f_X(x)dx = +\infty$ y $\int_{-\infty}^{0} x f_X(x)dx > -\infty$.
3. $\int_0^{\infty} x f_X(x)dx < +\infty$ y $\int_{-\infty}^{0} x f_X(x)dx = -\infty$.

En el caso 1 no se puede definir la esperanza de $X$.

En el caso 2 se define $E(X) = +\infty$.

En el 3 $E(X) = -\infty$.

# Esperanza infinita

Denotemos con Si $\int_{-\infty}^{+\infty} |x| f_X(x)dx < \infty$,

$$\int_{-\infty}^{+\infty} |x| f_X(x)dx = \int_{-\infty}^{0} x f_X(x)dx - \int_{0}^{+\infty} x f_X(x)dx$$

Si $\int_{-\infty}^{+\infty} |x| f_X(x)dx = \infty$. Entonces pueden ocurrir tres casos distintos.

1. $\int_0^\infty x f_X(x)dx = +\infty$ y $\int_{-\infty}^0 x f_X(x)dx = -\infty$.
2. $\int_0^\infty x f_X(x)dx = +\infty$ y $\int_{-\infty}^0 x f_X(x)dx > -\infty$.
3. $\int_0^\infty x f_X(x)dx < +\infty$ y $\int_{-\infty}^0 x f_X(x)dx = -\infty$.

En el caso 1 no se puede definir la esperanza de $X$.
En el caso 2 se define $E(X) = +\infty$.
En el 3 $E(X) = -\infty$.

# Esperanza infinita

Denotemos con Si $\int_{-\infty}^{+\infty} |x| f_X(x) dx < \infty$,

$$\int_{-\infty}^{+\infty} |x| f_X(x) dx = \int_{-\infty}^{0} x f_X(x) dx - \int_{0}^{+\infty} x f_X(x) dx$$

Si $\int_{-\infty}^{+\infty} |x| f_X(x) dx = \infty$. Entonces pueden ocurrir tres casos distintos.

1. $\int_0^\infty x f_X(x) dx = +\infty$ y $\int_{-\infty}^0 x f_X(x) dx = -\infty$.
2. $\int_0^\infty x f_X(x) dx = +\infty$ y $\int_{-\infty}^0 x f_X(x) dx > -\infty$.
3. $\int_0^\infty x f_X(x) dx < +\infty$ y $\int_{-\infty}^0 x f_X(x) dx = -\infty$.

En el caso 1 no se puede definir la esperanza de $X$.

En el caso 2 se define $E(X) = +\infty$.

En el 3 $E(X) = -\infty$.

# Esperanza infinita

Denotemos con Si $\int_{-\infty}^{+\infty} |x| f_X(x)dx < \infty$,

$$\int_{-\infty}^{+\infty} |x| f_X(x)dx = \int_{-\infty}^{0} x f_X(x)dx - \int_{0}^{+\infty} x f_X(x)dx$$

Si $\int_{-\infty}^{+\infty} |x| f_X(x)dx = \infty$. Entonces pueden ocurrir tres casos distintos.

1. $\int_0^\infty x f_X(x)dx = +\infty$ y $\int_{-\infty}^0 x f_X(x)dx = -\infty$.
2. $\int_0^\infty x f_X(x)dx = +\infty$ y $\int_{-\infty}^0 x f_X(x)dx > -\infty$.
3. $\int_0^\infty x f_X(x)dx < +\infty$ y $\int_{-\infty}^0 x f_X(x)dx = -\infty$.

En el caso 1 no se puede definir la esperanza de $X$.
En el caso 2 se define $E(X) = +\infty$.
En el 3 $E(X) = -\infty$.

# Esperanzas de variables continuas famosas

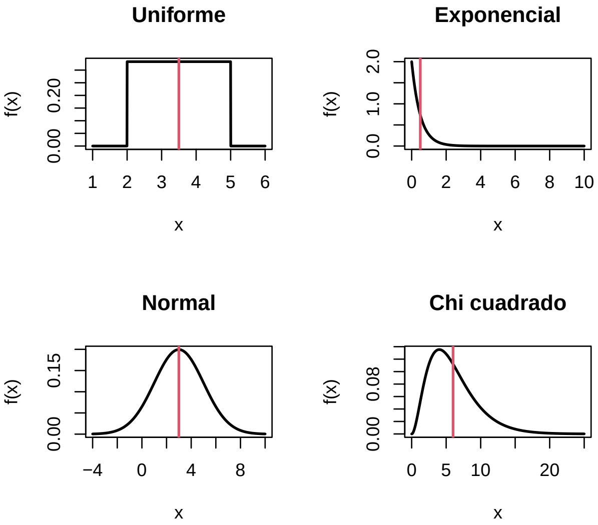
**Proposición 1.2**

1. *Si* $X \sim \mathcal{U}[a,b]$ *entonces* $E(X) = (a+b)/2$.
2. *Si* $X \sim \mathcal{E}(\lambda)$ *entonces* $E(X) = 1/\lambda$.
3. *Si* $X \sim \mathcal{N}(0,1)$ *entonces* $E(X) = 0$. *Más aún, si* $X \sim \mathcal{N}(\mu, \sigma^2)$ *entonces* $E(X) = \mu$
4. *Si* $X \sim \Gamma(\alpha, \lambda)$ *entonces* $E(X) = \alpha/\lambda$.
5. *Si* $X \sim \chi^2_k$ *entonces* $E(X) = n$.

**Demostración.**

Vemos algunas en el pizarrón, las demás quedan como ejercicio $\square$

# Densidades y esperanzas



# Ley del estadístico inconciente

## Teorema 10

*Sea* $(X_1, X_2, \ldots, X_k)$ *un vector aleatorio n-dimensional. Sea* $g : \mathbb{R}^n \to \mathbb{R}$ *una función boreliana. Consideremos la variable aleatoria* $Y = g(X_1, X_2, \ldots, X_k)$*. Entonces, si* $(X_1, X_2, \ldots, X_k)$ *es continuo y*

$$\int \ldots \int |g(x_1, \ldots, x_k)| \, f_{X_1 x_2 \ldots x_k}(x_1, \ldots, x_k) \, dx_1 dx_2 \ldots dx_k < +\infty$$

*entonces*

$$E[g(X_1, \ldots, X_k)] = \int \ldots \int g(x_1, \ldots, x_k) \, f_{X_1 \ldots x_k}(x_1, \ldots, x_k) \, dx_1 \cdots dx_k$$

## Demostración.

Ejercicio (similar al caso discreto) $\square$

## Ejemplo 11 (Distancia de frenado)

La distancia de frenado (la distancia recorrida por un auto desde el momento en que se aplica el freno hasta el momento en que se detiene) en pies es una función cuadrática de la velocidad del coche (en mph), es decir, $g(x) = cx^2$ donde $c$ es una constante que depende de las características físicas del coche, sus frenos, la superficie de la carretera, etc. Para los propósitos de este ejemplo, tomemos un valor realista $c = 9 \cdot 10^{-6}$. En un estudio de tráfico, la distribución de las velocidades de los coches al inicio del frenado se determina empíricamente como una v.a. $X$ uniformemente distribuida en el intervalo [60, 90], medido en millas por hora. Hallar $E(Y)$.

## Resolución.

Pizarrón.

Resultado: $E(X) = 0{,}05159$, es decir que, en promedio, los autos de esta autopista tardarían 0,05159 millas en detenerse completamente si tuvieran que frenar. Es decir, aproximadamente 83 metros. Tengamos en cuenta además que el intervalo $[60, 90]$ en millas es aprox $[97, 145]$ $\square$

# Outline

## Proposición 1.3

*Sean $X$ e $Y$ v.a. definidas en $(\Omega, \Sigma, P)$ con esperanza finita.*

1. $E(aX) = aE(X)$ *para todo* $a \in \mathbb{R}$
2. $E(X + Y) = E(X) + E(Y)$

## Demostración.

Pizarrón $\square$

## Corolario 1.1

*Si* $X \sim N(\mu, \sigma^2)$ *entonces* $E(X) = \mu$.

## Proposición 1.4

*Sean* $X_1 \ldots, X_k$ *v.a. definidas en* $(\Omega, \Sigma, P)$ *con esperanza finita y* $a_1, \ldots, a_k \in \mathbb{R}$, *entonces* $E(\sum_{i=1}^n a_i X_i) = \sum_{i=1}^n a_i E(X_i)$

## Proposición 1.3

*Sean $X$ e $Y$ v.a. definidas en $(\Omega, \Sigma, P)$ con esperanza finita.*

1. $E(aX) = aE(X)$ *para todo* $a \in \mathbb{R}$
2. $E(X + Y) = E(X) + E(Y)$

## Demostración.

Pizarrón $\square$

## Corolario 1.1

*Si* $X \sim N(\mu, \sigma^2)$ *entonces* $E(X) = \mu$.

## Proposición 1.4

*Sean* $X_1 \ldots, X_k$ *v.a. definidas en* $(\Omega, \Sigma, P)$ *con esperanza finita y* $a_1, \ldots, a_k \in \mathbb{R}$, *entonces* $E(\sum_{i=1}^{n} a_i X_i) = \sum_{i=1}^{n} a_i E(X_i)$

## Proposición 1.3

*Sean $X$ e $Y$ v.a. definidas en $(\Omega, \Sigma, P)$ con esperanza finita.*

1. $E(aX) = aE(X)$ *para todo* $a \in \mathbb{R}$
2. $E(X + Y) = E(X) + E(Y)$

## Demostración.

Pizarrón □

## Corolario 1.1

*Si* $X \sim N(\mu, \sigma^2)$ *entonces* $E(X) = \mu$.

## Proposición 1.4

*Sean* $X_1 \ldots, X_k$ *v.a. definidas en* $(\Omega, \Sigma, P)$ *con esperanza finita y* $a_1, \ldots, a_k \in \mathbb{R}$, *entonces* $E(\sum_{i=1}^n a_i X_i) = \sum_{i=1}^n a_i E(X_i)$

### Definición 1.1

*Sean $X$ e $Y$ v.a. definidas en $(\Omega, \Sigma, P)$, decimos que $X \leq Y$ si $X(\omega) \leq Y(\omega)\ \forall \omega \in \Omega$*

### Proposición 1.5

*Sean $X$ e $Y$ v.a. definidas en $(\Omega, \Sigma, P)$ con esperanza finita tales que $X \leq Y$. Entonces $E(X) \leq E(Y)$*

### Demostración.

Pizarrón $\square$

# Ejemplo:distancia de frenado

**Ejemplo 12 (Distancia de frenado)**

Se sabe que la distancia de frenado es es una función cuadrática de la velocidad del coche $g(x) = cx^2$. Supongamos que por mejoras en el asfalto, la constante $c$ baja de $9 \cdot 10^{-6}$ a $8 \cdot 10^{-6}$. Sea $Y_1 = 9 \cdot 10^{-6} X^2$, $Y_2 = 8 \cdot 10^{-6} X^2$, con $X \sim \mathcal{U}[60, 90]$. Entonces $E(Y_2) < E(Y_1)$.

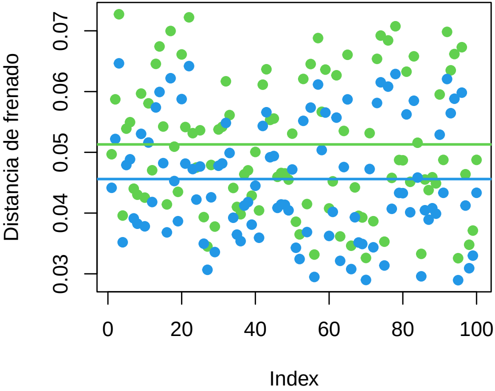
# Ejemplo:distancia de frenado

```
x <- runif(100, 60,90)
c1 <- 9*10^(-6)
c2 <- 8*10^(-6)
y1 <- c1*x^2
y2 <- c2*x^2
plot(y1, pch=19, col=3, ylim=c(c2*60^2,c1*90^2),
     ylab="Distancia de frenado")
points(y2, col =4, pch=19)

mu1 <- (90^3-60^3)/90 * c1
mu2 <- (90^3-60^3)/90 * c2

abline(h=mu1, col =3, lwd = 2)
abline(h=mu2, col =4, lwd = 2)
```

# Ejemplo:distancia de frenado




Distancias de frenado de 100 autos antes de las mejoras en el asfalto (verde) y después (celeste). Las líneas indican la esperanza en cada caso.

## Proposición 1.6

*Sean* $X \geq 0$ *y* $E(X) = 0$, *entonces* $P(X = 0) = 1$

## Demostración.

Pizarrón $\square$

**Proposición 1.7**

*Sean $X$ e $Y$ v.a. independientes definidas en $(\Omega, \Sigma, P)$ con esperanza finita. Entonces $E(XY)$ también es finita y $E(XY) = E(X)E(Y)$*

**Demostración.**

Pizarrón $\square$

# Outline

**Ejemplo 13**

El tiempo, en horas por día, que los estudiantes de una sede universitaria destinan a viajar desde sus hogares hasta la sede es una v.a. $Y$ con distribución exponencial de parámetro $\lambda = 1$. Hallar la constante que mejor aproxima a $Y$.

# La esperanza como mejor predictor

Qué significa “que mejor aproxima”?

**Definición 1.2 (Error cuadrático medio)**

*Sea $Y$ una variable aleatoria. El error cuadrático medio de la constante $c$ como predictor de $Y$ se define como*
$\mathrm{ECM}(c, Y) = E\left((Y - c)^2\right)$

**Definición 1.3 (Error absoluto medio)**

*Sea $Y$ una variable aleatoria. El error absoluto medio de la constante $c$ como predictor de $Y$ se define como*
$\mathrm{EAM}(c, Y) = E\left(|Y - c|\right)$

La constante que mejor aproxima a $Y$ puede definirse como la que minimiza el $\mathrm{ECM}$ o como la que minimiza el $\mathrm{EAM}$. ¿Serán iguales?

# La esperanza como mejor predictor

Qué significa "que mejor aproxima"?

**Definición 1.2 (Error cuadrático medio)**

*Sea $Y$ una variable aleatoria. El error cuadrático medio de la constante $c$ como predictor de $Y$ se define como*
$\mathrm{ECM}(c, Y) = E\left((Y - c)^2\right)$

**Definición 1.3 (Error absoluto medio)**

*Sea $Y$ una variable aleatoria. El error absoluto medio de la constante $c$ como predictor de $Y$ se define como*
$\mathrm{EAM}(c, Y) = E\left(|Y - c|\right)$

La constante que mejor aproxima a $Y$ puede definirse como la que minimiza el $\mathrm{ECM}$ o como la que minimiza el $\mathrm{EAM}$. ¿Serán iguales?

# La esperanza como mejor predictor

Qué significa "que mejor aproxima"?

## Definición 1.2 (Error cuadrático medio)

*Sea $Y$ una variable aleatoria. El error cuadrático medio de la constante $c$ como predictor de $Y$ se define como*
$\mathrm{ECM}(c, Y) = E\left((Y - c)^2\right)$

## Definición 1.3 (Error absoluto medio)

*Sea $Y$ una variable aleatoria. El error absoluto medio de la constante $c$ como predictor de $Y$ se define como*
$\mathrm{EAM}(c, Y) = E\left(|Y - c|\right)$

La constante que mejor aproxima a $Y$ puede definirse como la que minimiza el $\mathrm{ECM}$ o como la que minimiza el $\mathrm{EAM}$. ¿Serán iguales?

# La esperanza como mejor predictor

## Ejemplo 14

El tiempo, en horas por día, que los estudiantes de una sede universitaria destinan a viajar desde sus hogares hasta la sede es una v.a. $Y$ con distribución exponencial de parámetro $\lambda = 2$. Hallar la constante que mejor aproxima a $Y$.

## Resolución.

Calculemos la constante que minimiza el $\mathrm{ECM}$

$$\mathrm{ECM}(c, Y) = E\left((Y - c)^2\right) = E(Y^2) - 2cE(Y) + c^2$$

Se minimiza en $c = E(Y)$, que en este caso es $1/2$. $\square$

Esto se puede hacer para cualquier distribución, siempre que $E(Y^2) < \infty$.

## Resolución (continuación).

Calculemos la constante que minimiza el $\mathrm{EAM}$

$$H(c) = E(|Y-c|) = \int_{-\infty}^{c} (c-y) f_Y(y)dy + \int_{c}^{+\infty} (y-c) f_Y(y)dy =$$

$$c\int_{-\infty}^{c} f_Y(y)dy - \int_{-\infty}^{c} y f_Y(y)dy + \int_{c}^{+\infty} y f_Y(y)dy - c\int_{c}^{+\infty} f_Y(y)dy$$

$$H'(c) = \int_{-\infty}^{c} f_Y(y)dy + cf_Y(c) - cf_Y(c) - cf_Y(c) - \int_{c}^{+\infty} f_Y(y)dy + cf_Y(c)$$

$$= \int_{-\infty}^{c} f_Y(y)dy - \int_{c}^{+\infty} f_Y(y)dy$$

$$H'(c) = 0 \Leftrightarrow \int_{-\infty}^{c} f_Y(y)dy = \int_{c}^{+\infty} f_Y(y)dy$$

Como la suma es 1, debe ser

$$\int_{-\infty}^{c} f_Y(y)dy = 0{,}5,$$

o sea que $c = \mathrm{med}(Y)$ $\square$

## Resolución (continuación).

Calculemos la mediana para nuestra $\mathcal{E}(2)$.
Buscamos $m$ tal que

$$\int_0^m \lambda e^{-\lambda x} dx = 0{,}5$$

$$-e^{-\lambda x}\Big|_0^m = 1 - e^{-\lambda m} = 0{,}5$$

$$e^{-\lambda m} = 0{,}5$$

$$-\lambda m = \ln(0{,}5)$$

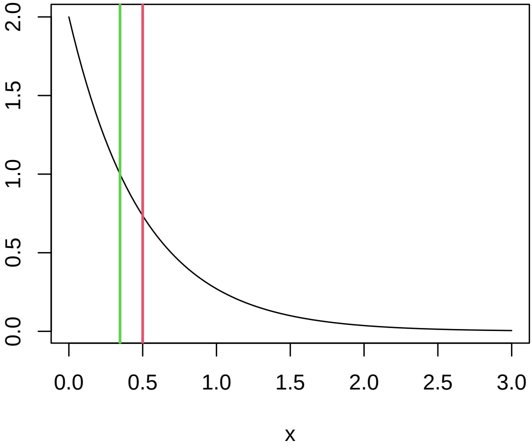
$$\lambda m = -\ln(0{,}5) = \ln(2) \Rightarrow m = \ln(2)/\lambda$$

Por lo tanto,

$$Y \sim \mathcal{E}(2) \Rightarrow \mathrm{Med}(Y) \approx 0{,}346$$

El mejor predictor del tiempo de viaje de los estudiantes con el criterio del EAM es de aprox 21 minutos. $\square$

# Densidad exponencial, media y mediana



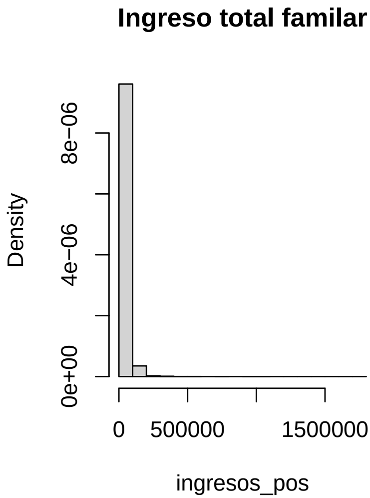

Densidad exponencial, media (en rojo) y mediana (en verde).

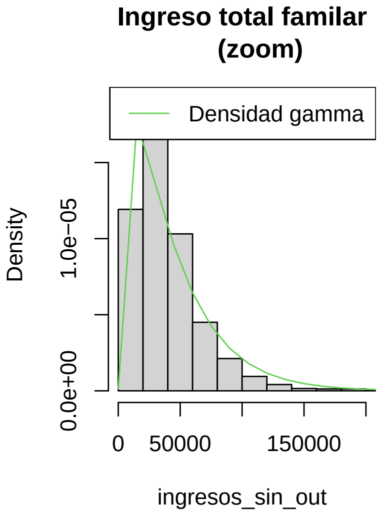
## Observación 1.1

1. *La constante que mejor aproxima a una variable aleatoria $Y$, en el sentido de que minimiza el $\mathrm{ECM}$ es $E(X)$.*
2. *La constante que mejor aproxima a una variable aleatoria $Y$, en el sentido de que minimiza el $\mathrm{EAM}$ es la mediana de $X$ (esto lo probamos sólo para continuas, pero también se puede hacer fácilmente para discretas).*

# Ejemplo: Ingreso total familiar en Argentina (datos de 2016)

(Fuente: Encuesta permanente de hogares, INDEC, [www.indec.gob.ar])



Supongamos que el ingreso total familar en la Argentina sigue una distribución

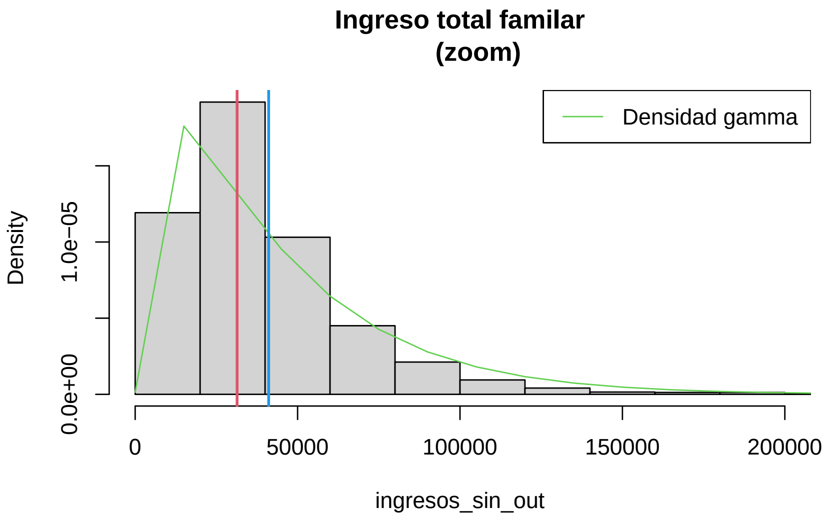
$$\Gamma \left( 1{,}33, \frac{1}{30901} \right)$$

Calcular el mejor predictor del ITF, según el criterio del ECM y según el criterio del EAM.

```
> mediana <- qgamma(0.5, shape=1.33, rate = 1/30901)
> mediana
[1] 31379.4
> media <- 1.33*30901
> media
[1] 41098.33
```

**Interpretación:**

El ingreso promedio de los argentinos en 2016 era de 41098.33 pesos.
El ingreso del 50 % de los argentinos era de 31379.4 pesos en 2016
La diferencia entre media y mediana indica la asimetría en la distribución de los ingresos.




Ingreso total familiar en Argentina en 2016. Las rectas verticales indican la media (en azul) y la mediana (en rojo)

# Outline

# Varianza

Acabamos de mostrar que la esperanza es la mejor constante para aproximar a nuestra variable aleatoria. Vamos a querer ahora medir cual es el precio que cometemos al hacer esta aproximacíón. Es decir, queremos calcular cuanto vale $\mathrm{ECM}(x, Y)$ cuando $c = E[X]$.

**Definición 15**

Dada una variable aleatoria $X$ con esperanza $\mu$, definimos su varianza mediante la formula

$$V(X) = E\left[\left(X - \mu\right)^2\right]. \tag{3}$$

La varianza de la variable aleatoria esta midiendo cuan dispersa está esta alrededor de su esperanza. Notemos que $V(X) \geq 0$ y además $V(X) = 0$ si y solo si la variable aleatoria es constate $X = \mu$. El siguiente lemma, da una fórmula alternativa para computar la varianza de una variable aleatoria.

# Varianza

Acabamos de mostrar que la esperanza es la mejor constante para aproximar a nuestra variable aleatoria. Vamos a querer ahora medir cual es el precio que cometemos al hacer esta aproximacíón. Es decir, queremos calcular cuanto vale $\mathrm{ECM}(x, Y)$ cuando $c = E[X]$.

## Definición 15

Dada una variable aleatoria $X$ con esperanza $\mu$, definimos su varianza mediante la formula

$$V(X) \;=\; E\left[\left(X-\mu\right)^2\right]. \tag{3}$$

La varianza de la variable aleatoria esta midiendo cuan dispersa está esta alrededor de su esperanza. Notemos que $V(X) \geq 0$ y además $V(X) = 0$ si y solo si la variable aleatoria es constate $X = \mu$. El siguiente lemma, da una fórmula alternativa para computar la varianza de una variable aleatoria.

# Varianza

Acabamos de mostrar que la esperanza es la mejor constante para aproximar a nuestra variable aleatoria. Vamos a querer ahora medir cual es el precio que cometemos al hacer esta aproximacíón. Es decir, queremos calcular cuanto vale $\mathrm{ECM}(x, Y)$ cuando $c = E[X]$.

**Definición 15**

Dada una variable aleatoria $X$ con esperanza $\mu$, definimos su varianza mediante la formula

$$V(X) \; = \; E\left[\left(X - \mu\right)^2\right]. \tag{3}$$

La varianza de la variable aleatoria esta midiendo cuan dispersa está esta alrededor de su esperanza. Notemos que $V(X) \geq 0$ y además $V(X) = 0$ si y solo si la variable aleatoria es constate $X = \mu$. El siguiente lemma, da una fórmula alternativa para computar la varianza de una variable aleatoria.

## Lema 16

*Sea $X$ una variable aleatoria con esperanza $\mu$, entonces vale que*

$$V(X) \; = \; E\left[X^2\right] - \mu^2 \; .$$

## Demostración.

Pizarrón $\square$

## Ejemplo 17

Se tira un dado. Si sale par se ganan 2 pesos. Si sale 1,3 o 5 se pierden 1,3 o 5 pesos respectivamente. Hallar el desvío estandar de la ganancia.

## Resolución.

Pizarrón □

### Lema 18

*Dados $\alpha$ y $\beta$ números reales, tenemos que*

$$V(\alpha X + \beta) = \alpha^2 V(X) .$$

### Demostración.

Ejercicio $\square$

# Ejemplo

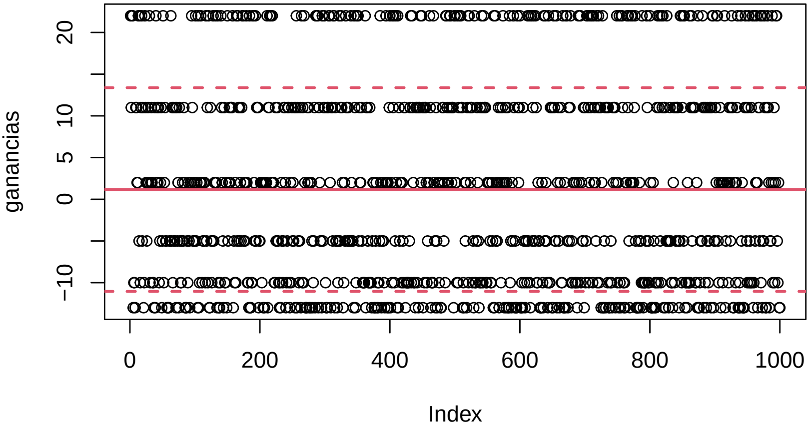
## Ejemplo 19

Me invitan a jugar el siguiente juego de azar. Se tira un dado. Se gana tantos pesos como el cuadrado de los puntos que indica el dado. Se pagan 14 pesos para participar. ¿Cuál es la varianza de la ganancia?

## Resolución.

Pizarrón □

# Ganancias simuladas, media y varianza




1000 ganancias simuladas, esperanza (linea llena) y esperanza $\pm$ desvioestandar (líneas punteadas)

# Varianza de variables famosas

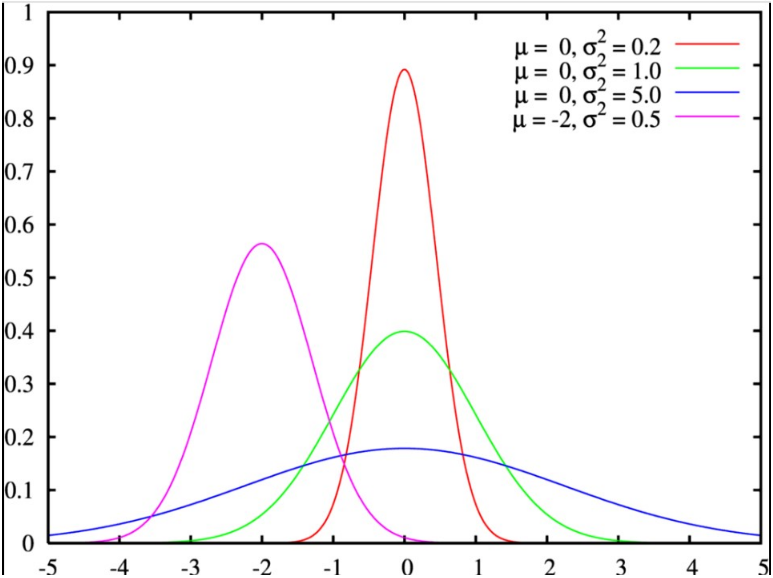
## Proposición 2.1

1. *Si* $X \sim B(1,p)$ *entonces* $V(X) = p(1-p)$
2. *Si* $X \sim B(n,p)$ *entonces* $E(X) = np(1-p)$
3. *Si* $X \sim \mathcal{P}(\lambda)$ *entonces* $V(X) = \lambda$
4. *Si* $X \sim \mathcal{U}[a,b]$ *entonces* $V(X) = (b-a)^2/12$.
5. *Si* $X \sim \mathcal{E}(\lambda)$ *entonces* $E(X) = 1/\lambda^2$.
6. *Si* $X \sim \mathcal{N}(0,1)$ *entonces* $E(X) = 1$. *Más aún, si* $X \sim \mathcal{N}(\mu, \sigma^2)$ *entonces* $E(X) = \sigma^2$

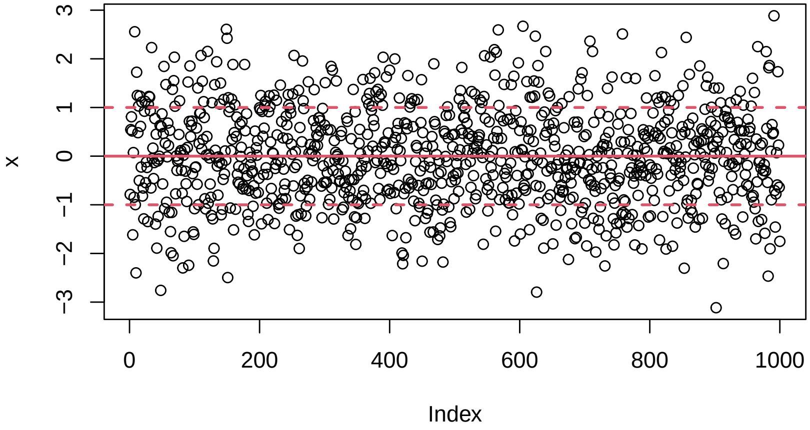
## Demostración.

Pizarrón □

# Densidades normales



# Observaciones $N(0,1)$ simuladas, media y varianza




1000 observaciones $N(0,1)$ simuladas, esperanza (linea llena) y esperanza $\pm$ desvioestandar (líneas punteadas)

# Outline

# Outline

# El espacio $\mathcal{L}^2$

## Definición 20

Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad. Definimos el espacio $\mathcal{L}^2(\Omega, \mathcal{F}, P)$ por

$$\mathcal{L}^2(\Omega, \mathcal{F}, P) = \{X : X \text{ es una variable aleatoria y } E\left[X^2\right] < +\infty\}.$$

## Proposición 3.1

$\mathcal{L}^2(\Omega, \mathcal{F}, P)$ *es un espacio vectorial con las operaciones*
- $a \in \mathbb{R}, X \in \mathcal{L}^2(\Omega, \mathcal{F}, P)$, *entonces* $aX$ *también está en* $\mathcal{L}^2(\Omega, \mathcal{F}, P)$
- $X, Y \in \mathcal{L}^2(\Omega, \mathcal{F}, P)$, *entonces* $X + Y \in \mathcal{L}^2(\Omega, \mathcal{F}, P)$.

## Demostración.

Para probarlo necesitamos probar que se verifican todas las propiedades de espacios vectoriales, por ejemplo:

a) $X = 0 \in \mathcal{L}^2(\Omega, \mathcal{F}, P)$

b) $X, Y \in \mathcal{L}^2(\Omega, \mathcal{F}, P), a, b \in \mathbb{R}$, entonces $aX + bY \in \mathcal{L}^2(\Omega, \mathcal{F}, P)$. O sea que sabiendo que $E\left[X^2\right] < +\infty, E\left[Y^2\right] < +\infty$ queremos probar que $E\left[(aX + bY)^2\right] < +\infty$.

Sabiendo que $E\left[X^2\right] < +\infty, E\left[Y^2\right] < +\infty$ qvq $E\left[(aX + bY)^2\right] < +\infty$.

Sabemos que para todo $\alpha, \beta \in \mathbb{R}$, vale

$$(\alpha + \beta)^2 \leq 4\alpha^2 + 4\beta^2 \quad \text{(pizarrón)}$$

Luego $(aX + bY)^2 \leq 4a^2X^2 + 4b^2Y^2$. Por la monotonía de la esperanza, y la linealidad, tenemos

$$E\left[(aX + bY)^2\right] \leq E\left[4a^2X^2 + 4b^2Y^2\right] = 4a^2E\left[X^2\right] + 4b^2E\left[Y^2\right] < +\infty$$

$\square$

# Producto interno en $\mathcal{L}^2(\Omega, \mathcal{F}, P)$

## Proposición 3.2

$\langle X, Y \rangle = E[XY]$ *define un producto interno en* $\mathcal{L}^2(\Omega, \mathcal{F}, P)$.

## Demostración.

Veamos que
- Si $X, Y \in \mathcal{L}^2(\Omega, \mathcal{F}, P)$, $E(XY) < \infty$.
- $\langle X, Y \rangle$ es bilineal, es decir, $E[(aX + bY)Z] = aE[XZ] + bE[YZ]$
- $\langle X, Y \rangle$ es simétrico: trivial
- $\langle X, Y \rangle$ es definido positivo: sea $X$ no nula (o sea, $P(X = 0) \neq 1$ ), entonces $\langle X, X \rangle > 0$. $\square$

## Observación 3.1

*En el espacio* $\mathcal{L}^2(\Omega, \mathcal{F}, P)$, $X = 0$ *quiere decir* $P(X = 0) = 1$ *y* $X = Y$ *quiere decir* $P(X = Y) = 1$.

# Producto interno en $\mathcal{L}^2(\Omega, \mathcal{F}, P)$

## Proposición 3.2

*$\langle X, Y \rangle = E[XY]$ define un producto interno en $\mathcal{L}^2(\Omega, \mathcal{F}, P)$.*

## Demostración.

Veamos que
- Si $X, Y \in \mathcal{L}^2(\Omega, \mathcal{F}, P)$, $E(XY) < \infty$.
- $\langle X, Y \rangle$ es bilineal, es decir, $E[(aX + bY)Z] = aE[XZ] + bE[YZ]$
- $\langle X, Y \rangle$ es simétrico: trivial
- $\langle X, Y \rangle$ es definido positivo: sea $X$ no nula (o sea, $P(X = 0) \neq 1$ ), entonces $\langle X, X \rangle > 0$. $\square$

## Observación 3.1

*En el espacio $\mathcal{L}^2(\Omega, \mathcal{F}, P)$, $X = 0$ quiere decir $P(X = 0) = 1$ y $X = Y$ quiere decir $P(X = Y) = 1$.*

# Desigualdad de Cauchy - Schwarz

## Lema 21 (Desigualdad de Cauchy - Schwarz)

*En todo $\mathbb{R}-$espacio vectorial con producto interno $\mathbb{V}$ vale que, para todo $x, y \in \mathbb{V}$,*

$$\langle x, y\rangle^2 \leq \langle x, x\rangle \langle y, y\rangle$$

*La igualdad vale si y sólo si $x$ e $y$ son linealmente dependientes.*

## Lema 22 (desigualdad de Cauchy - Schwarz)

*$X, Y \in \mathcal{L}^2(\Omega, \mathcal{F}, P)$, entonces $[E(XY)]^2 \leq E\left(X^2\right) E\left(Y^2\right)$. La igualdad vale si y sólo si existe $a \in \mathbb{R}$ tal que $P(Y = aX) = 1$.*

## Demostración.

Pizarrón □

## Corolario 23

*Si $X \in \mathcal{L}^2(\Omega, \mathcal{F}, P)$ entonces $E(|X|) < \infty$*

# Desigualdad de Cauchy - Schwarz

## Lema 21 (Desigualdad de Cauchy - Schwarz)

*En todo* $\mathbb{R}-$*espacio vectorial con producto interno* $\mathbb{V}$ *vale que, para todo* $x, y \in \mathbb{V}$,

$$\langle x, y \rangle^2 \leq \langle x, x \rangle \langle y, y \rangle$$

*La igualdad vale si y sólo si* $x$ *e* $y$ *son linealmente dependientes.*

## Lema 22 (desigualdad de Cauchy - Schwarz)

$X, Y \in \mathcal{L}^2(\Omega, \mathcal{F}, P)$, *entonces* $[E(XY)]^2 \leq E\left(X^2\right) E\left(Y^2\right)$. *La igualdad vale si y sólo si existe* $a \in \mathbb{R}$ *tal que* $P(Y = aX) = 1$.

## Demostración.

Pizarrón $\square$

## Corolario 23

*Si* $X \in \mathcal{L}^2(\Omega, \mathcal{F}, P)$ *entonces* $E(|X|) < \infty$

## Definición 24

Dadas dos variables aleatorias $X$ e $Y$ definimos la covarianza entre ellas mediante la fórmula

$$Cov(X, Y) = E\Big[(X - \mu_X)(Y - \mu_Y)\Big] \,,$$

siendo $\mu_X = E[X]$ y $\mu_Y = E[Y]$.

## Lema 25

*Fórmula reducida para la covarianza:*

$$Cov(X, Y) = E[XY] - E[X]E[Y] \,.$$

**Teorema 26**

*Si $X$ e $Y$ son independientes, entonces*

$$Cov(X, Y) = 0$$

Cuidado! covarianza cero NO garantiza independencia!

**Ejemplo 3.1**

*Sea $X \sim \mathcal{U}(-1, 1)$. Sea $Y = X^2$. Demuestre que $Cov(X, Y) = 0$. son las variables independientes?*

## Lema 27

*La covarianza verifica las siguientes propiedades:*

1. $Cov(X, Y) = Cov(Y, X)$
2. $Cov(X, X) = V(X)$
3. $Cov(X + Y, Z) = Cov(X, Z) + Cov(Y, Z)$
4. $Cov(aX, Y) = a\, Cov(X, Y)$
5. $Cov(\sum_{i=1}^{n} a_i\, X_i, \sum_{j=1}^{m} b_j\, Y_j) = \sum_{i=1}^{n} \sum_{j=1}^{m} a_i\, b_j\, Cov(X_i, Y_j)$
6. $V(X \pm Y) = V(X) + V(Y) \pm 2\,\mathrm{Cov}(X, Y)$

# Desigualdad de Cauchy-Schwartz para la covarianza

**Teorema 28 (Cauchy-Schwarz para covarianza)**

$$|\operatorname{Cov}(X, Y)|^2 \leq V(X)V(Y)$$

*La igualdad vale si y sólo si existen* $a, b \in \mathbb{R}$ *con* $a \neq 0$ *tales que* $P(Y = aX + b) = 1$.

## Corolario 3.1

*Si* $X_i, \cdots, X_n$ *son independientes, entonces*

$$V(\sum_{i=1}^{n} X_i) \; = \; \sum_{i=1}^{n} V(X_i) \quad .$$

*Además, si todas tienen la misma varianza* $V(X_i) = \sigma^2$ *para todo* $i$

$$V(\sum_{i=1}^{n} X_i) \; = \; n\sigma^2 \quad .$$

# Interpretación del signo de la covarianza

Si las variables aleatorias están positivamente asociadas, es decir, si cuando $X$ es mayor que su media, $Y$ tiende a ser mayor que su media, entonces, la covarianza será positiva. Si la asociación es negativa, es decir, cuando $X$ es mayor que su media, $Y$ tiende a ser menor que su media, la covarianza es negativa.

## Definición 29

Dadas dos variables aleatorias $X$ e $Y$ definimos el coeficiente de correlacion entre ellas mediante la fórmula

$$\rho(X, Y) = \frac{Cov(X, Y)}{\sqrt{V(X)V(Y)}}.$$

# Propiedades de la correlación

**Lema 30**

$$-1 \leq \rho(X, Y) \leq 1$$

**Demostración.**

Es consecuencia de la desigualdad de Cauchy-Schwartz □

**Lema 31**

$$\rho(aX, bY) = \rho(X, Y)$$

# Covarianza y correlación de una normal bivariada

## Proposición 3.3

*Si $(X, Y)$ es un vector aleatorio normal bivariado con parámetros $\mu_1$, $\mu_2$, $\sigma_1$, $\sigma_2$, $\rho$ entonces $E(X) = \mu_1$, $E(Y) = \mu_2$, $V(X) = \sigma_1^2$, $V(Y) = \sigma_2^2$, $\text{Cov}(X, Y) = \rho\sigma_1\sigma_2$ y $\rho(X, Y) = \rho$.*

## Demostración.

$$f_{XY}(x, y) = \frac{1}{C} e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right)\right]}.$$

con $C = 2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}$.

$$\text{Cov}(X, Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\mu_X)(y-\mu_Y) f_{XY}(x, y)dxdy$$

Cambio de variables $u = (x-\mu_X)/\sigma_X$, $v = (y-\mu_Y)/\sigma_Y$ con $J = \sigma_X\sigma_Y$

$$\frac{\sigma_X\sigma_Y}{2\pi\sqrt{1-\rho^2}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} uv \exp\left[-\frac{1}{2(1-\rho^2)}\left(u^2+v^2-2\rho uv\right)\right] dudv$$

# Covarianza y correlación de una normal bivariada

## Proposición 3.3

*Si* $(X, Y)$ *es un vector aleatorio normal bivariado con parámetros* $\mu_1$, $\mu_2$, $\sigma_1$, $\sigma_2$, $\rho$ *entonces* $E(X) = \mu_1$, $E(Y) = \mu_2$, $V(X) = \sigma_1^2$, $V(Y) = \sigma_2^2$, $\mathrm{Cov}(X, Y) = \rho\sigma_1\sigma_2$ *y* $\rho(X, Y) = \rho$.

## Demostración.

$$f_{XY}(x, y) = \frac{1}{C} e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right)\right]}.$$

con $C = 2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}$.

$$\mathrm{Cov}(X, Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x - \mu_X)(y - \mu_Y) f_{XY}(x, y) dxdy$$

Cambio de variables $u = (x - \mu_X)/\sigma_X$, $v = (y - \mu_Y)/\sigma_Y$ con $J = \sigma_X\sigma_Y$

$$\frac{\sigma_X\sigma_Y}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} uv \exp\left[-\frac{1}{2(1-\rho^2)}\left(u^2 + v^2 - 2\rho uv\right)\right] dudv$$

# Covarianza y correlación de una normal bivariada

## Proposición 3.3

*Si* $(X, Y)$ *es un vector aleatorio normal bivariado con parámetros* $\mu_1$, $\mu_2$, $\sigma_1$, $\sigma_2$, $\rho$ *entonces* $E(X) = \mu_1$, $E(Y) = \mu_2$, $V(X) = \sigma_1^2$, $V(Y) = \sigma_2^2$, $\mathrm{Cov}(X, Y) = \rho\sigma_1\sigma_2$ *y* $\rho(X, Y) = \rho$.

## Demostración.

$$f_{XY}(x, y) = \frac{1}{C} e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right)\right]}.$$

con $C = 2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}$.

$$\mathrm{Cov}(X, Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x - \mu_X)(y - \mu_Y) f_{XY}(x, y) dxdy$$

Cambio de variables $u = (x - \mu_X)/\sigma_X$, $v = (y - \mu_Y)/\sigma_Y$ con $J = \sigma_X\sigma_Y$

$$\frac{\sigma_X\sigma_Y}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} uv \exp\left[-\frac{1}{2(1-\rho^2)}\left(u^2 + v^2 - 2\rho uv\right)\right] dudv$$

# Covarianza y correlación de una normal bivariada

## Proposición 3.3

*Si* $(X, Y)$ *es un vector aleatorio normal bivariado con parámetros* $\mu_1$, $\mu_2$, $\sigma_1$, $\sigma_2$, $\rho$ *entonces* $E(X) = \mu_1$, $E(Y) = \mu_2$, $V(X) = \sigma_1^2$, $V(Y) = \sigma_2^2$, $\mathrm{Cov}(X, Y) = \rho\sigma_1\sigma_2$ *y* $\rho(X, Y) = \rho$.

## Demostración.

$$f_{XY}(x, y) = \frac{1}{C} e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right)\right]}.$$

con $C = 2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}$.

$$\mathrm{Cov}(X, Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x - \mu_X)(y - \mu_Y) f_{XY}(x, y) dxdy$$

Cambio de variables $u = (x - \mu_X)/\sigma_X$, $v = (y - \mu_Y)/\sigma_Y$ con $J = \sigma_X\sigma_Y$

$$\frac{\sigma_X\sigma_Y}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} uv \exp\left[-\frac{1}{2(1-\rho^2)}\left(u^2 + v^2 - 2\rho uv\right)\right] dudv$$

# Covarianza y correlación de una normal bivariada

## Proposición 3.3

*Si* $(X, Y)$ *es un vector aleatorio normal bivariado con parámetros* $\mu_1$, $\mu_2$, $\sigma_1$, $\sigma_2$, $\rho$ *entonces* $E(X) = \mu_1$, $E(Y) = \mu_2$, $V(X) = \sigma_1^2$, $V(Y) = \sigma_2^2$, $\mathrm{Cov}(X, Y) = \rho\sigma_1\sigma_2$ *y* $\rho(X, Y) = \rho$.

## Demostración.

$$f_{XY}(x,y) = \frac{1}{C} e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right)\right]}.$$

con $C = 2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}$.

$$\mathrm{Cov}(X,Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\mu_X)(y-\mu_Y) f_{XY}(x,y)\,dxdy$$

Cambio de variables $u = (x-\mu_X)/\sigma_X$, $v = (y-\mu_Y)/\sigma_Y$ con $J = \sigma_X\sigma_Y$

$$\frac{\sigma_X\sigma_Y}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} uv \exp\left[-\frac{1}{2(1-\rho^2)}\left(u^2+v^2-2\rho uv\right)\right] dudv$$

# Covarianza y correlación de una normal bivariada

## Proposición 3.3

*Si* $(X, Y)$ *es un vector aleatorio normal bivariado con parámetros* $\mu_1$, $\mu_2$, $\sigma_1$, $\sigma_2$, $\rho$ *entonces* $E(X) = \mu_1$, $E(Y) = \mu_2$, $V(X) = \sigma_1^2$, $V(Y) = \sigma_2^2$, $\text{Cov}(X, Y) = \rho\sigma_1\sigma_2$ *y* $\rho(X, Y) = \rho$.

## Demostración.

$$f_{XY}(x, y) = \frac{1}{C} e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right)\right]}.$$

con $C = 2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}$.

$$\text{Cov}(X, Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x - \mu_X)(y - \mu_Y) f_{XY}(x, y) dxdy$$

Cambio de variables $u = (x - \mu_X)/\sigma_X$, $v = (y - \mu_Y)/\sigma_Y$ con $J = \sigma_X\sigma_Y$

$$\frac{\sigma_X\sigma_Y}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} uv \exp\left[-\frac{1}{2(1-\rho^2)}\left(u^2 + v^2 - 2\rho uv\right)\right] dudv$$

## Continuación.

Para calcular la integral, completamos cuadrados en u. Usamos la igualdad:

$$u^2 + v^2 - 2\rho uv = (u - \rho v)^2 + v^2\left(1 - \rho^2\right)$$

Entonces, podemos reemplazar en la integral que teníamos

$$\frac{\sigma_X \sigma_Y}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} uv \exp\left[-\frac{1}{2\left(1-\rho^2\right)}\left(u^2 + v^2 - 2\rho uv\right)\right] dudv$$

$$\frac{\sigma_X \sigma_Y}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} v \exp\left(-v^2/2\right)\left(\int_{-\infty}^{\infty} u \exp\left[-\frac{1}{2\left(1-\rho^2\right)}(u-\rho v)^2\right] du\right)$$

La integral interior es la esperanza de una v.a. $\mathcal{N}\left(\rho v, \left(1-\rho^2\right)\right)$, a la que únicamente le falta la constante de normalización $\left[2\pi\left(1-\rho^2\right)\right]^{-1/2}$, y por lo tanto tenemos

$$\mathrm{Cov}(X,Y) = \rho\sigma_X\sigma_Y \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} v^2 e^{-v^2/2} dv = \rho\sigma_X\sigma_Y$$

ya que la varianza de la normal estándar es 1

## Continuación.

Para calcular la integral, completamos cuadrados en u. Usamos la igualdad:

$$u^2 + v^2 - 2\rho uv = (u - \rho v)^2 + v^2\left(1 - \rho^2\right)$$

Entonces, podemos reemplazar en la integral que teníamos

$$\frac{\sigma_X \sigma_Y}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} uv \exp\left[-\frac{1}{2\left(1-\rho^2\right)}\left(u^2 + v^2 - 2\rho uv\right)\right] dudv$$

$$\frac{\sigma_X \sigma_Y}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} v \exp\left(-v^2/2\right)\left(\int_{-\infty}^{\infty} u \exp\left[-\frac{1}{2\left(1-\rho^2\right)}(u-\rho v)^2\right] du\right)$$

La integral interior es la esperanza de una v.a. $\mathcal{N}\left(\rho v, \left(1-\rho^2\right)\right)$, a la que únicamente le falta la constante de normalización $\left[2\pi\left(1-\rho^2\right)\right]^{-1/2}$, y por lo tanto tenemos

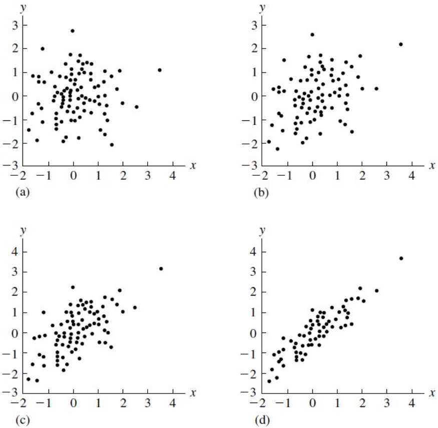
$$\mathrm{Cov}(X, Y) = \rho\sigma_X\sigma_Y \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} v^2 e^{-v^2/2} dv = \rho\sigma_X\sigma_Y$$

ya que la varianza de la normal estándar es 1 $\square$

# Interpretación de la correlación

- Si las variables aleatorias están positivamente asociadas, es decir, si cuando $X$ es mayor que su media, $Y$ tiende a ser mayor que su media, entonces, la correlación será positiva. Si la asociación es negativa, es decir, cuando $X$ es mayor que su media, $Y$ tiende a ser menor que su media, la correlación es negativa.
- La correlación indica la fuerza de la relación lineal entre dos variables aleatorias
- La correlación no varía si se cambian las unidades en las que se miden $X$ e $Y$, en cambio la covarianza si.

# Correlación de normales bivariadas




100 observaciones de normales bivariadas independientes, (a) $\rho = 0$, (b) $\rho = .3$, (c) $\rho = .6$, (d) $\rho = .9$.

Figura extraída del libro de John A. Rice "Mathematical Statistics and Data Analysis ".