PROBABILIDAD I

alessandro.ferreiro@vam.es

coordinador:
Julián de la Horra

1^{er} PARCIAL – 15 marzo
2º PARCIAL – 26 abril

} media ≥ 5 y nota en cada uno ≥ 4

APROBADOS

T1. – Espacios de probabilidad
- Teorema de Bayes

T2. – Variables aleatorias
- Función de distribución
- Modelo de probabilidades (quizá en tema 3)
  - normal o gaussiana
  - Binomial
  - Poisson

T3. – Vectores aleatorios
- Es el estudio de la relación entre diferentes variables aleatorias
- Covarianza y correlación
- Dependencia e independencia de variables aleat.

T4. – Teorema Central y Convergencias de variables aleatorias
- Teorema Central Límite
- Convergencias de variables aleatorias (OPCIONAL)

TEMA 1 ESPACIOS DE PROBABILIDAD

DEFINICIÓN: Un ESPACIO MUESTRAL Ω es el conjunto de todos los posibles resultados del fenómeno aleatorio que deseamos estudia

Ejemplos: (1.) Espacio muestral Ω de lanzar una moneda.

$$\Omega = \{\text{"cara"}, \text{"cruz"}\}$$

(2.) Lanzar un dado

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

DEFINICIÓN: Una σ-ÁLGEBRA DE SUCESO, $\mathcal{S}$, es un subconjunto de partes de Ω que verifica las siguientes propiedades:

a) Si $A \in \mathcal{S} \Rightarrow A^c = \Omega \setminus A \in \mathcal{S}$.

b) Si $A_1, A_2, \dots, \in \mathcal{S} \Rightarrow \bigcup_{m \in \mathbb{N}}^{\infty} A_m \in \mathcal{S}$.

La σ-álgebra representa las informaciones disponibles sobre el fenómeno aleatorio

Ejemplos: (1.) Si $\Omega = \{1,2,3,4,5,6\}$ es el espacio muestral asociado al fenómeno aleatorio "lanzar un dado", entonces $\mathcal{S} = \mathcal{P}(\Omega)$.
Para cualquier evento, e.d. $A \in \mathcal{P}(\Omega)$, sabemos su probabilid de ocurrencia.

PROPIEDADES:

1.- $\emptyset \in \mathcal{S}$ y $\Omega \in \mathcal{S}$ (si $\mathcal{S} \neq \emptyset$) $\Rightarrow$ la σ-álgebra mínima es $\{\emptyset, \Omega$

2.- Si $A_1, A_2, \dots, \in \mathcal{S} \Rightarrow \bigcap_{m=1}^{\infty} A_m \in \mathcal{S}$

Demostración 2:

$A_1^c, A_2^c, \dots, \in \mathcal{S}$ por a) $\overset{(b)}{\Rightarrow} \bigcup_m A_m^c \in \mathcal{S}$ $\qquad \bigcup_m A_m^c = \left(\bigcap_m A_m\right)^c$

$\underset{(a)}{\Rightarrow} \bigcap_m A_m \in \mathcal{S}$

3.- Si $A_1, A_2 \in S \Rightarrow A_1 \setminus A_2 \in S$, $A_1 \setminus A_2 = A_1 \cap A_2^c$

4.- Si $A_1, A_2, \dots \in S \Rightarrow \overline{\lim_{m\to\infty}} A_m := \bigcap_{m=1}^{\infty} \left[\bigcup_{i=m}^{\infty} A_i \right] \in S$ y

también $\underline{\lim_{m\to\infty}} A_m := \bigcup_{m=1}^{\infty} \left[\bigcap_{i=m}^{\infty} A_i \right] \in S$.

<u>Observación</u>: $\lim_{m\to\infty} A_m$ se define como $\lim A_m$ si $\overline{\lim} A_m = \underline{\lim} A_m = A$

y si $\exists \lim_{m\to\infty} A_m \Rightarrow \lim_{m\to\infty} A_m \in S$.

5.- La intersección de una familia de σ-álgebras S_α es una σ-álgebra, e.d. $\bigcap_\alpha S_\alpha$ es σ-álgebra.

<u>DEFINICIÓN</u>: Se define la σ-álgebra generada por un subconjunto $B \subset \mathcal{P}(\Omega)$: $\quad \sigma(B) := \bigcap \{ S : S \supset B \text{ y } S\ \sigma\text{-álgebra} \}$

- $\sigma(B)$ es la σ-álgebra más pequeña que contiene a B.

▶ Si Ω es finito ó discreto, normalmente la σ-álgebra utilizada es $S = \mathcal{P}(\Omega)$

▶ Si Ω es continua, por ejemplo $\Omega = \mathbb{R}$, la σ-álgebra más utilizada es la σ-álgebra de BOREL, que es la $\sigma\big((-\infty, x] : x \in \mathbb{R} \big)$.

▶ Si $\Omega = \mathbb{R}^n$, se utiliza la σ-álgebra $\sigma\big(\{ (-\infty, x_1] \times \dots \times (-\infty, x_n] : x_1, \dots, x_n \in \mathbb{R} \} \big)$

<u>DEFINICIÓN</u>: Sea Ω un espacio muestral y S un σ-álgebra sobre Ω. Una FUNCIÓN DE PROBABILIDAD $\mathbb{P}$, es una función $\mathbb{P}: S \to [0,1]$ que verifica:

a) $\mathbb{P}(\Omega) = 1$.

b) Si $A_1, A_2, \ldots \in S$ <u>disjuntos</u> $\Rightarrow \mathbb{P}\left(\bigcup_{m=1}^{\infty} A_m\right) = \sum_{m=1}^{\infty} \mathbb{P}(A_m)$.

espacio muestral, σ-álgebra, funci[ón] de pro[ba]bilid[ad]

<u>DEFINICIÓN</u>: Un ESPACIO DE PROBABILIDAD es la terna $(\Omega, S, \mathbb{P})$

<u>PROPIEDADES</u>:

1.- Si $A \in S$, $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$

<u>demostración</u>

$A \cup A^c = \Omega$ y $A \cap A^c = \emptyset$ (e.d. son disjuntos)

por lo anterior (b) $\underbrace{\mathbb{P}(A \cup A^c)}_{\mathbb{P}(\Omega) = 1} = \mathbb{P}(A) + \mathbb{P}(A^c) \Big\} \Rightarrow \mathbb{P}(A^c) = 1 - \mathbb{P}(A)$

2.- $\mathbb{P}(\emptyset) = 0$

3.- Si $A_1, A_2 \in S$, y $A_1 \subset A_2 \Rightarrow \mathbb{P}(A_1) \leq \mathbb{P}(A_2)$ y además

$\mathbb{P}(A_2 \setminus A_1) = \mathbb{P}(A_2) - \mathbb{P}(A_1)$

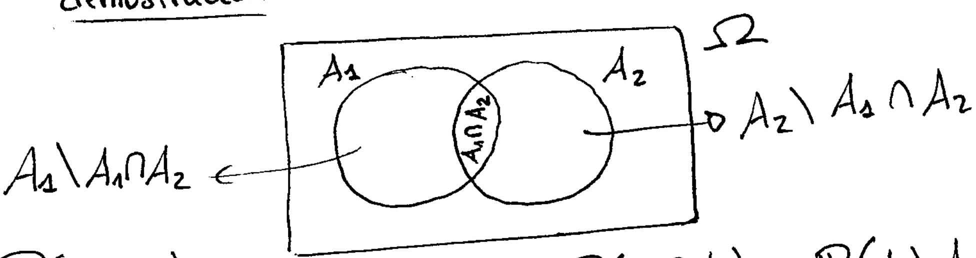
<u>demostración</u> unión de conjuntos disjuntos

$A_2 = (A_2 \setminus A_1) \cup A_1 \Rightarrow \mathbb{P}(A_2) = \overbrace{\mathbb{P}(A_1) + \mathbb{P}(A_2 \setminus A_1)}^{\mathbb{P}(A_2)} \Rightarrow$

$\Rightarrow \mathbb{P}(A_2 \setminus A_1) = \mathbb{P}(A_2) - \mathbb{P}(A_1)$

4.- Si $A_1, A_2 \in S$, $\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2)$

<u>demostración</u>



$\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1 \setminus A_1 \cap A_2) + \mathbb{P}(A_1 \cap A_2) + \mathbb{P}(A_2 \setminus A_1 \cap A_2) =$

$$\Rightarrow \mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2)$$

5.- Si $A_1, A_2, A_3 \in \mathcal{S} \Rightarrow \mathbb{P}(A_1 \cup A_2 \cup A_3) = \mathbb{P}(A_1) + \mathbb{P}(A_2) + \mathbb{P}(A_3) -$

$$- \mathbb{P}(A_1 \cap A_2) - \mathbb{P}(A_1 \cap A_3) - \mathbb{P}(A_2 \cap A_3) + \mathbb{P}(A_1 \cap A_2 \cap A_3)$$

6.- Si $A_1, \dots, A_n \in \mathcal{S} \Rightarrow \mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathbb{P}(A_i) - \sum_{i_1 < i_2} \mathbb{P}(A_{i_1} \cap A_{i_2}) +$

$$+ \sum_{i_1 < i_2 < i_3} \mathbb{P}(A_{i_1} \cap A_{i_2} \cap A_{i_3}) - \cdots + (-1)^{n+1} \mathbb{P}\left(\bigcap_{i=1}^{n} A_i\right)$$

7.- Si $A_1, \dots, A_n \in \mathcal{S} \Rightarrow \mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} \mathbb{P}(A_i)$

<u>TEOREMA</u>: a) Sea $A_1, A_2, \dots \in \mathcal{S}$ una sucesión <u>creciente</u> $(A_n \subseteq A_{n+1} \ \forall n \in \mathbb{N})$

Entonces, $\mathbb{P}\left(\lim_{n \to \infty} A_n\right) = \lim_{n \to \infty} \mathbb{P}(A_n)$

b) Sea $A_1, A_2, \dots \in \mathcal{S}$ una sucesión <u>decreciente</u> $(A_n \supseteq A_{n+1}, \forall n \in \mathbb{N}$

Entonces, $\mathbb{P}\left(\lim_{n \to \infty} A_n\right) = \lim_{n \to \infty} \mathbb{P}(A_n)$

<u>demostración</u>

a) $\lim_{n \to \infty} A_n = \bigcup_{n=1}^{\infty} A_n$ (¿Por qué? $\overline{\lim} A_n = \bigcap_{n=1}^{\infty} \left(\bigcup_{i=n}^{\infty} A_i\right) = \bigcap_{n=1}^{\infty} \left(\bigcup_{i=1}^{\infty} A_i\right) = \bigcup_{i=1}^{\infty} A_i$)

sucesión creciente

$B_1 := A_1$; $B_2 := A_2 \setminus A_1$; $B_3 := A_3 \setminus A_2$; ... B_i disjuntos !

$$\mathbb{P}\left(\lim_{n \to \infty} A_n\right) = \mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \mathbb{P}\left(\bigcup_{n=1}^{\infty} B_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(B_n) = \lim_{n \to \infty} \sum_{i=1}^{n} \mathbb{P}(B_i) =$$

$$= \lim_{n \to \infty} \mathbb{P}\left(\bigcup_{i=1}^{n} B_i\right) = \lim_{n \to \infty} \mathbb{P}(A_n) \quad \square$$

TEOREMA: Sean $A_1, A_2, \ldots \in S$. Entonces $\mathbb{P}(\underline{\lim}\, A_n) \underset{(1)}{\leq} \underline{\lim}\, \mathbb{P}(A_n)$ y, además, $\overline{\lim}\, \mathbb{P}(A_n) \overset{(2)}{\leq} \mathbb{P}(\overline{\lim}\, A_n)$.

Demostración (1)

$$\mathbb{P}(\underline{\lim}\, A_n) = \mathbb{P}\left(\bigcup_{n=1}^{\infty} \bigcap_{i=n}^{\infty} A_i\right) = \mathbb{P}\left(\lim_{n\to\infty} \bigcap_{i=n}^{\infty} A_i\right) = \lim_{n\to\infty} \mathbb{P}\left(\bigcap_{i=n}^{\infty} A_i\right) \leq (*)$$

por otro lado: $\forall i \geq n$, $\mathbb{P}\left(\bigcap_{i=n}^{\infty} A_i\right) \leq \mathbb{P}(A_i)$. Por tanto obtenemos:

$$\mathbb{P}\left(\bigcap_{i=n}^{\infty} A_i\right) \leq \inf_{i\geq n} \mathbb{P}(A_i)$$

Entonces:

$$(*) \leq \lim_{n\to\infty} \underbrace{\inf_{i\geq n} \mathbb{P}(A_i)}_{\substack{\shortparallel \\ a_n \\ a_1 \leq a_2 \leq a_3 \cdots}} = \sup_{n\geq 1} \inf_{i\geq n} \mathbb{P}(A_i) = \underline{\lim_{n\to\infty}}\, \mathbb{P}(A_n)$$

□

Observación: como $\underline{\lim}\, \mathbb{P}(A_n) \leq \overline{\lim}\, \mathbb{P}(A_n) \Rightarrow$ por el teorema anterior $\Rightarrow$

$$\Rightarrow \mathbb{P}(\underline{\lim}\, A_n) \overset{(1)}{\leq} \underline{\lim}\, \mathbb{P}(A_n) \leq \overline{\lim}\, \mathbb{P}(A_n) \overset{(2)}{\leq} \mathbb{P}(\overline{\lim}\, A_n)$$

COROLARIO: Si $A_1, A_2, \ldots \in S$ son una sucesión de eventos que converge, e.d., $\exists \lim A_n \Rightarrow \mathbb{P}(\lim A_n) = \lim \mathbb{P}(A_n)$

ESPACIO MUESTRAL DISCRETO

- $\Omega = \{a_1, a_2, \ldots\}$
- $S = \mathcal{P}(\Omega)$
- $\mathbb{P} : \mathcal{P}(\Omega) \longrightarrow [0,1]$

DEFINICIÓN: Una FUNCIÓN DE MASA $p : \Omega \longrightarrow [0,1]$ que verifica $\sum_{n=1}^{\infty} p(a_n) = 1$
Una función de masa p define una probabilidad por medio de la relación: si $A \subseteq \Omega$, $\mathbb{P}(A) := \sum_{a\in A} p(a)$.

Si $\mathbb{P}$ es la probabilidad sobre un espacio muestral discreto $\Rightarrow$
$\Rightarrow \forall a \in \Omega$, $p(a) := \mathbb{P}(\{a\})$.

Si Ω es finito, existe una función de masa estándar (modelo de LAPLACE)

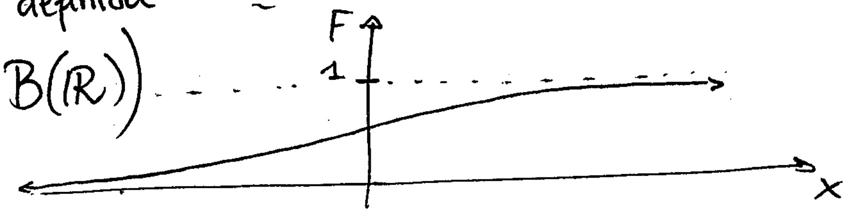
$$p(a) := \frac{1}{\text{card}(\Omega)}, \ \forall a \in A \Rightarrow \mathbb{P}(A) = \sum_{a\in A} p(a) = \frac{1}{\text{card}(\Omega)} \sum_{a\in A} 1 = \frac{\text{card}(A)}{\text{card}(\Omega)} = \frac{\#\text{casos favo}}{\#\text{casos posibles}}$$

MODELOS SOBRE $\mathbb{K}$

- $\Omega = \mathbb{R}$
- $S = \mathcal{B}(\mathbb{R}) = \sigma\left(\{(-\infty, x] : x \in \mathbb{R}\}\right)$

Observación: ¿Por qué no se puede utilizar una función de masa sobre $\Omega = \mathbb{R}$?
$\sum_{a \in \Omega} p(a) \neq 1$ no puede ser porque una serie de una familia infinita no numerable > 0 no converge, es $+\infty$!

DEFINICIÓN: La FUNCIÓN DE DISTRIBUCIÓN asociada a una probabilidad $\mathbb{P}$ sobre $\Omega = \mathbb{R}$ es la función $F : \mathbb{R} \longrightarrow [0,1]$ definida por $F(x) := \mathbb{P}\left((-\infty, x]\right)$ $\left(S = \mathcal{B}(\mathbb{R})\right)$



PROPIEDADES:

a) $\lim_{x \to -\infty} F(x) = 0$

b) $\lim_{x \to \infty} F(x) = 1$

c) F es una función creciente

d) F es continua por la derecha.

demostración a

$$\lim_{x \to -\infty} F(x) = \lim_{x \to -\infty} \mathbb{P}(-\infty, x]) = \mathbb{P}\Big(\overbrace{\lim_{x \to -\infty} (-\infty, x]}^{\emptyset}\Big) = \mathbb{P}(\{\emptyset\}) = 0$$

demostración b

$$\lim_{x \to \infty} F(x) = \lim_{x \to \infty} \mathbb{P}\left((-\infty, x]\right) = \mathbb{P}\Big(\overbrace{\lim_{x \to \infty} (-\infty, x]}^{\mathbb{R}}\Big) = \mathbb{P}(\{\mathbb{R}\}) = \mathbb{P}(\Omega) = 1$$

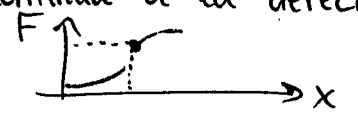
demostración c

Si $a \leq b$, $F(b) = \mathbb{P}\left((-\infty, b]\right) = \mathbb{P}\left((-\infty, a] \cup (a, b]\right) =$

$$= \mathbb{P}\left((-\infty, a]\right) + \mathbb{P}\left((a, b]\right) \geq \mathbb{P}\left((-\infty, a]\right) = F(a)$$

demostración d

[F es continua a la derecha en x, $F : \mathbb{R} \to \mathbb{R}$ si $\lim_{\substack{h \to 0 \\ h > 0}} F(x+h) = F(x)$]



$$\lim_{\substack{h \to 0 \\ h > 0}} F(x+h) = \lim_{h \to 0^+} F(x+h) = \lim_{h \to 0^+} \mathbb{P}\left((-\infty, x+h]\right) = \mathbb{P}\left(\lim_{h \to 0^+} (-\infty, x+h]\right) =$$

$$\overset{\text{decreciente}}{=} \mathbb{P}\Big(\underbrace{\bigcap_{h > 0} (-\infty, x+h]}_{(-\infty, x]}\Big) = \mathbb{P}\left((-\infty, x]\right) = F(x)$$

Observación: Si F es una función que verifica las anteriores propiedades, entonces se puede definir una probabilidad $\mathbb{P}$ asociada a F sobre $\mathcal{B}(\mathbb{R})$ como $\mathbb{P}((-\infty, x]) := F(x)$.

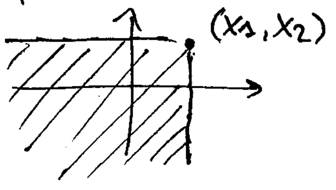
▷ Si $\Omega = \mathbb{R}^n$, $\mathcal{S} = \mathcal{B}(\mathbb{R}^n) = \sigma\left(\{(-\infty, \dots, x_1] \times \cdots \times (-\infty, \dots, x_n] : x_i \in \mathbb{R}\}\right)$

MODELOS SOBRE $\mathbb{R}^n$

DEFINICIÓN: Se define la FUNCIÓN DE DISTRIBUCIÓN asociada a $\mathbb{P}$ sobre $\mathcal{B}(\mathbb{R}^n)$ como $F: \mathbb{R}^n \longrightarrow [0,1]$.

$$F(x_1, \dots, x_n) := \mathbb{P}\left((-\infty, x_1] \times \cdots \times (-\infty, x_n]\right)$$

Ejemplo: $\mathbb{R}^2$



PROPIEDADES:

a) $\lim\limits_{(x_1,\dots,x_n) \to -\infty} F(x_1, \dots, x_n) = 0$ $\quad (x_1,\dots,x_n) \to$ alguno (al menos uno)

b) $\lim\limits_{(x_1,\dots,x_n) \to \infty} F(x_1, \dots, x_n) = 1$ $\quad (x_1,\dots,x_n) \to$ todos

c) F es creciente en cada una de sus variables.

d) F es continua por la derecha en cada una de sus variables.

Observación: Si F es una función de distribución, es decir, F verifica las propiedades anteriores, entonces puedo definir una probabilidad sobre $\mathcal{B}(\mathbb{R}^n)$ por $\mathbb{P}\left((-\infty, x_1] \times \cdots \times (-\infty, x_n]\right) := F(x_1, \dots, x_n)$

ROBABILIDAD CONDICIONADA Y EVENTOS INDEPENDIENTES

EFINICIÓN: La probabilidad de un evento A CONDICIONADA a un evento B e define como: $\mathbb{P}(A|B) := \dfrac{\mathbb{P}(A\cap B)}{\mathbb{P}(B)}$

Ejemplo: $\Omega = \{1,2,3,4,5,6\}$, $\mathcal{S} = \mathcal{P}(\Omega)$, $\mathbb{P}(\{x_i\}) = \frac{1}{6}$ $\forall i=1,\dots,6$.

$A =$ "sale 1"

$B =$ "sale un número impar"

$$\left.\begin{array}{l}\mathbb{P}(A) = \frac{1}{6} \\ \mathbb{P}(B) = \frac{1}{2}\end{array}\right| \quad \mathbb{P}(A|B) = \frac{\mathbb{P}(A\cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A)}{\mathbb{P}(B)} = \frac{1/6}{1/2} = 1/3$$

DEFINICIÓN: Los eventos A y B son INDEPENDIENTES si $\mathbb{P}(A\cap B) = \mathbb{P}(A)\mathbb{P}(B)$

ROPOSICIÓN: Los eventos A y B son independientes $\Longleftrightarrow$

$\Longleftrightarrow \mathbb{P}(A|B) = \mathbb{P}(A)$ $\left(\text{ó } \mathbb{P}(B|A) = \mathbb{P}(B)\right)$.

demostración:

$\boxed{\Rightarrow}$ Si A y B son independientes: $\mathbb{P}(A|B) = \dfrac{\mathbb{P}(A\cap B)}{\mathbb{P}(B)} = \dfrac{\mathbb{P}(A)\cdot\mathbb{P}(B)}{\mathbb{P}(B)} = \mathbb{P}(A)$

$\boxed{\Leftarrow}$ $\mathbb{P}(A\cap B) = \mathbb{P}(A|B)\cdot\mathbb{P}(B) = \mathbb{P}(A)\cdot\mathbb{P}(B)$

$\left(\mathbb{P}(B|A)\cdot\mathbb{P}(A)\right)$

Ejemplo: lanzar dos veces una moneda

$\Omega = \{(c,c), (c,x), (x,c), (x,x)\}$

$A =$ "sale cara en el primer lanzamiento"

$B =$ "sale cara en el segundo lanzamiento"

$\mathbb{P}(A) = \mathbb{P}(\{(c,c),(c,x)\}) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$

$\mathbb{P}(B) = \mathbb{P}(\{(c,c),(x,c)\}) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$

$\mathbb{P}(A\cap B) = \mathbb{P}(\{(c,c)\}) = \frac{1}{4}$

Como $\mathbb{P}(A\cap B) = \mathbb{P}(A).\mathbb{P}(B)$, ya que $\frac{1}{4} = \frac{1}{2}\cdot\frac{1}{2} \Rightarrow A$ y B son independientes.

Ejemplo: lanzar dos veces una moneda

$$\Omega = \{(c,c), (c,x), (x,c), (x,x)\}$$

D = "sale exactamente una cara en los dos lanzamientos"

¿Son D y B independientes?

$$\mathbb{P}(D) = \mathbb{P}(\{(c,x),(x,c)\}) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

$$\mathbb{P}(B \cap D) = \mathbb{P}(\{(x,c)\}) = \frac{1}{4}$$

Como $\mathbb{P}(B \cap D) = \mathbb{P}(B) \cdot \mathbb{P}(D)$ porque $\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4} \Rightarrow$ B y D son independien

DEFINICIÓN: Los eventos $\{A_i\}_{i \in I}$ son INDEPENDIENTES si $\forall \{i_1, i_2, \dots, i_k\} \subset$

tenemos $\mathbb{P}(A_{i_1} \cap \dots \cap A_{i_k}) = \mathbb{P}(A_{i_1}) \bullet \dots \bullet \mathbb{P}(A_{i_k})$

Ejemplo:

Si $\operatorname{card}(I) = 3$ $\qquad$ A, B, C son independientes si:

$$\begin{cases} \mathbb{P}(A \cap B \cap C) = \mathbb{P}(A) \cdot \mathbb{P}(B) \cdot \mathbb{P}(C) \\ \mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B) \\ \mathbb{P}(A \cap C) = \mathbb{P}(A) \cdot \mathbb{P}(C) \\ \mathbb{P}(B \cap C) = \mathbb{P}(B) \cdot \mathbb{P}(C) \end{cases}$$

EGLAS PARA CALCULAR PROBABILIDADES

TEOREMA: (Regla de multiplicación): Sean $A_1, A_2, \ldots, A_n$ eventos, entonces,

$$\mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_n) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2|A_1) \cdot \mathbb{P}(A_3|A_1 \cap A_2) \cdot \ldots \cdot \mathbb{P}(A_n|A_1 \cap \ldots \cap A_{n-1})$$

<u>demostración</u>

Si $n=3$,

$$\underbrace{\mathbb{P}(A_3|A_1\cap A_2)}_{\frac{\mathbb{P}(A_1\cap A_2\cap A_3)}{\mathbb{P}(A_1\cap A_2)}} \cdot \underbrace{\mathbb{P}(A_2|A_1)}_{\frac{\mathbb{P}(A_2\cap A_1)}{\mathbb{P}(A_1)}} \cdot \mathbb{P}(A_1) = \frac{\mathbb{P}(A_1\cap A_2\cap A_3)}{\cancel{\mathbb{P}(A_1\cap A_2)}} \cdot \frac{\cancel{\mathbb{P}(A_2\cap A_1)}}{\cancel{\mathbb{P}(A_1)}} \cdot \cancel{\mathbb{P}(A_1)}$$

$$= \mathbb{P}(A_1 \cap A_2 \cap A_3) \quad \square$$

<u>Ejemplo</u>: Una urna con $\left\{\begin{array}{l}20 \text{ bolas blancas} \\ 10 \text{ bolas negras}\end{array}\right\}$ y hacemos 3 extracciones sin devolución.

- ¿Cuál es la probabilidad de que las 3 bolas sean blancas?

$A_i =$ "La i-ésima bola es blanca" $\to$ buscamos $\mathbb{P}(A_1 \cap A_2 \cap A_3)$

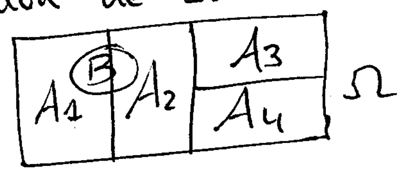
$$\mathbb{P}(A_1 \cap A_2 \cap A_3) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2|A_1) \cdot \mathbb{P}(A_3|A_1\cap A_2) = \frac{20}{30} \cdot \frac{19}{29} \cdot \frac{18}{28}$$

TEOREMA (Regla de la probabilidad total): Sean $A_1, \ldots, A_n$ eventos tales que:

a) $\bigcup_{i=1}^{n} A_i = \Omega$

b) $A_i \cap A_j = \emptyset \quad \forall i,j : i \neq j$ (disjuntos)

Partición de Ω



Entonces:

$$\mathbb{P}(B) = \sum_{i=1}^{n} \mathbb{P}(A_i) \cdot \mathbb{P}(B|A_i)$$

<u>demostración</u>:

$$\mathbb{P}(B) = \mathbb{P}\Big(B \cap \big(\overbrace{\bigcup_{i=1}^{n} A_i}^{\Omega}\big)\Big) = \mathbb{P}\Big(\bigcup_{i=1}^{n} (B \cap A_i)\Big) \overset{A_i \text{ disjuntos}}{=} \sum_{i=1}^{n} \mathbb{P}(B \cap A_i) =$$

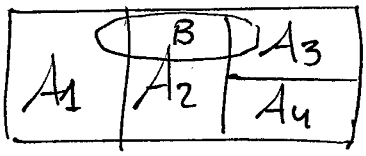
$$= \sum_{i=1}^{n} \mathbb{P}(B|A_i) \cdot \mathbb{P}(A_i) \quad \square$$

TEOREMA (Regla de Bayes): Sean $A_1, \ldots, A_n$ eventos tales que:

a) $\bigcup_{i=1}^{n} A_i = \Omega$

b) $A_i \cap A_j = \emptyset \quad \forall i,j : i \neq j$ (disjuntos)

} Partición de Ω



Ω

Entonces:

$$\mathbb{P}(A_i \mid B) = \frac{\mathbb{P}(A_i) \cdot \mathbb{P}(B \mid A_i)}{\sum_{k=1}^{n} \mathbb{P}(A_k) \cdot \mathbb{P}(B \mid A_k)}$$

demostración

$$\mathbb{P}(A_i \mid B) = \frac{\mathbb{P}(A_i) \cdot \mathbb{P}(B \mid A_i)}{\mathbb{P}(B)} =$$

$$= \frac{\mathbb{P}(A_i) \cdot \mathbb{P}(B \mid A_i)}{\sum_{k=1}^{n} \mathbb{P}(A_k) \cdot \mathbb{P}(B \mid A_k)} \qquad \square$$

Regla de Bayes

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(B \mid A) \cdot \mathbb{P}(A)}{\mathbb{P}(B)}$$

demostración

$$\frac{\mathbb{P}(B \mid A) \cdot \mathbb{P}(A)}{\mathbb{P}(B)} = \frac{\frac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)} \cdot \mathbb{P}(A)}{\mathbb{P}(B)} = \mathbb{P}(A \mid B)$$

$\square$

Observación ¿Por qué $\mathbb{P}(A \mid B) \neq \mathbb{P}(B \mid A)$?

$A =$ "sale 1" $\qquad \mathbb{P}(A) = 1/6$

$B =$ "sale número impar" $\qquad \mathbb{P}(B) = 1/2$

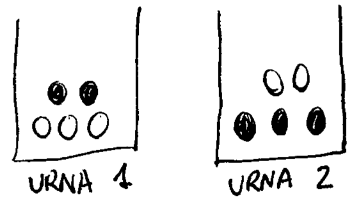
$\mathbb{P}(A \mid B) = 1/3$ $\qquad \mathbb{P}(A \mid B) \neq \mathbb{P}(B \mid A)$

$\mathbb{P}(B \mid A) = 1$

Según la Regla de Bayes:

$$\mathbb{P}(B \mid A) = \frac{\mathbb{P}(A \mid B) \cdot \mathbb{P}(B)}{\mathbb{P}(A)} = \frac{1/3 \cdot 1/2}{1/6} = 1$$

Ejemplo: Tenemos dos urnas: urna 1 tiene 3 bolas blancas y 2 negras; urna 2 tiene 2 bolas blancas y 3 negras.



$U1$ = "meter la mano en la urna 1"

$U2$ = "meter la mano en la urna 2"

$\mathbb{P}(U1) = 1/3 \quad , \quad \mathbb{P}(U2) = 2/3$

1. ¿Cuál es la probabilidad que la bola extraída sea blanca?

2. Si sabemos que la bola extraida es blanca, ¿cuál es la probabilidad que la extración se haya hecho en la U1?

1. $$\mathbb{P}(\overbrace{\text{"bola sea blanca"}}^{B}) = \underbrace{\mathbb{P}(B|U1)}_{3/5} \cdot \underbrace{\mathbb{P}(U1)}_{1/3} + \underbrace{\mathbb{P}(B|U2)}_{2/5} \cdot \underbrace{\mathbb{P}(U2)}_{2/5} = \frac{7}{15}$$

2. $$\mathbb{P}(U1|B) = \frac{\mathbb{P}(B|U1) \cdot \mathbb{P}(U1)}{\mathbb{P}(B)} = \frac{3/5 \cdot 1/3}{7/15} = \frac{3}{7}$$

Ejemplo: Tenemos plantas de flores rojas homocigóticas y heterocigóticas y flores blancas homocigóticas:

$$\text{rojas} \begin{cases} RR & \text{homocigóticas} \\ Rb & \text{heterocigóticas} \end{cases}$$

$\mathbb{P}(Rb|\text{"flores rojas"}) = 70\%$ blancas: bb

- Nos traen una planta de flores rojas.
- La cruzamos con una planta bb (flores blancas) y obtenemos 5 plantas de flores rojas.
- ¿Cuál es la probabilidad que la planta fuera: RR?

$\mathbb{P}(\text{"5 plantas de flores rojas"}|RR) = ? = 1$ (cruzando una RR con una $bb \to Rb$)

$\mathbb{P}(\text{"5 plantas de flores rojas"}|Rb) = ? = \left(\frac{1}{2}\right)^5$ (cruzando una Rb con $bb \to \begin{cases} Rb \\ bb \end{cases}$)

$$\mathbb{P}(RR|\text{"5 cruces de flores rojas"}) = \frac{\overbrace{\mathbb{P}(\text{"5 cruces de flores rojas"}|RR)}^{1} \cdot \overbrace{\mathbb{P}(RR)}^{0'3}}{\left[\underbrace{\mathbb{P}(RR)}_{0'3} \cdot \underbrace{\mathbb{P}(\text{" "}|RR)}_{1} + \underbrace{\mathbb{P}(Rb)}_{0'7} \cdot \underbrace{\mathbb{P}(\text{" "}|Rb)}_{(\frac{1}{2})^5}\right]^{(*)}} =$$

$$= 0'932$$

(*) denominador (Regla prob. total) $= \mathbb{P}(\text{"5 cruces de flores rojas"})$

TEMA 2: VARIABLES ALEATORIAS

DEFINICIÓN: Una VARIABLE ALEATORIA X sobre un espacio de probabilidad $(\Omega, S, \mathbb{P})$ es una función $X : \Omega \longrightarrow \mathbb{R}$ tal que $\underset{\{w \in \Omega :\, X(w) = A\}}{\underset{\parallel}{X^{-1}(A)}} \in S$, $\forall A \in B(\mathbb{R}) = \sigma(\{(-\infty, x] : x \in \mathbb{R}\})$.

DEFINICIÓN: Se define $\mathbb{P}_X$, una PROBABILIDAD SOBRE $B(\mathbb{R})$, como:

$$\mathbb{P}_X(A) := \mathbb{P}(X^{-1}(A)) \quad , \forall A \in B(\mathbb{R})$$

Con esta definición, $(\mathbb{R}, B(\mathbb{R}), \mathbb{P}_X)$ es un espacio de probabilidad.

PROPIEDADES:

i) $\mathbb{P}_X(A) \in [0,1]$, $\forall A \in B(\mathbb{R})$ ($\mathbb{P}$ es una probabilidad)

ii) $\mathbb{P}_X(\mathbb{R}) = 1$

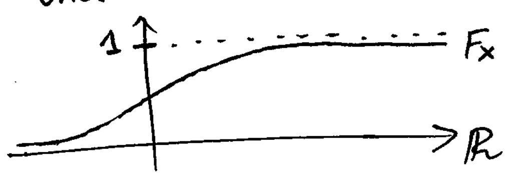
demostración: $\mathbb{P}_X(\mathbb{R}) = \mathbb{P}(X^{-1}(\mathbb{R})) = \mathbb{P}(\Omega) = 1$

iii) Si $A_1, A_2, \ldots, A_n \in B(\mathbb{R})$ son disjuntos.

$$\mathbb{P}_X\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}_X(A_i)$$

demostración: $\mathbb{P}\left(X^{-1}\left(\bigcup_{i=1}^{\infty} A_i\right)\right) = \mathbb{P}\left(\bigcup_{i=1}^{\infty} X^{-1}(A_i)\right) = \sum_{i=1}^{\infty} \overbrace{\mathbb{P}(X^{-1}(A_i))}^{\mathbb{P}_X(A_i)}$

DEFINICIÓN: La FUNCIÓN DE DISTRIBUCIÓN de una variable aleatoria X es $x \in \mathbb{R}$, $F_X(x) := \mathbb{P}_X((-\infty, x])$



$X^{-1}((-\infty, x]) = \{\omega \in \Omega : X(\omega) \in (-\infty, x]\} =$
$= \{\omega \in \Omega : X(\omega) \leq x\}$

$$\boxed{F_X(x) := \mathbb{P}_X((-\infty, x]) = \mathbb{P}(X^{-1}((-\infty, x])) = \mathbb{P}(X \leq x)}$$

variable aleatoria

valor

Ejemplo: tenemos 3 lanzamientos de moneda.

$$\Omega = \{ccc, ccx, cxc, xcc, cxx, xcx, xxc, xxx\}$$

$$\mathbb{P}(\{\omega_i\}) = \frac{1}{8}$$

$$S = \mathbb{P}(\Omega) \Rightarrow (\Omega, S = \mathbb{P}(\Omega), \mathbb{P})$$

▸ variable aleatoria X = "nº de caras" $\in \{0,1,2,3\}$

$$\mathbb{P}(X=0) = \mathbb{P}_X(\{0\}) = \frac{1}{8}$$

$$\mathbb{P}(X=1) = \mathbb{P}_X(\{1\}) = \frac{3}{8}$$

$$\mathbb{P}(X=2) = \mathbb{P}_X(\{2\}) = \frac{3}{8}$$

$$\mathbb{P}(X=3) = \mathbb{P}_X(\{3\}) = \frac{1}{8}$$

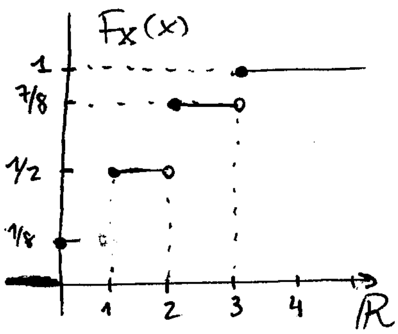
$$(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}_X)$$

$\downarrow$ en este ejemplo

$$(\overbrace{\{0,1,2,3\}}^{=C}, \mathcal{B}(C), \mathbb{P}_X)$$

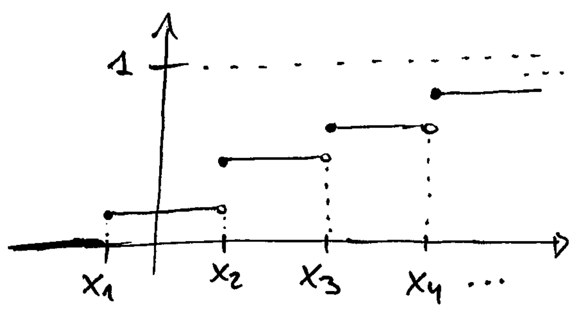
$x \in \mathbb{R}$

$$F_X(x) = \mathbb{P}(X \leq x) = \begin{cases} 0 & \text{si } x < 0 \\ 1/8 & \text{si } 0 \leq x < 1 \\ 1/8 + \frac{3}{8} & \text{si } 1 \leq x < 2 \\ 1/8 + 3/8 + 3/8 & \text{si } 2 \leq x < 3 \\ 1 & \text{si } 3 \leq x \end{cases}$$



DEFINICIÓN: Las VARIABLES ALEATORIAS DISCRETAS son las variables aleatorias que pueden tomar solo una cantidad finita o numerable de valores de $\mathbb{R}$.

- Si X es discreta, la variable aleatoria definida por su función de masa, es decir $p: \{x_1, x_2, \dots\} \longrightarrow [0,1]$, $p(x_i) = \mathbb{P}(X = x_i)$



DEFINICIÓN: La MEDIA (o ESPERANZA) de una variable aleatoria discreta X se define como:

$$\mu = \mathbb{E}(x) = \mathbb{E}X = \sum_{n=1}^{\infty} x_n \cdot \mathbb{P}(X = x_n)$$

(si por ejemplo 90% parajas → 0 hijo
10% parejas → 1 hijo
$\Rightarrow \mu = 0 \cdot 90\% + 1 \cdot 10\% = 0'1$)

DEFINICIÓN: Una MODA de una variable aleatoria discreta X se define como x_i tal que $p(X_i) = \max\{p(x_k) : k \in \mathbb{N}\}$

DEFINICIÓN: La MEDIANA de una variable aleatoria discreta X es el valor M tal que $F_X(M^-) = \lim_{\substack{h \to 0 \\ h > 0}} F_X(M-h) \leq 1/2$ y $F_X(M) \geq 1/2$.

DEFINICIÓN: La VARIAZA de una variable aleatoria discreta X es:

$$\sigma^2 = V(x) = \mathbb{E}\left((X - \mathbb{E}(X))^2\right) = \sum_n (x_n - \mu)^2 \cdot \mathbb{P}(X = x_n)$$

DEFINICIÓN: La DESVIACIÓN TÍPICA es $\sigma = \sqrt{\sigma^2} = \sqrt{V(X)}$.

DEFINICIÓN: El MOMENTO NO CENTRADO de ORDEN $k = 1, 2, 3, \dots$ es $\mathbb{E}(X^k) = \sum_n x_n^k \cdot \mathbb{P}(X = x_n)$

DEFINICIÓN: El MOMENTO CENTRADO es:

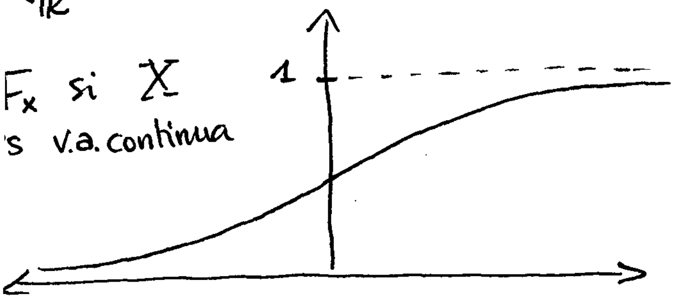
$$\mathbb{E}\left((X - \mathbb{E}(X))^k\right) = \sum_n (x_n - \mu)^k \cdot \mathbb{P}(X = x_n)$$

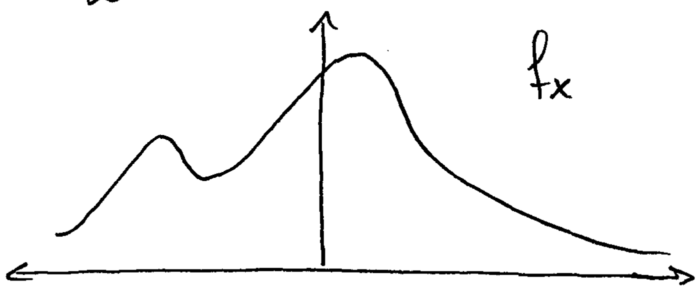
VARIABLES ALEATORIAS CONTINUAS

Son las variables aleatorias $X : \Omega \to \mathbb{R}$ que pueden tomar cualquier valor real, y que existe una función $f_X : \mathbb{R} \to [0, \infty)$ que se llama <u>densidad</u> tal que

$$\int_{\mathbb{R}} f_X(x)\,dx = 1 \quad \text{y} \quad \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f_X(y)\,dy \Rightarrow F_X'(x) = f_X(x)$$

F_X si X es v.a. continua





f_X

<u>Observación</u>:

- $\mathbb{P}(X = x_n)$ si X es v.a. discreta es > 0.
- Si X es v.a. continua $\mathbb{P}(X = x) = 0$.
- $f_X(x) = F_X'(x) = \lim_{h\to 0} \dfrac{F_X(x+h) - F_X(x)}{h} = \lim_{h \to 0} \dfrac{\mathbb{P}(X \leq x+h) - \mathbb{P}(X \leq x)}{h} =$

$$= \lim_{h\to 0} \frac{\mathbb{P}(x < X \leq x+h)}{h} \to \mathbb{P}(X = x)$$

<u>DEFINICIÓN</u>: La MEDIA de una variable aleatoria continua es:

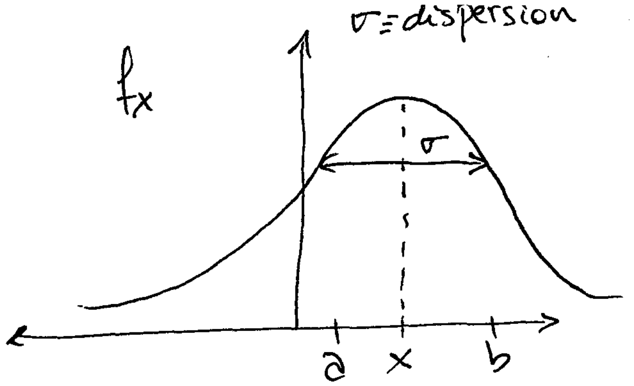
$$\mu = \mathbb{E}(X) = \int_{\mathbb{R}} x \cdot \underbrace{f_X(x)\,dx}_{\approx \mathbb{P}(X = x)}$$

<u>DEFINICIÓN</u>: La VARIANZA de una variable aleatoria continua es:

$$\sigma^2 = \mathbb{E}\left((X - \mathbb{E}(X))^2\right) = \int_{\mathbb{R}} (x-\mu)^2 \cdot f_X(x)\,dx$$

<u>DEFINICIÓN</u>: El MOMENTO NO CENTRADO de ORDEN $K = 1, 2, 3, \ldots$ es:

$$\mathbb{E}(X^K) = \int_{\mathbb{R}} x^K \cdot f_X(x)\,dx$$

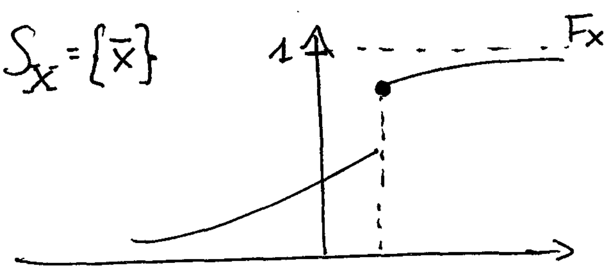


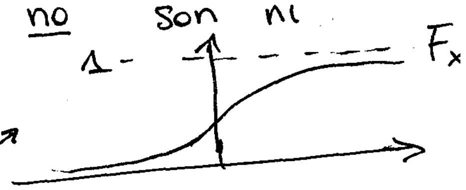
$$F_X(x) = \mathbb{P}(X \le x) = \int_{-\infty}^{x} f_X(y)\,dy$$

$$F_X(b) - F_X(a) = \mathbb{P}(a < X < b) = \int_a^b f_X(y)\,dy$$

<u>Observación</u>: Existen variables aleatorias X que <u>no</u> son ni continuas ni discretas.

$S_X = \{\bar{x}\}$

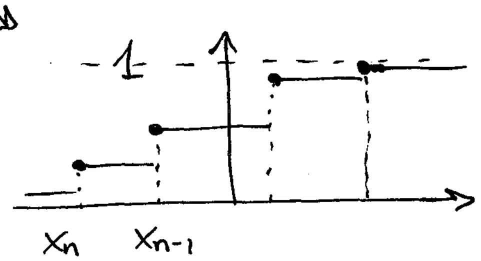




X ni continua ni discreta

como esta

como esta



Todas las variables aleatorias X tienen un conjunto $S_X := \{x \in \mathbb{R} \mid F_X(x) - F_X(x^-) > 0\}$ que es el máximo numerable y $x \in \mathbb{R}$, $F_X(x) = \int_{-\infty}^{x} f_X(y)\,dy + \sum_{\substack{x_i \in S_X \\ x_i \le x}} \left[F_X(x_i) - F_X(x_i^-)\right]$

donde $f_X(y) = F_X'(y)$, $\forall y \in \mathbb{R} \setminus S_X$

$$\boxed{F_X(x^-) = \lim_{\substack{h \to 0 \\ h < 0}} F(x+h)}$$

- Si X es variable continua $\Rightarrow S_X = \emptyset$
- Si X es variable discreta $\Rightarrow S_X = \{x_n\}_{n \in \mathbb{N}}$ y $f_X \equiv 0$.

$$\mathbb{E}(X) = \int_{\mathbb{R} \setminus S_X} x f_X(x)\,dx + \sum_{x_i \in S_X} x_i \cdot \left[F_X(x_i) - F_X(x_i^-)\right]$$

$$\shortparallel$$

$\int_{\mathbb{R}} x \cdot f_X(x)\,dx$ porque S_X es al máximo numerable.

$$\mathbb{V}(X) = \int_{\mathbb{R}\setminus S_X} [x - \mathbb{E}(X)]^2 f_X(x)\,dx + \sum_{x_i \in S_X} (x_i - \mathbb{E}(X))^2 \underbrace{[F_X(x_i) - F_X(x_i^-)]}_{\mathbb{P}(X = x_i)}$$

DESIGUALDAD DE CHEBYCHEV

Sea X una variable aleatoria y $g : \mathbb{R} \longrightarrow [0, \infty]$ no negativa entonces, $\forall \varepsilon > 0$, $\boxed{\mathbb{P}_X(\{x \in \mathbb{R} : g(x) \geq \varepsilon\}) \leq \frac{\mathbb{E}(g(x))}{\varepsilon}}$

Demostración

$$\mathbb{E}(g(x)) = \int_{\mathbb{R}} x \cdot f_{g(x)}(x)\,dx = \int_{\mathbb{R}} g(x) \cdot f_X(x)\,dx =$$

$$= \int_{\{x \in \mathbb{R} : g(x) \geq \varepsilon\}} g(x) \cdot f_X(x)\,dx + \int_{\{x \in \mathbb{R} : g(x) < \varepsilon\}} g(x) \cdot f_X(x)\,dx \geq$$

$g(x) f(x) \geq 0$ $\qquad g(x) \geq \varepsilon$

$$\geq \int_{\{x \in \mathbb{R} : g(x) \geq \varepsilon\}} g(x) \cdot f_X(x)\,dx \geq \varepsilon \int_{\{x \in \mathbb{R} : g(x) \geq \varepsilon\}} f_X(x)\,dx = \varepsilon \cdot \mathbb{P}_X(\{x \in \mathbb{R} : g(x) \geq \varepsilon\})$$

□

COROLARIO: Si $\mathbb{E}(X) = \mu$ y $\mathbb{V}(X) = \sigma^2 \Rightarrow$

$\Rightarrow \forall K > 0$, $\mathbb{P}_X(\{x \in \mathbb{R} : |x - \mu| \geq K\sigma\}) \leq \frac{1}{K^2}$

Demostración

$g(x) := (x - \mu)^2$

$$\mathbb{P}_X(\{x \in \mathbb{R} : |x - \mu| \geq K \cdot \sigma\}) = \mathbb{P}(\{x \in \mathbb{R} : (x - \mu)^2 \geq K^2 a^2\}) \leq$$

$$\leq \frac{\mathbb{E}((x - \mu)^2)}{K^2 \sigma^2} = \frac{\sigma^2}{K^2 \sigma^2} = \frac{1}{K^2}$$ □

Si $k=2$, $\mathbb{P}_X(\{x \in \mathbb{R} : |x-\mu| \geq 2\sigma\}) \leq \frac{1}{4}$