
TEORIA DE LA PROBABILITAT

ApuntsFME

BARCELONA, OCTUBRE 2018

Darrera modificació: 16 d'octubre de 2018.

This work is licensed under a [Creative Commons](#)
“[Attribution-NonCommercial-ShareAlike 4.0 International](#)”
license.



Continguts

1	Espai de probabilitat	1
1.1	Definició axiomàtica de probabilitat	1
	Desigualtats de Bonferroni	2
1.2	Probabilitat condicionada	4
1.3	Independència	5
1.4	Espai producte	5
1.5	Lema de Borel-Cantelli	6
	Lema de Borel-Cantelli	7
2	Variables aleatòries	9
2.1	Definició i propietats bàsiques de les variables aleatòries	9
	Teorema de l'existència d'una funció de distribució	12
2.2	Esperança d'una variable aleatòria. Desigualtats de Markov i Chebyshev	13
	Desigualtat de Markov	15
	Desigualtat de Chebyshev	15
2.3	Vectors de variables aleatòries. Independència de variables aleatòries . . .	16
3	Variables aleatòries discretes	19
3.1	Definició i objectes relacionats	19
3.2	Funció generadora de probabilitat	20
3.3	Models de variables aleatòries discretes	23
	Índex alfabètic	27

Tema 1

Espai de probabilitat

1.1 Definició axiomàtica de probabilitat

Definició 1.1.1. Un espai de probabilitat és un espai de mesura format per la terna (Ω, \mathcal{A}, p) tal que $p(\Omega) = 1$. Diem que

- Ω és l'espai mostral,
- \mathcal{A} és el conjunt d'esdeveniments o de successos,
- p és la funció de probabilitat.

Observació 1.1.2. Recordem que (Ω, \mathcal{A}) és un espai mesurable si $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ és una σ -àlgebra d' Ω , és a dir,

- i) $\emptyset \in \mathcal{A}$,
- ii) $A \in \mathcal{A} \iff \bar{A} \in \mathcal{A}$,
- iii) Si $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{A}$, aleshores $\bigcup_{i \in \mathbb{N}} A_i \in \mathcal{A}$.

I que $(\Omega, \mathcal{A}, \mu)$ és un espai de mesura si μ és una mesura sobre l'espai mesurable (Ω, \mathcal{A}) , és a dir,

- i) $\mu(\emptyset) = 0$,
- ii) $\forall A \in \mathcal{A}, \quad \mu(A) \geq 0$,
- iii) (σ -additivitat) Si $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{A}$ és tal que $\forall i \neq j, A_i \cap A_j = \emptyset$, aleshores

$$\mu \left(\bigcup_{i \in \mathbb{N}} A_i \right) = \sum_{i \in \mathbb{N}} \mu(A_i).$$

Proposició 1.1.3. Sigui (Ω, \mathcal{A}, p) un espai de probabilitat. Aleshores,

- i) Si $A_1, \dots, A_r \in \mathcal{A}$ t. q. $\forall i \neq j, A_i \cap A_j = \emptyset$, aleshores $p \left(\bigcap_{i=1}^r A_i \right) = \sum_{i=1}^r p(A_i)$.
- ii) $A \in \mathcal{A} \implies p(\bar{A}) = 1 - p(A)$.

iii) $A, B \in \mathcal{A}, A \subseteq B \implies p(B \setminus A) = p(B) - p(A)$.

iv) $A, B \in \mathcal{A}, A \subseteq B \implies p(A) \leq p(B)$.

v) Successions monòtones:

a) Si $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{A}$ són tals que $A_i \subseteq A_{i+1}$, aleshores $p\left(\bigcup_{i \in \mathbb{N}} A_i\right) = \lim_{i \rightarrow \infty} p(A_i)$.

b) Si $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{A}$ són tals que $A_i \supseteq A_{i+1}$, aleshores $p\left(\bigcap_{i \in \mathbb{N}} A_i\right) = \lim_{i \rightarrow \infty} p(A_i)$.

Demostració.

1. Conseqüència directa de la σ -additivitat.

2. Conseqüència directa de ii) usant que $\mathcal{A} = A \cup \bar{A}$.

3. Com que $A \subseteq B$, $B = (B \setminus A) \cup A$ i, per tant, $p(B \setminus A) = p(B) - p(A)$.

4. Conseqüència directa de iii) ja que $p(B \setminus A) \geq 0$.

5.

a) Sigui $B_0 = A_0$ i per $i > 0$ sigui $B_i = A_i \setminus A_{i-1}$. Aleshores, es compleix que $\forall i \neq j, B_i \cap B_j = \emptyset$ i que $\bigcup_{i \in \mathbb{N}} B_i = \bigcup_{i \in \mathbb{N}} A_i$, de manera que

$$\begin{aligned} p\left(\bigcup_{i \in \mathbb{N}} A_i\right) &= p\left(\bigcup_{i \in \mathbb{N}} B_i\right) = \sum_{i \in \mathbb{N}} p(B_i) = \\ &= \lim_{N \rightarrow \infty} \sum_{i=0}^N p(B_i) = \lim_{N \rightarrow \infty} p\left(\bigcup_{i=0}^N B_i\right) = \lim_{N \rightarrow \infty} p(A_N). \end{aligned}$$

b) Anàleg al cas anterior.

□

Observem que l'apartat v) només es pot aplicar en casos molt particulars. En general, si tenim A_1, \dots, A_r successos, hi ha estimacions per a $p(\bigcup_{i=1}^r A_i)$ com es veu al teorema següent.

Teorema 1.1.4. *Desigualtats de Bonferroni.*

Siguin $A_1, \dots, A_r \in \mathcal{A}$, i per $I \subseteq \{1, \dots, r\}$ sigui $A_I = \bigcap_{i \in I} A_i$. Definim

$$S_k = \sum_{I \in \{1, \dots, r\}, \#I=k} p(A_I),$$

això és, $S_1 = \sum p(A_i)$, $S_2 = \sum_{i \neq j} p(A_i \cap A_j)$, Aleshores:

i) Si t és parell,

$$p\left(\bigcup_{i=1}^r A_i\right) \geq \sum_{i=1}^t (-1)^{i+1} S_i$$

ii) Si t és senar,

$$p\left(\bigcup_{i=1}^r A_i\right) \leq \sum_{i=1}^t (-1)^{i+1} S_i$$

Observació 1.1.5. Amb els casos $t = 1$ (desigualtat de Boole) i $t = 2$ es poden donar fites inferiors i superiors.

Exemple 1.1.6.

1. Espais de probabilitat numerables.

Prenem un conjunt numerable $\Omega = \{a_i\}_{i \geq 1}$. Prenem $\mathcal{A} = \mathcal{P}(\Omega)$ (que és una σ -àlgebra). Per a definir la probabilitat sobre (Ω, \mathcal{A}) prenem una successió $\{p_i\}_{i \geq 1}$ tal que $0 \leq p_i \leq 1$ que compleix que $\forall i, p(a_i) = p_i$ i $\sum p_i = 1$. Per tant, per a qualsevol element $A \in \mathcal{A}$, tenim que

$$p\left(\bigcup_{a \in A} \{a\}\right) = p(A) = \sum_{i \geq 1} p(\{a_i\}).$$

Si, a més, $|\Omega| < +\infty$, $\mathcal{A} = \mathcal{P}(\Omega)$ té $2^{|\Omega|}$ elements i si prenem $\Omega = \{a_i\}_{i=1}^N$ i $p_1 = p_2 = \dots = p_N = \frac{1}{N}$ obtenim un espai clàssic de probabilitat.

2. Espai de probabilitat en $[a, b] \subseteq \mathbb{R}$.

Signi $\Omega = [a, b]$ i prenem $\mathcal{A} = \mathcal{B} \cap [a, b]$ amb \mathcal{B} un borelià i com a funció de probabilitat $p = \frac{\lambda}{b-a}$, on λ és la mesura de Lebesgue. Observem que no podem prendre tot \mathbb{R} perquè no podem normalitzar $\lambda(\mathbb{R})$. Malgrat això, usant λ construirem més endavant funcions de probabilitat sobre $(\mathbb{R}, \mathcal{B})$.

3. Tirada indefinida d'una moneda.

En aquest cas tenim que $\Omega = \{a_i\}_{i \geq 1}$, $a_i \in \{0, 1\}$ de la forma

$$\begin{array}{l} 00010001110110\dots \\ 01001110101101\dots \\ 10010111110010\dots \end{array}$$

sent 0 creu i 1 cara. Aquest conjunt és no numerable fàcilment demostrable amb l'argument de la diagonal de Cantor. Per a construir una σ -àlgebra sobre Ω trobem una "bijecció" amb $[0, 1]$ de la forma

$$\begin{aligned} \varphi: \Omega &\rightarrow [0, 1] \subseteq \mathbb{R} \\ a = a_1 a_2 \dots &\mapsto 0.a_1 a_2 \dots \end{aligned}$$

on $0.a_1 a_2 \dots$ és un nombre en binari. No és una bijecció completa ja que hi ha elements diferents que van a la mateixa imatge degut als nombres que acaben en 1 periòdic, però al ser tots racionals, el conjunt d'aquests nombres és numerable i per tant té mesura nul·la. És per això que podem definir una σ -àlgebra sobre Ω prenent $\{\varphi^{-1}(A)\}_{A \subseteq \mathcal{B} \cup [0, 1]}$. Similarment ho fem amb la mesura.

1.2 Probabilitat condicionada

Definició 1.2.1. Sigui (Ω, \mathcal{A}, p) un espai de probabilitat i siguin $A, B \in \mathcal{A}$. Definim la probabilitat d' A condicionada a B com

$$p(A | B) = \frac{p(A \cap B)}{p(B)}.$$

Observació 1.2.2. Sigui (Ω, \mathcal{A}, p) un espai de probabilitat i sigui $B \in \mathcal{A}$ t. q. $p(B) > 0$. Aleshores, l'aplicació

$$\begin{aligned} p_B: \mathcal{A} &\rightarrow \mathbb{R} \\ A &\mapsto p_B(A) := p(A | B) \end{aligned}$$

defineix un espai de probabilitat $(\Omega, \mathcal{A}, p_B)$.

Proposició 1.2.3. Sigui I un conjunt numerable o finit i siguin $\{A_i\}_{i \in I} \subseteq \mathcal{A}$ tals que

- i) $p(A_i) > 0$,
- ii) $i \neq j \implies A_i \cap A_j = \emptyset$,
- iii) $\bigcup_{i \in I} A_i = \Omega$.

Aleshores,

- 1) Probabilitat total:

$$p(B) = \sum_{i \in I} p(B | A_i) p(A_i), \quad \forall B \in \mathcal{A}.$$

- 2) Fórmula de Bayes:

$$p(A_i | B) = \frac{p(B | A_i) p(A_i)}{\sum_{j \in I} p(B | A_j) p(A_j)}, \quad \forall B \in \mathcal{A} \text{ amb } p(B) > 0.$$

Demostració.

- 1) Com que els A_i són disjunts i $\bigcup_{i \in I} A_i = \Omega$, $\forall B \in \mathcal{A}$, $B = \bigcup_{i \in I} B \cap A_i$, i la unió és disjunta. Es té

$$p(B) = p\left(\bigcup_{i \in I} B \cap A_i\right) \stackrel{\sigma\text{-add.}}{=} \sum_{i \in I} p(B \cap A_i) = \sum_{i \in I} p(B | A_i) p(A_i).$$

- 2)

$$\begin{aligned} p(A_i | B) \sum_{j \in I} p(B | A_j) p(A_j) &\stackrel{i)}{=} p(A_i | B) p(B) = \\ \frac{p(B \cap A_i)}{p(B)} p(B) &= p(B \cap A_i) = p(B | A_i) p(A_i). \end{aligned}$$

□

Problema 1.2.4. *Ruïna del jugador.* Partim d'un capital de k unitats i, en cada jugada (sense memòria) augmenta o disminueix el capital en una unitat, amb probabilitats $1/2$ i $1/2$. El joc acaba si ens quedem sense capital o si assolim un objectiu N ($N > k$). Quina és la probabilitat de perdre tot el capital?

Solució. Sigui A_k el succés “el jugador, començant amb capital k , perd”. Condicionem A_k a la primera tirada de la moneda i definim B com el succés “la primera tirada surt cara”. Aleshores,

$$\begin{aligned} p(A_k) &= p(A_k|B)p(B) + p(A_k|\bar{B})p(\bar{B}) = \frac{p(A_k|B)}{2} + \frac{p(A_k|\bar{B})}{2} \\ \implies 2p(A_k) &= p(A_{k-1}) + p(A_{k+1}) \implies p(A_k) - p(A_{k-1}) = p(A_{k+1}) - p(A_k) = C, \end{aligned}$$

el que ens diu que la diferència entre nivells és constant. Per tant $p(A_k) = p(A_0) + kC$. Sabent que $p(A_0) = 1$ i $p(A_N) = 0$ ens queda que

$$0 = 1 + CN \implies C = -\frac{1}{n} \implies p(A_k) = 1 - \frac{k}{N}.$$

1.3 Independència

Definició 1.3.1. Sigui (Ω, \mathcal{A}, p) un espai de probabilitat, sigui I un conjunt finit o numerable i sigui $\{A_i\}_{i \in I} \subseteq \mathcal{A}$. Diem que els esdeveniments A_i són independents si $\forall J \subseteq I$ amb $|J| \in \mathbb{N}$ es té que

$$p\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} p(A_j).$$

Exemple 1.3.2.

1. \emptyset, Ω són independents entre si.
2. A és independent amb si mateix si i només si $p(A) = 1$ o $p(A) = 0$.
3. No tenim independència si només les interseccions dos a dos compleixen que la probabilitat de la intersecció és el producte de probabilitats.
4. A i B independents $\iff A$ i \bar{B} independents, ja que

$$\begin{aligned} p(A \cap \bar{B}) &= p(A) - p(A \cap B) = p(A) - p(A)(B) \\ &= p(A) \left(p(B) + p(\bar{B}) - p(B) \right) = p(A)p(\bar{B}) \end{aligned}$$

1.4 Espai producte

Donats dos espais de probabilitat $(\Omega_1, \mathcal{A}_1, p_1)$ i $(\Omega_2, \mathcal{A}_2, p_2)$, volem construir un nou espai de probabilitat $(\Omega_3, \mathcal{A}_3, p_3)$ que codifiqui els dos espais de probabilitat inicials. A aquest espai de probabilitat l'anomenarem espai de probabilitat producte.

Definició 1.4.1. Siguin $(\Omega_1, \mathcal{A}_1, p_1)$ i $(\Omega_2, \mathcal{A}_2, p_2)$ dos espais de probabilitat. Anomenem espai de probabilitat producte a la terna $(\Omega_3, \mathcal{A}_3, p_3)$ tal que

- i) $\Omega_3 = \Omega_1 \times \Omega_2$
- ii) $\mathcal{A}_3 = \sigma(\mathcal{A}_1 \times \mathcal{A}_2)$ (σ -àlgebra generada per $\mathcal{A}_1 \times \mathcal{A}_2$)
- iii) p_3 és una funció de probabilitat que compleix que $\forall A_1, A_2$ t. q. $A_1 \times A_2 \in \mathcal{A}_1 \times \mathcal{A}_2$ aleshores $p_3(A_1 \times A_2) = p_1(A_1)p_2(A_2)$.

Observació 1.4.2. p_3 està ben definida ja que pel Teorema d'extensió de Carathéodory podem construir una σ -àlgebra sobre $\Omega_1 \times \Omega_2$ a partir d'una extensió de $\sigma(\mathcal{A}_1 \times \mathcal{A}_2)$ i restringir p_3 segons [iii](#)).

Observació 1.4.3. Podem estendre λ (la mesura de Lebesgue) a \mathbb{R}^2 de la següent forma. Sabem que $([0, 1], \mathcal{B} \cap [0, 1], \lambda_{[0, 1]})$ és un espai de probabilitat. Aleshores

$$([0, 1] \times [0, 1], \sigma(\mathcal{B} \cap [0, 1] \times \mathcal{B} \cap [0, 1]), \lambda_{[0, 1] \times [0, 1]})$$

defineix un espai de probabilitat a \mathbb{R}^2 .

Problema 1.4.4. *Agulla de Buffon.* Considerem el pla \mathbb{R}^2 tesel·lat amb línies paral·leles indefinides separades per una distància L . Llancem una agulla de longitud $l \leq L$ sobre el pla. Trobar quina és la probabilitat que l'agulla toqui una de les línies.

Solució. Considerarem dues variables: x com la distància del centre de l'agulla a la línia més propera i θ com l'angle de l'agulla amb la direcció de les línies. Tenim que $x \in [0, \frac{L}{2}]$ i $\theta \in [0, \pi)$ i per tant, $\Omega = [0, \frac{L}{2}] \times [0, \pi)$, \mathcal{A} són els borelians del conjunt i p la mesura de Lebesgue normalitzada en \mathcal{A} . Sigui $A \in \mathcal{A}$ l'esdeveniment "l'agulla talla una recta" i $\omega \in \Omega$ una tirada. Aleshores $\omega \in A \iff x \leq \frac{l}{2} \sin \theta$. Per tant,

$$p(A) = \frac{\int_0^\pi \frac{l}{2} \sin \theta d\theta}{\frac{L\pi}{2}} = \frac{2l}{L\pi}.$$

1.5 Lema de Borel-Cantelli

Siguin (Ω, \mathcal{A}, p) un espai de probabilitat i $\{A_n\}_{n \geq 1} \subseteq \mathcal{A}$. Volem donar-li un sentit a "límit de $\{A_n\}_{n \geq 1}$ ". Farem com a \mathbb{R} i definirem els límits superior i inferior (que sempre existiran) i, si coincideixen, aquest serà el límit.

Definició 1.5.1. Sigui (Ω, \mathcal{A}, p) un espai de probabilitat. Donats $\{A_n\}_{n \geq 1} \subseteq \mathcal{A}$, definim els límits superior i inferior de la successió de successos $\{A_n\}_{n \geq 1}$ com

$$\limsup_{n \rightarrow \infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k,$$

$$\liminf_{n \rightarrow \infty} A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k.$$

Observació 1.5.2. Els dos límits pertanyen a \mathcal{A} ja que són unió i intersecció numerable de successos.

Proposició 1.5.3. Sigui (Ω, \mathcal{A}, p) un espai de probabilitat i siguin $\{A_n\}_{n \geq 1} \subseteq \mathcal{A}$. Aleshores,

- i) $\liminf_{n \rightarrow \infty} A_n = \{\omega \in \Omega : \exists m \equiv m(\omega) \text{ amb } \omega \in A_r \ \forall r \geq m(\omega)\},$
- ii) $\limsup_{n \rightarrow \infty} A_n = \{\omega \in \Omega : \omega \text{ pertany a un nombre infinit dels } A_n\},$
- iii) $\liminf_{n \rightarrow \infty} A_n \subseteq \limsup_{n \rightarrow \infty} A_n.$

Demostració.

- i) $\omega \in \liminf A_n \iff \omega \in \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k \iff \exists m \equiv m(\omega) \text{ t. q. } \omega \in \bigcap_{k=m(\omega)}^{\infty} A_k,$ és a dir, si $\exists m \equiv m(\omega) \text{ t. q. } \omega \in A_r \ \forall r \geq m(\omega).$
- ii) $\omega \in \limsup A_n \iff \omega \in \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k \iff \omega \in \bigcup_{k=n}^{\infty} A_k \ \forall n \iff \forall n, \exists n_0 \geq n \text{ t. q. } \omega \in A_{n_0} \iff \omega \text{ pertany a un nombre infinit dels } A_n.$
- iii) Si $\omega \in \liminf A_n$, aleshores $\omega \in A_r, \forall r \geq m(\omega)$, de manera que pertany a un nombre infinit dels A_n i, en conseqüència, pertany a $\limsup A_n$.

□

Proposició 1.5.4. Sigui (Ω, \mathcal{A}, p) un espai de probabilitat i siguin $\{A_n\}_{n \geq 1} \subseteq \mathcal{A}$, amb $\lim A_n = A$. Aleshores, $p(A) = p(\lim A_n) = \lim p(A_n)$ i aquest límit existeix.

Demostració. Definim $B_n = \bigcup_{k \geq n} A_k$ i $C_n = \bigcap_{k \geq n} A_k$. Observem que $\{B_n\}_{n \geq 1}$ és decreixent i que $\{C_n\}_{n \geq 1}$ és creixent. Naturalment tenim que, $C_n \subseteq A_n \subseteq B_n$, $\limsup A_n = \bigcap_{n \geq 1} B_n$ i $\liminf A_n = \bigcup_{n \geq 1} C_n$.

Vegem que $p(\liminf A_n) \leq \liminf p(A_n)$.

$$p(\liminf A_n) = p\left(\bigcup_{n \geq 1} C_n\right) = \lim p(C_n) = \lim p\left(\bigcap_{k \geq n} A_k\right) \leq \liminf p(A_n).$$

Al darrer pas hem utilitzat el fet que $p\left(\bigcap_{k \geq n} A_k\right) \leq p(A_n)$. Anàlogament, tenim que $\limsup p(A_n) \leq p(\limsup A_n)$. Així doncs,

$$p(\liminf A_n) \leq \liminf p(A_n) \leq \limsup p(A_n) \leq p(\limsup A_n).$$

Atès que $p(\liminf A_n) = p(\limsup A_n) = p(A)$, concloem que

$$\liminf p(A_n) = \limsup p(A_n) = \lim p(A_n) = p(A).$$

□

Teorema 1.5.5. *Lema de Borel-Cantelli.*

Sigui (Ω, \mathcal{A}, p) un espai de probabilitat i siguin $\{A_n\}_{n \geq 1} \subseteq \mathcal{A}$. Aleshores,

- i) $\sum_{n \geq 1} p(A_n) < \infty \implies p(\limsup A_n) = 0.$
- ii) Si $\{A_n\}_{n \geq 1}$ és independent, $\sum_{n \geq 1} p(A_n) = \infty \implies p(\limsup A_n) = 1.$

Demostració. Anomenem $A = \limsup A_n = \bigcap_{n \geq 1} \bigcup_{k \geq n} A_k$.

i) Sabem que

$$0 \leq p(A) \leq p\left(\bigcup_{k \geq n} A_k\right) \leq \sum_{k \geq n} p(A_k), \forall n \in \mathbb{N}$$

i que $\sum_{n \geq 1} p(A_n) < \infty$, de manera que $\lim_{n \rightarrow \infty} \sum_{k \geq n} p(A_k) = 0$ i immediatament deduïm que $p(A) = 0$.

ii) Observem primer que $\bar{A} = \overline{\bigcap_{n \geq 1} \bigcup_{k \geq n} A_k} = \bigcup_{n \geq 1} \bigcap_{k \geq n} \bar{A}_k = \liminf \bar{A}_n$. Veurem que

$p(\bar{A}) = 0$. Calculem $p\left(\bigcap_{m \geq n} \bar{A}_m\right)$.

$$\begin{aligned} 0 \leq p\left(\bigcap_{m \geq n} \bar{A}_m\right) &= \lim_{r \rightarrow \infty} p\left(\bigcap_{m=n}^r \bar{A}_m\right) = \lim_{r \rightarrow \infty} \prod_{m=n}^r \left(p(\bar{A}_m)\right) = \\ &= \lim_{r \rightarrow \infty} \prod_{m=n}^r (1 - p(A_m)) \leq \lim_{r \rightarrow \infty} \prod_{m=n}^r \left(e^{-p(A_m)}\right) = \\ &= \lim_{r \rightarrow \infty} e^{-\sum_{m=n}^r p(A_m)} = 0, \end{aligned}$$

de manera que $p\left(\bigcap_{m \geq n} \bar{A}_m\right) = 0, \forall n \in \mathbb{N}$. Finalment,

$$0 \leq p(\bar{A}) = p\left(\bigcup_{n \geq 1} \bigcap_{m \geq n} \bar{A}_m\right) \leq \sum_{n \geq 1} p\left(\bigcap_{m \geq n} \bar{A}_m\right) = 0 + 0 + \dots = 0,$$

i concloem que $p(A) = 1$.

□

Tema 2

Variables aleatòries

2.1 Definició i propietats bàsiques de les variables aleatòries

Definició 2.1.1. Siguin $(\Omega_1, \mathcal{A}_1)$ i $(\Omega_2, \mathcal{A}_2)$ espais mesurables. Diem que $X: \Omega_1 \rightarrow \Omega_2$ és una variable aleatòria si

$$X^{-1}(A_2) \in \mathcal{A}_1, \forall A_2 \in \mathcal{A}_2.$$

En aquest curs, sempre prendrem $(\Omega_2, \mathcal{A}_2) = (\mathbb{R}, \mathcal{B})$. Per tant, quan parlem de variable aleatòria ens estarem referint a una aplicació $X: \Omega \rightarrow \mathbb{R}$ amb $B \in \mathcal{B} \implies X^{-1}(B) \in \mathcal{A}$, on (Ω, \mathcal{A}) és un espai de mesura.

Exemple 2.1.2.

1. Sigui (Ω, \mathcal{A}) un espai de mesura. Aleshores, $\forall c \in \mathbb{R}$, l'aplicació

$$\begin{aligned} X: \Omega &\rightarrow \mathbb{R} \\ \omega &\mapsto c \end{aligned}$$

és una variable aleatòria, atès que, $\forall B \in \mathcal{B}$, es té que

$$X^{-1}(B) = \begin{cases} \Omega, & \text{si } c \in B, \\ \emptyset, & \text{si } c \notin B. \end{cases}$$

2. Siguin X i Y variables aleatòries. Aleshores, també son variables aleatòries les següents funcions.

- $X + Y$
- $X - Y$
- $aX, \forall a \in \mathbb{R}$
- XY
- $|X|$
- $\max\{X, Y\}$
- $\min\{X, Y\}$

- X^+
 - X^-
 - $g(X, Y)$, on $g: \mathbb{R}^2 \rightarrow \mathbb{R}$ és una funció mesurable.
3. Sigui (Ω, \mathcal{A}) un espai de mesura i sigui $A \in \mathcal{A}$. Definim la variable aleatòria indicadora d' A com

$$\mathbb{I}_A \equiv \mathbb{1}_A: \Omega \rightarrow \mathbb{R}$$

$$\omega \mapsto \mathbb{I}_A(\omega) = \begin{cases} 0, & \text{si } \omega \notin A, \\ 1, & \text{si } \omega \in A. \end{cases}$$

Vegem que, efectivament, es tracta d'una variable aleatòria. Sigui $B \in \mathcal{B}$. Aleshores,

$$\mathbb{I}_A^{-1}(B) = \begin{cases} \Omega, & \text{si } 0 \in B, 1 \in B, \\ \overline{A}, & \text{si } 0 \in B, 1 \notin B, \\ A, & \text{si } 0 \notin B, 1 \in B, \\ \emptyset, & \text{si } 0 \notin B, 1 \notin B. \end{cases}$$

Observació 2.1.3. A partir d'ara, emprarem la notació següent. Sigui (Ω, \mathcal{A}, p) un espai de probabilitat i sigui $B \in \mathcal{B}$, escrivim

$$p(X \in B) := p\left(\{\omega \in \Omega \mid \omega \in X^{-1}(B)\}\right).$$

Exemple 2.1.4. $p(X \leq 2) = p\left(\{\omega \in \Omega \mid X(\omega) \leq 2\}\right)$.

Observació 2.1.5. Sigui (Ω, \mathcal{A}, p) un espai de probabilitat i sigui X una variable aleatòria. X induïx una funció de probabilitat P_X sobre l'espai de mesura $(\mathbb{R}, \mathcal{B})$

$$P_X(B) := p(X \in B).$$

És a dir, $(\mathbb{R}, \mathcal{B}, P_X)$ és un espai de probabilitat. Comprovem, primer, que és un espai de mesura.

- i) $P_X(\emptyset) = p\left(\{\omega \in \Omega \mid \omega \in X^{-1}(\emptyset)\}\right) = p(\emptyset) = 0$, atès que p és una funció de probabilitat.
- ii) $0 \leq p\left(\{\omega \in \Omega \mid \omega \in X^{-1}(B)\}\right) = P_X(B)$, atès que p és una funció de probabilitat.
- iii) Si $\{B_i\}_{i \in \mathbb{N}} \subseteq \mathcal{B}$ són disjunts dos a dos, aleshores $\{X^{-1}(B_i)\}_{i \in \mathbb{N}} \subseteq \mathcal{A}$ també són disjunts dos a dos. I, per ser p una funció de probabilitat, es té que

$$\begin{aligned} P_X\left(\bigcup_{i \in \mathbb{N}} B_i\right) &= p\left(\left\{\omega \in \Omega \mid \omega \in X^{-1}\left(\bigcup_{i \in \mathbb{N}} B_i\right)\right\}\right) = \\ &= \sum_{i \in \mathbb{N}} p\left(\{\omega \in \Omega \mid \omega \in X^{-1}(B_i)\}\right) = \\ &= \sum_{i \in \mathbb{N}} P_X(B_i). \end{aligned}$$

A més a més, per ser p una funció de probabilitat,

$$P_X(\mathbb{R}) = p\left(\{\omega \in \Omega \mid \omega \in X^{-1}(\mathbb{R})\}\right) = p(\Omega) = 1$$

i $(\mathbb{R}, \mathcal{B}, P_x)$ és un espai de probabilitat.

Observació 2.1.6. Sigui (Ω, \mathcal{A}) un espai mesurable. Recordem que $X: \Omega \rightarrow \mathbb{R}$ és una funció mesurable si i només si $X^{-1}((-\infty, a]) \in \mathcal{A}$, $\forall a \in \mathbb{R}$.

Definició 2.1.7. Sigui (Ω, \mathcal{A}, p) un espai de probabilitat i sigui X una variable aleatòria. Anomenem funció de distribució de probabilitat d' X a l'aplicació

$$F_X: \mathbb{R} \rightarrow [0, 1] \\ x \mapsto F_X(x) = p(X \leq x) = P_X((-\infty, x]).$$

Proposició 2.1.8. Sigui (Ω, \mathcal{A}, p) un espai de probabilitat i sigui F_X la funció de distribució de probabilitat d'una variable aleatòria X sobre (Ω, \mathcal{A}, p) . Aleshores,

- i) $x_1 \leq x_2 \implies F_X(x_1) \leq F_X(x_2)$.
- ii) $\lim_{x \rightarrow -\infty} F_X(x) = 0$ i $\lim_{x \rightarrow \infty} F_X(x) = 1$.
- iii) F_X és contínua per la dreta, és a dir, $\lim_{h \rightarrow 0^+} F_X(x+h) = F_X(x)$.

Demostració.

- i) $F_X(x_1) = p\left(\{\omega \in \Omega \mid X(\omega) \leq x_1\}\right) \leq p\left(\{\omega \in \Omega \mid X(\omega) \leq x_2\}\right) = F_X(x_2)$, atès que $\{\omega \in \Omega \mid X(\omega) \leq x_1\} \subseteq \{\omega \in \Omega \mid X(\omega) \leq x_2\}$ i que p és una funció mesurable.
- ii) Vegem que $\forall \{x_n\}_{n \in \mathbb{N}}$ tal que $\lim_{n \rightarrow \infty} x_n = -\infty$, es té que $\lim_{n \rightarrow \infty} F_X(x_n) = 0$. Definim $A_n = \{\omega \in \Omega \mid X(\omega) \leq x_n\}$. Tenim que $\emptyset \subseteq \liminf A_n \subseteq \limsup A_n$. A més, $\limsup A_n = \emptyset$ perquè, altrament, hi hauria un nombre infinit de conjunts A_n contenint un $\omega \in \Omega$ determinat. Per tant,

$$\lim_{n \rightarrow \infty} F_X(x_n) = \lim_{n \rightarrow \infty} p(A_n) = p\left(\lim_{n \rightarrow \infty} A_n\right) = p(\emptyset) = 0.$$

Anàlogament, es demostra que $\lim_{x \rightarrow \infty} F_X(x) = 1$.

- iii) Fixat x , volem veure que $\lim_{h \rightarrow 0^+} F_X(x+h) = F_X(x)$.

Prenem $C_n = \{\omega \in \Omega \mid X(\omega) \leq x + h_n\}$, on $\{h_n\}$ és una successió de reals no negatius amb límit zero. Aleshores, $\liminf C_n = \limsup C_n = \{\omega \in \Omega \mid X(\omega) \leq x\}$. Això ens diu que

$$\lim_{n \rightarrow \infty} F_X(x + h_n) = \lim_{n \rightarrow \infty} p(C_n) = p\left(\lim_{n \rightarrow \infty} C_n\right) = p(C) = F_X(x).$$

Com això és cert $\forall h$ t. q. $\{h_n\} \rightarrow 0$, tenim que $\lim_{h \rightarrow 0^+} F_X(x+h) = F_X(x)$.

□

Observació 2.1.9. En general no podem assegurar que sigui contínua per l'esquerra. Fent la mateixa prova prenent $x - h_n$ amb $h_n \rightarrow 0^+$ en comptes de $x + h_n$, obtenim que $C = \{\omega \in \Omega \mid X(\omega) < x\}$ i, per tant

$$\lim_{h \rightarrow 0^-} F_X(x + h) = p(X < x) = F_X(x) - p(X = x).$$

Lema 2.1.10. Sigui $f: \mathbb{R} \rightarrow \mathbb{R}$ una funció creixent i fitada. Aleshores f és mesurable Lebesgue.

Demostració. Suposem que f té un nombre no numerable de discontinuïtats. Observem que totes les discontinuïtats són de salt. Sigui $D \subseteq \mathbb{R}$ el conjunt de punts on f és discontinua. Aleshores, tenim que, per tots els punts $x_d \in D$, existeixen els límits $\lim_{x \rightarrow x_d^+} f(x)$ i $\lim_{x \rightarrow x_d^-} f(x)$. Definim, per tot $n \in \mathbb{N}$, els conjunts

$$A_n = \left\{ x_d \in D \mid \frac{1}{n+1} \leq \lim_{x \rightarrow x_d^+} f(x) - \lim_{x \rightarrow x_d^-} f(x) < \frac{1}{n} \right\},$$

on cometem l'abús de notació $\frac{1}{0} = \infty$. Com que D és no numerable, hi ha un nombre numerable de conjunts A_n i $\cup_{n \in \mathbb{N}} A_n = D$, necessàriament $\exists n \in \mathbb{N}$ tal que $|A_n| \notin \mathbb{N}$. Per tant, hi ha un nombre infinit de salts de, com a mínim $\frac{1}{n+1}$, la qual cosa contradiu la hipòtesi que f és fitada. Per tant, f té un nombre numerable de discontinuïtats i és, doncs, mesurable. \square

Teorema 2.1.11. *Teorema de l'existència d'una funció de distribució.*

Sigui $F: \mathbb{R} \rightarrow [0, 1]$ una funció de probabilitat tal que

- i) $x_1 \leq x_2 \implies F_X(x_1) \leq F(x_2)$.
- ii) $\lim_{x \rightarrow -\infty} F(x) = 0$ i $\lim_{x \rightarrow \infty} F(x) = 1$.
- iii) F és contínua per la dreta, és a dir, $\lim_{h \rightarrow 0^+} F(x + h) = F(x)$.

Aleshores, existeixen un espai de probabilitat (Ω, \mathcal{A}, p) i una variable aleatòria $X: \Omega \rightarrow \mathbb{R}$ tals que $F_X(x) = F(x)$.

Demostració. Prenem $(\Omega, \mathcal{A}, p) = ([0, 1], \mathcal{B} \cap [0, 1], \lambda_{[0, 1]})$ i definim

$$X: [0, 1] \rightarrow [0, 1] \\ \omega \mapsto X(\omega) = \sup \{y \in \mathbb{R} \mid F(y) \leq \omega\}.$$

Observem que a tots els punts on F és contínua X també ho és, de manera que X és una funció mesurable. Vegem que $F_X(x) = F(x)$, $\forall x \in \mathbb{R}$. Donat $x \in \mathbb{R}$, definim els conjunts

$$A = \{\omega \in [0, 1] \mid X(\omega) \leq x\}, \\ B = \{\omega \in [0, 1] \mid \omega \leq F(x)\}$$

i observem que

$$P(A) = P(X \leq x) = F_X(x), \\ P(B) = \lambda([0, F(x)]) = F(x).$$

Si demostrem que $A = B$, haurem acabat.

- $\omega \in B \implies \omega \leq F(x) \implies x \notin \{y \in \mathbb{R} \mid F(y) < \omega\} \implies x \geq X(\omega) \implies \omega \in A.$
- $\omega \notin B \implies \omega > F(x) \implies \exists \varepsilon > 0 \text{ t.q. } \omega > F(x + \varepsilon) \implies x(\omega) \geq x + \varepsilon > x \implies X(\omega) > x \implies \omega \notin A.$

□

2.2 Esperança d'una variable aleatòria. Desigualtats de Markov i Chebyshev

Definició 2.2.1. Sigui (Ω, \mathcal{A}, p) un espai de probabilitat i sigui $X: \Omega \rightarrow \mathbb{R}$ una variable aleatòria. Com ja sabem, X indueix una probabilitat P_X sobre $(\mathbb{R}, \mathcal{B})$. Definim l'esperança de la variable aleatòria d' X , $\mathbb{E}[X]$ com

$$\mathbb{E}[X] = \int_{\Omega} X \, dp = \int_{\mathbb{R}} x \, dP_X,$$

si existeix aquesta integral.

Observació 2.2.2. La demostració que aquestes dues integrals són iguals resulta de l'aplicació de la definició de la integral de Lebesgue, però escapa dels objectius d'aquest curs i no l'escriurem.

Observació 2.2.3. Igual que es va veure al curs de teoria de la mesura, pot ser que $\mathbb{E}[X]$ no existeixi o que sigui infinita. No obstant això, atès que $|\int_{\Omega} f \, dp| \leq \int_{\Omega} |f| \, dp$, sovint demanarem que $\mathbb{E}[|X|] \leq +\infty$ per poder afirmar que $\mathbb{E}[X] \leq +\infty$.

Exemple 2.2.4. Sigui (Ω, \mathcal{A}, p) un espai de probabilitat i sigui $A \in \mathcal{A}$. Considerem la variable aleatòria

$$X(\omega) = \mathbb{I}_A(\omega) = \begin{cases} 0, & \text{si } \omega \in A, \\ 1, & \text{si } \omega \notin A. \end{cases}$$

Observem que, donat $B \in \mathcal{B}$,

$$P_{\mathbb{I}_A}(B) = \begin{cases} 1, & \text{si } 0, 1 \in B, \\ p(A), & \text{si } 0 \notin B, 1 \in B, \\ p(\overline{A}), & \text{si } 0 \in B, 1 \notin B, \\ 0, & \text{si } 0, 1 \notin B. \end{cases}$$

Així, podem calcular l'esperança amb les dues integrals i comprovar que el resultat és el mateix.

$$\mathbb{E}[\mathbb{I}_A] = \int_{\Omega} \mathbb{I}_A \, dp = 1 \cdot p(A) + 0 \cdot p(\overline{A}) = p(A),$$

$$\mathbb{E}[\mathbb{I}_A] = \int_{\mathbb{R}} x \, dP_{\mathbb{I}_A} = 1 \cdot P_{\mathbb{I}_A}(\{1\}) + 0 \cdot P_{\mathbb{I}_A}(\{0\}) + \int_{\mathbb{R} \setminus \{0,1\}} x \, dP_{\mathbb{I}_A} = P_{\mathbb{I}_A}(\{1\}) = p(A).$$

Proposició 2.2.5. Sigui $f: \mathbb{R} \rightarrow \mathbb{R}$ una funció mesurable i sigui X una variable aleatòria. Aleshores, $f(X)$ és una variable aleatòria i

$$\mathbb{E}[f(X)] = \int_{\Omega} f(X) \, dp = \int_{\mathbb{R}} x \, dP_{f(X)}.$$

Demostració. Si f és mesurable, aleshores $f(X)$ també, de manera que $f(X)$ és una variable aleatòria i la resta segueix de la definició. \square

Definició 2.2.6. Sigui (Ω, \mathcal{A}, p) un espai de probabilitat i sigui $X: \Omega \rightarrow \mathbb{R}$ una variable aleatòria. Aleshores, definim

- Moment d'ordre r d' X :

$$\mathbb{E}[X^r],$$

on $r \in \mathbb{R}$ i hem suposat que $\mathbb{E}[|X|^r] < +\infty$.

- Moment factorial d'ordre r d' X :

$$\mathbb{E}[(X)_r] = X(X-1)\cdots(X-r+1),$$

on $r \in \mathbb{N}$.

- Variància d' X :

$$\mathbb{V}\text{ar}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2].$$

- Desviació típica d' X :

$$\sigma = \sqrt{\mathbb{V}\text{ar}[X]}.$$

Proposició 2.2.7. Siguin X, Y variables aleatòries, siguin $a, b \in \mathbb{R}$ i sigui $A \in \mathcal{A}$. Es tenen les següents propietats de l'esperança.

- $\mathbb{E}[a] = a$,
- $\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$,
- $\mathbb{E}[\mathbb{I}_A] = p(A)$,
- $|\mathbb{E}[X]| \leq \mathbb{E}[|X|]$.

Es tenen les següents propietats de la variància.

- $\mathbb{V}\text{ar}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2 + \mathbb{E}[X]^2 - 2X\mathbb{E}[X]] = \mathbb{E}[X^2] + \mathbb{E}[X]^2 - 2\mathbb{E}[X]^2 = \mathbb{E}[X^2] - \mathbb{E}[X]^2$,
- $\mathbb{V}\text{ar}[a] = 0$,
- $\mathbb{V}\text{ar}[a + X] = \mathbb{V}\text{ar}[X]$,
- $\mathbb{V}\text{ar}[aX] = a^2 \mathbb{V}\text{ar}[X]$.

Proposició 2.2.8. *Desigualtat de Holder.* Siguin X, Y variables aleatòries i siguin $p, q \in \mathbb{R}$ tals que $\frac{1}{p} + \frac{1}{q} = 1$. Si $\mathbb{E}[|X|^p], \mathbb{E}[|Y|^q] < +\infty$, aleshores

$$\mathbb{E}[|XY|] \leq \mathbb{E}[|X|^p]^{\frac{1}{p}} \mathbb{E}[|Y|^q]^{\frac{1}{q}} < +\infty.$$

Desigualtat de Cauchy-Schwarz. Siguin X, Y variables aleatòries. Si $\mathbb{E}[|X|^2], \mathbb{E}[|Y|^2] < +\infty$, aleshores

$$\mathbb{E}[|XY|] \leq \mathbb{E}[|X|^2]^{\frac{1}{2}} \mathbb{E}[|Y|^2]^{\frac{1}{2}} < +\infty.$$

Desigualtat de Minkowsky. Siguin X, Y variables aleatòries i sigui $p \in \mathbb{R}$. Si $\mathbb{E}[|X|^p], \mathbb{E}[|Y|^p] < +\infty$, aleshores

$$\mathbb{E}[|X + Y|^p]^{\frac{1}{p}} \leq \mathbb{E}[|X|^p]^{\frac{1}{p}} + \mathbb{E}[|Y|^p]^{\frac{1}{p}} < +\infty.$$

Demostració. Tots aquests resultats són l'aplicació de les desigualtats corresponents demostrades al curs de teoria de la mesura. \square

Observació 2.2.9. La desigualtat de Cauchy-Schwarz és el cas particular $p = q = 2$ de la desigualtat de Holder.

Teorema 2.2.10. *Desigualtat de Markov.*

Sigui (Ω, \mathcal{A}, p) un espai de probabilitat, sigui $X: \Omega \rightarrow \mathbb{R}$ una variable aleatòria amb $X > 0$ i sigui $a \in \mathbb{R}^+$. Aleshores,

$$P(X \geq a) \leq \frac{\mathbb{E}[X]}{a}.$$

Demostració. Sigui $A = \{\omega \in \Omega \mid X(\omega) \geq a\}$. Com que X és mesurable, A és un succés. Observem que

$$a\mathbb{I}_A(\omega) \leq X(\omega), \forall \omega \in \Omega.$$

Aleshores,

$$ap(X \geq a) = \mathbb{E}[a\mathbb{I}_A(\omega)] \leq \mathbb{E}[X(\omega)] = \mathbb{E}[X].$$

\square

Teorema 2.2.11. *Desigualtat de Chebyshev.*

Sigui X una variable aleatòria amb $\mathbb{E}[X] < +\infty$, $\text{Var}[X]$ i $\text{Var}[X] \neq 0$. Aleshores, per tot $k > 0$, es té que

$$P(|X - \mathbb{E}[X]| \geq k \sqrt{\text{Var}[X]}) \leq \frac{1}{k^2}.$$

Demostració. Posem $Y := |X - \mathbb{E}[X]|$. Observem que Y és una variable aleatòria. Tenim que, per tot $a > 0$,

$$\begin{aligned} P(Y \geq a) &= P(|X - \mathbb{E}[X]| \geq a) = \\ &= P((X - \mathbb{E}[X])^2 \geq a^2) \leq \\ &\leq \frac{\mathbb{E}[(X - \mathbb{E}[X])^2]}{a^2} = \\ &= \frac{\text{Var}[X]}{a^2}, \end{aligned}$$

on hem aplicat la desigualtat de Markov. Prenent ara $a = k \sqrt{\text{Var}[X]}$, deduïm el resultat volgut. \square

Observació 2.2.12. Quan $\text{Var}[X] = 0$, seguim tenint que, per tot $a > 0$,

$$0 \leq P(|X - \mathbb{E}[X]| \geq a) \leq \frac{\text{Var}[X]}{a^2} = 0,$$

i tenim que $P(X = \mathbb{E}[X]) = 1$.

2.3 Vectors de variables aleatòries. Independència de variables aleatòries

Definició 2.3.1. Sigui (Ω, \mathcal{A}, p) un espai de probabilitat. Diem que un vector de variables aleatòries o una variable aleatòria multidimensional de dimensió n és una funció mesurable $\vec{X}: \Omega \rightarrow \mathbb{R}^n$.

Observació 2.3.2. La funció

$$\begin{aligned} \pi_i: \mathbb{R}^n &\rightarrow \mathbb{R} \\ (x_1, \dots, x_n) &\mapsto x_i \end{aligned}$$

és una funció mesurable. Per tant, les components $\pi_i \circ \vec{X} = \pi_i(\vec{X}) = X_i$ d' \vec{X} són variables aleatòries.

Definició 2.3.3. Sigui $\vec{X} = (X_1, \dots, X_n)$ un vector de variables aleatòries de dimensió n . Anomenem funció de distribució de probabilitat d' \vec{X} a

$$F_{\vec{X}}(x_1, \dots, x_n) = P(X_1 \leq x_1, \dots, X_n \leq x_n).$$

Proposició 2.3.4. Sigui (Ω, \mathcal{A}, p) un espai de probabilitat i sigui $\vec{X} = (X, Y): \Omega \rightarrow \mathbb{R}^2$ un vector de variables aleatòries amb funció de distribució $F_{\vec{X}}(x, y)$. Aleshores,

- i) $\lim_{(x,y) \rightarrow (-\infty, -\infty)} F_{\vec{X}}(x, y) = 0$ i $\lim_{(x,y) \rightarrow (+\infty, +\infty)} F_{\vec{X}}(x, y) = 1$.
- ii) $x_1 \leq x_2 \implies F_{\vec{X}}(x_1, y) \leq F_{\vec{X}}(x_2, y), \forall y \in \mathbb{R}$.
- iii) $\lim_{(x,y) \rightarrow (x_0^+, y_0^+)} F_{\vec{X}}(x, y) = F_{\vec{X}}(x_0, y_0)$.
- iv) $\lim_{y \rightarrow +\infty} F_{\vec{X}}(x, y) = F_X(x)$. A aquestes funcions les anomenem distribucions marginals.
- v) $P(a < x \leq b, c < y \leq d) = F_{\vec{X}}(b, d) - F_{\vec{X}}(b, c) - F_{\vec{X}}(a, d) + F_{\vec{X}}(a, c) = \Delta_R F$,

on R és el rectangle que determinen a, b, c, d .

Demostració. Les demostracions de les tres primeres propietats son anàlogues a les de una sola dimensió. Les altres no les farem. \square

Proposició 2.3.5. Sigui una $F(x, y)$ una funció que satisfà les cinc propietats de la proposició anterior i tal que $\Delta_R F > 0$ per tot rectangle R . Aleshores,

$$\exists \vec{X} = (X, Y) \text{ variable aleatòria tal que } F(x, y) = F_{\vec{X}}.$$

Demostració. La demostració és llarga i pesada, i no la farem. \square

Definició 2.3.6. Sigui $\{X_i\}_{i \in I}$ una família de variables aleatòries sobre un espai de probabilitat (Ω, \mathcal{A}, p) . Diem que $\{X_i\}_{i \in I}$ és independent si per tot $J \subseteq I$ amb $|J| < +\infty$ i per qualssevol $B_1, \dots, B_{|J|} \in \mathcal{B}$ es té que

$$P\left(\bigcap_{j \in J} X_j \in B_j\right) = \prod_{j \in J} P(X_j \in B_j).$$

En general, no podem conèixer la distribució de X_1, \dots, X_n si coneixem només les seves distribucions marginals. Tanmateix, sí que podem conèixer-la si les variables aleatòries són independents.

Proposició 2.3.7. Si prenem $B_i = (-\infty, x_i)$ i una família finita de variables aleatòries independent $\{X_i\}_{i=1}^n$ (i posant $\vec{X} = (X_1, \dots, X_n)$), tenim que

$$F_{\vec{X}}(x_1, \dots, x_n) = \prod_{i=1}^n F_{X_i}(x_i).$$

Demostració.

$$\begin{aligned} F_{\vec{X}}(x_1, \dots, x_n) &= p\left(\bigcap_{i=1}^n X_i \in B_i\right) = p\left(\bigcap_{i=1}^n X_i \leq x_i\right) = \\ &= \prod_{i=1}^n P(X_i \leq x_i) = \prod_{i=1}^n F_{X_i}(x_i). \end{aligned}$$

□

Observació 2.3.8. Es pot demostrar que la impliació recíproca és certa, és a dir, que si una determinada família de variables aleatòries satisfà la igualtat anterior, aleshores és independent.

Observació 2.3.9. Siguin $\{X_i\}_{i=1}^n$ una família de variables aleatòries independent i siguin $f_1, \dots, f_n: \mathbb{R} \rightarrow \mathbb{R}$ funcions mesurables. Aleshores, $\{f_1(X_i)\}_{i=1}^n$ també són independents.

Proposició 2.3.10. Siguin X_1, \dots, X_n variables aleatòries independents. Aleshores

$$\mathbb{E}[X_1 \dots X_n] = \prod_{i=1}^n \mathbb{E}[X_i].$$

Definició 2.3.11. Siguin X, Y variables aleatòries. La covariància de X i Y és:

$$\text{Cov}(X, Y) = \mathbb{E}\left[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))\right]$$

Observació 2.3.12. Està ben definida sempre que $\mathbb{E}[|X|], \mathbb{E}[|Y|] < \infty$ i

$$\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

Demostració.

$$\begin{aligned} \text{Cov}(X, Y) &= \mathbb{E}[XY - \mathbb{E}[X]Y - \mathbb{E}[Y]X + \mathbb{E}[X]\mathbb{E}[Y]] = \\ &= \mathbb{E}[XY] - 2\mathbb{E}[X]\mathbb{E}[Y] + \mathbb{E}[X]\mathbb{E}[Y] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]. \end{aligned}$$

□

Observació 2.3.13. Si X i Y són independents, aleshores $\text{Cov}(X, Y) = 0$.

Observació 2.3.14. Per X, Y variables aleatòries, es satisfà $\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y] + 2\text{Cov}(X, Y)$. Si a més són independents, $\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y]$.

Demostració.

$$\begin{aligned} & \mathbb{E}[(X+Y)^2] - \mathbb{E}[X+Y]^2 = \\ &= \mathbb{E}[X^2] + \mathbb{E}[Y^2] + 2\mathbb{E}[XY] - \mathbb{E}[X]^2 - \mathbb{E}[Y]^2 - 2\mathbb{E}[X]\mathbb{E}[Y] = \\ &= \text{Var}[X] + \text{Var}[Y] + 2\text{Cov}(X, Y). \end{aligned}$$

□

Proposició 2.3.15. Siguin X, Y, Z variables aleatòries, i a, b, C constants. Aleshores

- i) $\text{Cov}(C, X) = 0$.
- ii) $\text{Cov}(X, X) = \text{Var}[X]$.
- iii) $\text{Cov}(X, Y) = \text{Cov}(Y, X)$.
- iv) $\text{Cov}(aX + bY, Z) = a\text{Cov}(X, Z) + b\text{Cov}(Y, Z)$.
- v) $\text{Cov}(X, Y)^2 \leq \text{Var}[X] \text{Var}[Y]$.

Demostració. Demostrarem només l'últim apartat. Els altres es deixen com exercici pel lector. Tenim doncs que

$$|\text{Cov}(X, Y)| = \left| \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] \right| \leq \mathbb{E} \left[|(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])| \right]$$

Que, aplicant Cauchy-Schwarz, és menor que

$$\sqrt{\mathbb{E}[(X - \mathbb{E}[X])^2]} \sqrt{\mathbb{E}[(Y - \mathbb{E}[Y])^2]} = \sqrt{\text{Var}[X]} \sqrt{\text{Var}[Y]}.$$

És a dir, $\text{Cov}(X, Y)^2 \leq \text{Var}[X] \text{Var}[Y]$. □

Observació 2.3.16. Com es conclou de la desigualtat de Cauchy-Schwarz, la desigualtat anterior és una igualtat si i només si $X - \mathbb{E}[X] = \lambda(Y - \mathbb{E}[Y])$.

Observació 2.3.17. $\text{Cov}(X, Y) = 0 \not\Rightarrow X, Y$ independents.

Demostració. En donem un contraexemple. Definim X, Y variables aleatòries tals que

$$X = \begin{cases} 1 & \text{amb probabilitat } \frac{1}{2} \\ -1 & \text{amb probabilitat } \frac{1}{2} \end{cases}, Y = \begin{cases} 0 & \text{si } X = -1 \\ \pm 1 & \text{amb probabilitat } \frac{1}{2} \text{ si } X = 1 \end{cases}.$$

En aquest cas,

$$\begin{aligned} \mathbb{E}[XY] &= \int_{\Omega} XY \, dp = \\ &= 0 \cdot P(X = -1) + 1 \cdot P(X = 1, Y = 1) - 1 \cdot P(X = 1, Y = -1) = \\ &= \frac{1}{4} - \frac{1}{4} = 0, \end{aligned}$$

i també tenim que

$$\mathbb{E}[X] = 1 \cdot P(X = 1) - 1 \cdot P(X = -1) = 0.$$

Per les últimes dues igualtats, $\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = 0$. Vegem però que no són independents:

$$P(X = 1)P(Y = 0) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}.$$

Però $P(X = 1, Y = 0) = 0$, és a dir, com volíem veure X i Y no són independents. □

Tema 3

Variables aleatòries discretes

3.1 Definició i objectes relacionats

Definició 3.1.1. Sigui (Ω, \mathcal{A}, p) un e. prob. i $X: \Omega \rightarrow \mathbb{R}$ una variable aleatòria. Diem que X és discreta si $\text{Im}(X)$ és numerable.

Observació 3.1.2. En la pràctica $\text{Im}(X) = \{x_1 < x_2 < \dots\}$ és un conjunt numerable ordenat, en els casos que veurem $\text{Im}(X) \subseteq \mathbb{Z}$. Escrivem $p_i = p(X = x_i)$

Observació 3.1.3. Sigui $A \subset \mathcal{B}$, $p_X(A) = p(x \in A) = \sum_{i \in A} p_i$.

Observació 3.1.4. Donat $\{x_i\}_{i \geq 1} \subset \mathbb{R}$ creixent i valors $\{p_i\}_{i \geq 1} \subset [0, 1]$ t. q. $\sum p_i = 1$, es pot definir una variable discreta X que pren valors a $\{x_i\}$ tals que $p(X = x_i) = p_i$.

Observació 3.1.5. La funció de distribució, amb $|\text{Im}(X)| < +\infty$, és

$$F_X(x) = \begin{cases} 0 & \text{si } x < x_1 \\ p_1 + \dots + p_j & \text{si } x_j \leq x < x_{j+1} \\ 1 & \text{si } x \geq x_n \end{cases}$$

Proposició 3.1.6. *Operador esperança.* Sigui X v.a. discreta,

i)

$$\mathbb{E}[X] = \sum_{i \geq 1} x_i p_i$$

ii) $g: \mathbb{R} \rightarrow \mathbb{R}$ mesurable,

$$\mathbb{E}[g(x)] = \sum_{i \geq 1} g(x_i) p_i$$

Demostració.

i)

$$\mathbb{E}[X] = \int_{\mathbb{R}} x \, dP_x = \sum x_i P_X(X = x_i) + \int_{\mathbb{R} \setminus \cup \{x_i\}} x \mathbb{I}(x) \, dP_X = \sum_{i \geq 1} x_i p_i + 0$$

perquè $P_X(\mathbb{R} \setminus \cup \{x_i\}) = 0$

ii) Exercici

□

Observació 3.1.7. $\text{Var}[X] = \mathbb{E}[X^2] + \mathbb{E}[X]^2 = \sum_{i \geq 1} x_i^2 p_i - (\sum_{i \geq 1} x_i p_i)^2$

Proposició 3.1.8. X, Y v.a. discretes, $\text{Im}(X) = \{x_i\}_{i \geq 1}$, $\text{Im}(Y) = \{y_i\}_{i \geq 1}$, $g: \mathbb{R}^2 \rightarrow \mathbb{R}$. Aleshores

$$\mathbb{E}[g(X, Y)] = \sum_{i, j \geq 1} g(x_i, y_j) p(X = x_i, Y = y_j)$$

Proposició 3.1.9. X, Y v.a. discretes. Són independents si i només si, $\forall x \in \text{Im}(X)$ i $\forall y \in \text{Im}(Y)$,

$$P(X = x, Y = y) = P(X = x)P(Y = y)$$

Definició 3.1.10. Sigui $(X, Y): \Omega \rightarrow \mathbb{R}^2$ un vector aleatori. Direm que és discret si $\text{Im}((X, Y))$ és numerable.

Observació 3.1.11. (X, Y) vector aleatori és discret $\iff X$ i Y són discretes.

Definició 3.1.12. Sigui (X, Y) vector aleatori discret, $\forall (X, Y) \in \text{Im}((X, Y))$ definim

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

que si X, Y són independents és $P(X = x, Y = y)$.

Lema 3.1.13. Si X, Y són v.a. discretes independents, amb $\mathbb{E}[|X|], \mathbb{E}[|Y|] < +\infty$,

$$\mathbb{E}[xy] = \mathbb{E}[x] \mathbb{E}[y]$$

Demostració.

$$\begin{aligned} \mathbb{E}[xy] &= \int_{\mathbb{R}} x \, dP_{XY} \\ &= \sum_{u \in \text{Im}(X, Y)} u P(XY = u) \\ &= \sum_{x \in \text{Im}(X) \setminus \{0\}} \left(\sum_{\frac{u}{x} \in \text{Im}(Y)} u P\left(X = x, Y = \frac{u}{x}\right) \right) \end{aligned}$$

□

3.2 Funció generadora de probabilitat

D'aquí en endavant, prendrem X variable aleatòria discreta amb $\text{Im}(X) \subseteq \mathbb{N}_{\geq 0}$.

Definició 3.2.1. La funció generadora de probabilitat de X és la sèrie formal de potències

$$G_X(z) = \sum_{n \geq 0} P(X = n) z^n.$$

Podem pensar-la com $\mathbb{E}[z^X]$.

Proposició 3.2.2. $G_X(z)$ satisfà les següents propietats:

- i) $G_X(z)$ és una funció holomorfa al voltant de $z = 0$ amb radi de convergència major o igual a 1.

ii) $G_X(0) = P(X = 0)$ i $G_X(1) = 1$.

iii) $\left. \frac{d^k G_X(z)}{dz^k} \right|_{z=1} = \mathbb{E} [X(X-1) \dots (X-k+1)] = \mathbb{E} [(X)_k]$.

Demostració.

i) Si $\rho \in \mathbb{C}$, $|\rho| < 1$, aleshores:

$$0 \leq |G_X(\rho)| = \left| \sum_{n \geq 0} P(X = n) \rho^n \right| \leq \sum_{n \geq 0} P(X = n) |\rho|^n$$

Que, quan $|\rho| \leq 1$, és menor o igual a

$$\sum_{n \geq 0} P(X = n) = 1.$$

Per tant $G_X(\rho)$ és analítica (es pot expressar com una sèrie de potències convergent) a $B_1(0)$, i per tant, com s'ha vist a variable complexa, $G_x(\rho)$ és holomorfa a $B_1(0)$ (per tant infinitament derivable en sentit complex).

ii) És directe.

iii) Si derivem terme a terme obtenim

$$\frac{d^k G_X(z)}{dz^k} = \frac{d^k \sum_{n \geq 0} P(X = n) z^n}{dz^k} = \sum_{n \geq 0} n(n-1) \dots (n-k+1) (P(X = n)) z^{n-k}.$$

Que, avaluat en $z = 1$, és

$$\sum_{n \geq 0} n(n-1) \dots (n-k+1) (P(X = n)) = \mathbb{E} [(X)_n].$$

□

Exemple 3.2.3. Definim X com una variable aleatòria discreta tal que $P(X = 0) = 0$ i $P(X = n) = \frac{6}{\pi^2} \cdot \frac{1}{n^2}$. Aleshores

$$G_X(z) = \frac{6}{\pi^2} \sum_{n \geq 1} \frac{1}{n^2} z^n,$$

que té radi de convergència igual a 1, i

$$G_X(1) = \frac{6}{\pi^2} \sum_{n \geq 1} \frac{1}{n^2} = \frac{6}{\pi^2} \cdot \frac{\pi^2}{6} = 1.$$

Finalment, en calculem la seva esperança,

$$\mathbb{E}[X] = \left. \frac{dG_X(z)}{dz} \right|_{z=1} = \frac{6}{\pi^2} \sum_{n \geq 1} \frac{1}{n} = \infty.$$

Observació 3.2.4. $G_X(z)$ codifica totes les probabilitats $P(X = n)$ i per tant coneixent $G_X(z)$ coneixem X .

L'aplicació més útil de les funcions generadores de probabilitat és que ens permet trobar convolucions discretes de variables aleatòries.

Observació 3.2.5. Siguin X, Y variables aleatòries discretes, amb $\text{Im}(X) = \text{Im}(Y) = \mathbb{N}_{\geq 0}$. Aleshores

$$P(X + Y = n) = \sum_{k=0}^n P(X + Y = n, X = k) = \sum_{k=0}^n P(Y = n - k, X = k).$$

Proposició 3.2.6. Si X, Y són variables aleatòries discretes independents amb $\text{Im}(X) = \text{Im}(Y) = \mathbb{N}_{\geq 0}$ aleshores:

$$G_{X+Y}(z) = G_X(z)G_Y(z)$$

Demostració.

$$G_X(z)G_Y(z) = \sum_{i \geq 0} P(X = i)z^i \sum_{j \geq 0} P(Y = j)z^j = \sum_{i,j \geq 0} P(X = i)P(Y = j)z^{i+j}$$

I, com X, Y són independents, podem unir el producte i obtenim

$$\begin{aligned} \sum_{i,j \geq 0} P(X = i, Y = j)z^{i+j} &= \sum_{n \geq 0} \sum_{i=0}^n P(X = i, Y = n - i)z^n = \\ &= \sum_{n \geq 0} \sum_{i=0}^n P(X = i, X + Y = n)z^n = \sum_{n \geq 0} \sum_{i=0}^n P(X = i | X + Y = n)P(X + Y = n)z^n = \\ &= \sum_{n \geq 0} P(X + Y = n)z^n \sum_{i=0}^n P(X = i | X + Y = n). \end{aligned}$$

Si observem que la suma interior val 1 perquè està sumant la probabilitat de tots esdeveniments possibles, ens queda

$$\sum_{n \geq 0} P(X + Y = n)z^n = G_{X+Y}(z).$$

□

Observació 3.2.7. Això és equivalent a que si X, Y són variables aleatòries discretes independents amb $\text{Im}(X) = \text{Im}(Y) = \mathbb{N}_{\geq 0}$ aleshores

$$\mathbb{E}[z^X] \mathbb{E}[z^Y] = \mathbb{E}[z^{X+Y}].$$

Observació 3.2.8. En general, si X_1, \dots, X_n són variables aleatòries discretes independents amb $\text{Im } X_i = \mathbb{N}_{\geq 0}$:

$$G_{X_1, \dots, X_n}(z) = \prod_{i=1}^n G_{X_i}(z)$$

3.3 Models de variables aleatòries discretes

En aquesta secció introduïrem les variables aleatòries discretes més comunes que trobarem

Observació 3.3.1. Escriurem $X \sim Y$ si X i Y tenen la mateixa distribució de probabilitat

Distribució de Bernoulli

Modela l'èxit o fracàs s'un experiment amb probabilitat p d'èxit

Definició 3.3.2.

$$X \sim B(p) \iff \begin{cases} p(X = 1) = p \\ p(X = 0) = 1 - p \end{cases}$$

s També es pot escriure $Be(p)$.

Proposició 3.3.3.

- i) $G_X(z) = (1 - p)z^0 + pz^1 = (1 - p) + p(z)$
- ii) $\mathbb{E}[X] = p$
- iii) $\mathbb{V}\text{ar}[X] = p(1 - p)$

Demostració.

$$\text{iii) } \mathbb{E}[x(x - 1)] = 0 = \mathbb{E}[x^2 - x] \implies \mathbb{E}[x^2] = p \implies \mathbb{V}\text{ar}[x] = \mathbb{E}[x^2] + \mathbb{E}[x]^2 = p - p^2 = p(1 - p)$$

□

Distribució binomial

Modela el nombre d'èxits en fer N experiments independents, on cadascun és $Be(p)$

Definició 3.3.4. $X \sim \text{Bin}(N, p) \iff X = X_1 + \dots + X_N$, on $\{X_i\}_{i=1}^N$ són independents i $X_i \sim B(p) \forall i$

Proposició 3.3.5.

- i) $p(X = i) = \binom{N}{i} p^i (1 - p)^{N-i}$
- ii) $G_X(z) = \sum_{i=0}^N \binom{N}{i} p^i (1 - p)^{N-i} z^i$
- iii) $\mathbb{E}[X] = N \mathbb{E}[X_1] = Np$
- iv) $\mathbb{V}\text{ar}[X] = N \mathbb{V}\text{ar}[X_1] = Np(1 - p)$

Demostració.

ii) Per ser X_i independents,

$$\begin{aligned} G_X(z) &= G_{X_1+\dots+X_N}(z) = \prod_{i=1}^n G_{X_i}(z) = (pz + (1-p))^N = \\ &= \sum_{i=0}^n \binom{N}{i} (pz)^i (1-p)^{N-i} = \sum_{i=0}^n \binom{N}{i} p^i (1-p)^{N-i} z^i \end{aligned}$$

□

Observació 3.3.6. La suma de dos variables amb aquesta distribució també té distribució binomial:

$$\begin{aligned} \left. \begin{array}{l} X \sim \text{Bin}(N_1, p) \\ Y \sim \text{Bin}(N_2, p) \end{array} \right\} &\implies \left. \begin{array}{l} G_X(z) = (pz + (1-p))^{N_1} \\ G_Y(z) = (pz + (1-p))^{N_2} \end{array} \right\} \implies \\ \implies G_{X+Y}(z) &= G_X(z)G_Y(z) = (pz + (1-p))^{N_1+N_2} \implies \\ &\implies X + Y \sim \text{Bin}(N_1 + N_2, p) \end{aligned}$$

Distribució uniforme

Definició 3.3.7. $X \sim U[1, n] \iff p(X = i) = \frac{1}{n}$ per $i = 1, \dots, n$.

Proposició 3.3.8. $G_X(z) = \sum_{n=1}^n \frac{1}{n} z^n = \frac{1}{n} (z + z^2 + \dots + z^n) = \frac{1}{n} \frac{z(z^n-1)}{z-1}$

Distribució de Poisson

S'usa per modelar successos “estranyos” (persones en una cua, emissió de partícules...).

Definició 3.3.9.

$$X \sim P(\lambda) \iff P(X = i) = \frac{1}{i!} \lambda^i e^{-\lambda}$$

També es pot escriure $Po(\lambda)$.

Observació 3.3.10. La distribució està ben definida:

$$\sum_{i=0}^{\infty} p(X = i) = \sum_{i=0}^{\infty} \frac{1}{i!} \lambda^i e^{-\lambda} = e^{\lambda} e^{-\lambda} = 1$$

Proposició 3.3.11.

- i) $G_X(z) = e^{\lambda(z-1)}$
- ii) $\mathbb{E}[X] = \lambda$
- iii) $\text{Var}[X] = \lambda$

Demostració.

i)

$$G_X(z) = \sum_{i=0}^{\infty} \frac{1}{i!} \lambda^i e^{-\lambda} z^i = e^{-\lambda} \sum_{i=0}^{\infty} \frac{1}{i!} (\lambda z)^i = e^{-\lambda} e^{\lambda z} = e^{\lambda(z-1)}$$

$$\text{ii) } \mathbb{E}[X] = \lambda e^{\lambda(z-1)}|_{z=1} = \lambda$$

$$\text{iii) } \mathbb{E}[X(X-1)] = \lambda^2 e^{\lambda(z-1)}|_{z=1} = \lambda^2 \implies \mathbb{E}[x^2] = \lambda^2 + \lambda \implies \text{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

□

Observació 3.3.12. La suma de dos variables amb aquesta distribució també té distribució de Poisson:

$$\begin{aligned} \left. \begin{array}{l} X \sim \text{Po}(\lambda_1) \\ Y \sim \text{Po}(\lambda_2) \end{array} \right\} &\implies \left. \begin{array}{l} G_X(z) = e^{\lambda_1(z-1)} \\ G_Y(z) = e^{\lambda_2(z-1)} \end{array} \right\} \implies \\ \implies G_{X+Y}(z) = G_X(z)G_Y(z) = e^{(\lambda_1+\lambda_2)(z-1)} &\implies \\ \implies X + Y \sim \text{Po}(\lambda_1 + \lambda_2) & \end{aligned}$$

Distribució geomètrica

Distribució binomial negativa

Índex alfabètic

- conjunt
 - d'esdeveniments, [1](#)
 - de successos, [1](#)
- covariància, [17](#)
- desviació típica, [14](#)
- distribució
 - binomial, [23](#)
 - de Bernoulli, [23](#)
 - de Poisson, [24](#)
 - marginal, [16](#)
 - uniforme, [24](#)
- esdeveniments independents, [5](#)
- espai
 - de probabilitat, [1](#)
 - producte, [5](#)
 - mostral, [1](#)
- esperança d'una variable aleatòria, [13](#)
- funció
 - de distribució de probabilitat, [16](#), [11](#)
 - de probabilitat, [1](#)
 - generadora de probabilitat, [20](#)
- límit
 - inferior d'esdeveniments, [6](#)
 - superior d'esdeveniments, [6](#)
- moment
 - d'ordre r , [14](#)
 - factorial d'ordre r , [14](#)
- probabilitat condicionada, [4](#)
- variància, [14](#)
- variable
 - aleatòria, [9](#)
 - discreta, [19](#)
 - multidimensional, [16](#)
- vector de variables aleatòries, [16](#)
 - discret, [20](#)