

# Teoria de la Probabilitat

## Continguts

<b>1</b>	<b>Espais de Probabilitat</b>	<b>3</b>
1.1	Definició axiomàtica d'espai de probabilitat . . . . .	3
	Teorema (Desigualtats de Bonferroni) . . . . .	4
1.2	Probabilitat condicionada . . . . .	5
1.3	Independència . . . . .	6
1.4	Espai producte . . . . .	7
	Teorema (d'extensió o de Carathéodory) . . . . .	7
1.5	Lema de Borel-Cantelli . . . . .	7
<b>2</b>	<b>Variables Aleatòries</b>	<b>9</b>
2.1	Definició de variable aleatòria. Llei d'una v.a. . . . .	9
2.2	Moments d'una v.a. Desigualtats de Markov i Chebyshev . . . . .	11
	Teorema (Desigualtat de Markov) . . . . .	12
	Teorema (Desigualtat de Chebyshev) . . . . .	12
2.3	Vectors de variables aleatòries. Independència de v.a. . . . .	12
	Teorema . . . . .	14
<b>3</b>	<b>Variables Aleatòries Discretes</b>	<b>16</b>
3.1	Definicions i conceptes relacionats. Funció generadora de probabilitat . . . . .	16
3.2	Models de variables aleatòries discretes . . . . .	18
3.3	Distribucions condicionades i esperança condicionada . . . . .	20
3.4	Suma de variables aleatòries discretes (Aplicació 1) . . . . .	20
3.5	Arbres de Galton-Watson (Aplicació 2) . . . . .	21
	Teorema (Galton) . . . . .	22
<b>4</b>	<b>Variables Aleatòries Contínues</b>	<b>23</b>
4.1	Mesures de probabilitat absolutament contínues. Funció de densitat . . . . .	23
	Teorema (Radon-Nikodym) . . . . .	23
4.2	Models de variables aleatòries absolutament contínues . . . . .	24
4.3	Distribucions conjuntes i marginals. Independència i distribucions condicionades	26
	4.3.1 Independència . . . . .	27
	4.3.2 Distribucions condicionades . . . . .	27
	4.3.3 Esperança condicionada . . . . .	27
4.4	Funcions de v.a. absolutament contínues i aplicacions . . . . .	28

4.4.1	Cas univariat . . . . .	28
4.4.2	Cas multivariat . . . . .	28
4.5	Distribució normal, multivariant i distribucions associades . . . . .	28
	Teorema (Moivre-Laplace) . . . . .	28
4.5.1	Distribucions associades a la normal . . . . .	29
4.5.2	Normal multivariant . . . . .	29
	Teorema . . . . .	29
	Teorema . . . . .	30
	Teorema (Fisher) . . . . .	30
<b>5</b>	<b>Funcions característiques i famílies exponencials</b>	<b>31</b>
5.1	Funció generadora de moments i funció característica . . . . .	31
5.1.1	Propietats generals de les funcions característiques . . . . .	33
	Teorema (d'inversió) . . . . .	33
5.2	Famílies exponencials . . . . .	33
<b>6</b>	<b>Convergència de variables aleatòries</b>	<b>36</b>
6.1	Modes de convergència i equivalències . . . . .	36
6.1.1	Diagrama modes de convergència . . . . .	37
6.2	Convergència quasi-segura i la llei forta dels grans nombres . . . . .	40
	Teorema (Llei forta dels grans nombres) . . . . .	40
6.3	Convergència en distribució i el teorema del límit central . . . . .	40

# 1 Espais de Probabilitat

## 1.1 Definició axiomàtica d'espai de probabilitat

### Definició 1.1.1

Un **espai de probabilitat** és un espai de mesura  $(\Omega, \mathcal{A}, p)$ , tal que  $p(\Omega) = 1$ .

- $\Omega$  s'anomena **espai mostral**.
- $\mathcal{A}$  se l'anomena **conjunt d'esdeveniments** o **successos**.
- $p$  se l'anomena **funció de probabilitat**.

**Observació 1.1.2**  $\mathcal{A} \subseteq \mathcal{P}(\Omega)$  és una  $\sigma$ -àlgebra:

$\sigma 1)$   $\emptyset \in \mathcal{A}$

$\sigma 2)$   $A \in \mathcal{A} \iff \bar{A} \in \mathcal{A}$

$\sigma 3)$  Si  $\{A_n\}_{n \geq 1}$  és una seqüència de successos en  $\mathcal{A} \implies \bigcup_{n \geq 1} A_n \in \mathcal{A}$

**Observació 1.1.3** Recordem que  $p$  és una mesura i, per tant:

$p 1)$   $p(\emptyset) = 0$

$p 2)$   $\forall A \in \mathcal{A}, p(A) \geq 0$

$p 3)$  Si  $\{A_n\}_{n \geq 1}$  és una seqüència de successos en  $\mathcal{A}$  disjunts 2 a 2 ( $A_i \cap A_j = \emptyset$  si  $i \neq j$ ), aleshores

$$p\left(\bigcup_{n \geq 1} A_n\right) = \sum_{n \geq 1} p(A_n)$$

Vegem les primeres propietats dels espais de probabilitat:

### Proposició 1.1.4

Per un espai de probabilitat  $(\Omega, \mathcal{A}, p)$  es compleix que:

(i)  $A \in \mathcal{A} \implies p(\bar{A}) = 1 - p(A)$

(ii) Si  $A, B \in \mathcal{A}, A \subseteq B \implies p(A) \leq p(B)$

(iii) Si  $A_1, \dots, A_r \in \mathcal{A}, i A_i \cap A_j \neq \emptyset (i \neq j)$ , aleshores  $p\left(\bigcup_{i=1}^r A_i\right) = \sum_{i=1}^r p(A_i)$

(iv) Si  $A, B \in \mathcal{A}, A \subseteq B \implies p(B - A) = p(B) - p(A)$

(v) Successions monòtones:

a) Si  $A_1 \subseteq A_2 \subseteq A_3 \subseteq \dots \subseteq A_i \in \mathcal{A} \implies p\left(\bigcup_{n \geq 1} A_n\right) = \lim_{n \rightarrow \infty} p(A_n)$

b) Si  $A_1 \supseteq A_2 \supseteq A_3 \supseteq \dots \supseteq A_i \in \mathcal{A} \implies p\left(\bigcap_{n \geq 1} A_n\right) = \lim_{n \rightarrow \infty} p(A_n)$

Si ara tenim un espai de probabilitat  $(\Omega, \mathcal{A}, p)$ , i successos  $A_1, A_2, \dots, A_r$  en general no disjunts, aleshores no és cert que  $p\left(\bigcup_{i=1}^r A_i\right) = \sum_{i=1}^r p(A_i)$ . En aquest cas, tenim la següent fita:

**Lema 1.1.5** (Fita de la unió)

Siguin  $A_1, A_2, \dots, A_r$  successos en  $(\Omega, \mathcal{A}, p)$ , aleshores

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

**Teorema** (Desigualtats de Bonferroni) (1.1.6)

Siguin  $A_1, \dots, A_r$  successos en  $(\Omega, \mathcal{A}, p)$ . Denotem per  $I \subseteq \{1, \dots, r\} : = [r]$ ,

$$A_I = \bigcap_{i \in I} A_i$$

$$S_k = \sum_{\substack{I \subseteq [r] \\ |I|=k}} p(A_I)$$

Aleshores, si:

- 1)  $t$  és parell,  $p\left(\bigcup_{i=1}^r A_i\right) \geq \sum_{i=1}^t (-1)^{i+1} \cdot S_i$
- 2)  $t$  és senar,  $p\left(\bigcup_{i=1}^r A_i\right) \leq \sum_{i=1}^t (-1)^{i+1} \cdot S_i$

**Exemple 1.1.7**

## 1.2 Probabilitat condicionada

### Definició 1.2.1

Sigui  $(\Omega, \mathcal{A}, p)$  un espai de probabilitat, i  $B \in \mathcal{A}$  amb  $p(B) > 0$ .

Per  $A \in \mathcal{A}$ , la **probabilitat condicionada** de  $A$  amb  $B$  és:

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

**Observació 1.2.2**  $p(A | B)$  mesura la probabilitat de que el succés  $A$  ocorri sabent que  $B$  ha succeït.

**Observació 1.2.3** Si prenem

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

Aleshores  $P_B$  és una funció de probabilitat sobre  $\Omega, \mathcal{A}$

De fet, si definim  $\mathcal{A}_B = \{A \cap B : A \in \mathcal{A}\}$ , aleshores  $\mathcal{A}_B$  és una  $\sigma$ -àlgebra i  $P_B$  també defineix una probabilitat sobre  $(\Omega, \mathcal{A}_B)$ .

### Proposició 1.2.4

Siguin  $A_1, \dots, A_r \in \mathcal{A}$ , tals que  $p(A_i) > 0$ ,  $A_i \cap A_j = \emptyset$  si  $i \neq j$  i  $\bigcup_{i=1}^r A_i = \Omega$   
( $\{A_i\}_{i=1}^r$  és una partició de  $\Omega$ )

(i) Teorema de la probabilitat total:  $\forall A \in \mathcal{A}, p(A) = \sum_{i=1}^r p(A | A_i) \cdot p(A_i)$ .

(ii) Fórmula de Bayes: si  $A \in \mathcal{A}$ ,  $p(A) > 0$ ,

$$p(A_j | A) = \frac{p(A | A_j) \cdot p(A_j)}{\sum_{i=1}^r p(A | A_i) \cdot p(A_i)}$$

Molts cops serà més senzill calcular probabilitats condicionades. Vegem un exemple:

### Exemple 1.2.5 (La ruïna del jugador)

Partim d'un capital de  $k$  unitats ( $k \geq 0$ ) i volem aconseguir un capital de  $N$  unitats ( $N \geq k$ ) de la següent forma:

Llançem una moneda equilibrada, guanyant (surta cara) o perdent (surta creu) una unitat amb probabilitat  $\frac{1}{2}$ . El joc acaba si:

1. Ens quedem sense capital.
2. Assolim un capital igual a  $N$

Calcularem la probabilitat de perdre.

No és bona idea intentar codificar tots els casos possibles i sumar les seves probabilitats ( $n^\circ$  tirades  $\rightarrow \infty$ ).

Anem a resoldre el problema usant el teorema de la probabilitat total.

$A_k$  = "el jugador, amb capital inicial igual a  $k$ , s'arruïna" (Volem calcular  $p(A_k)$ ).

Escrivim  $p(A_k) = p_k$ . Aleshores  $p_0 = 1$ ,  $p_N = 0$ .

Definim  $B$  = "la primera tirada de la moneda és cara". Aleshores  $B$  i  $\overline{B}$  defineixen una partició de  $\Omega$ .

Useu el teorema de probabilitat total:

$$\left. \begin{aligned} p(A_k) &= p(A_k | B)p(B) + p(A_k | \overline{B})p(\overline{B}) \\ &= p(A_k | B) \cdot \frac{1}{2} + p(A_k | \overline{B}) \cdot \frac{1}{2} \\ &= \frac{1}{2} \cdot \left( \underbrace{p(A_k | B)}_{p(A_{k+1})} + \underbrace{p(A_k | \overline{B})}_{p(A_{k-1})} \right) \end{aligned} \right\} p_k = \frac{1}{2}(p_{k+1} + p_{k-1}); \quad p_0 = 1, p_N = 0$$

Resolent la recurrència tenim:

$$p_k = 1 - \frac{k}{N}$$

Si  $k \in o(N)$ , aleshores asimptòticament el jugador s'acabarà arruïnant.

## 1.3 Independència

### Definició 1.3.1

Dos successos  $A$  i  $B$  (del mateix espai  $(\Omega, \mathcal{A}, p)$ ) són **independents** si  $p(A \cap B) = p(A) \cdot p(B)$ .

### Observació 1.3.2

- Si  $p(B) > 0$ ,  $A$  i  $B$  independents  $\iff p(A | B) = p(A)$
- El successos  $\emptyset$  i  $\Omega$  són independents amb qualsevol altre succés  $B$
- Si  $A$  i  $B$  són independents,  $A$  i  $\overline{B}$  també. (De fet tenim que  $A$  i  $B$  són independents  $\iff \overline{A}$  i  $\overline{B}$  són independents).

### Definició 1.3.3

Donada una família de successos  $\{A_i\}_{i \in I}$ , es diu que és **independent** si:

$$\forall J \subseteq I \text{ amb } |J| < \infty, \text{ es té: } p\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} p(A_j)$$

**Observació 1.3.4** Donats  $\{A_i\}_{i=1, \dots, r}$  successos, poden ser independents dos a dos però no com a conjunt.

## 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 combinar-los en un sol espai amb espai mostral  $\Omega_1 \times \Omega_2$ .

Això ja ho hem fet en el cas discret, però en general:

- Prenem com a espai mostral  $\Omega_1 \times \Omega_2$
- Prenem com a  $\sigma$ -àlgebra  $\sigma(\mathcal{A}_1 \times \mathcal{A}_2)$  (la més petita  $\sigma$ -àlgebra que conté  $\mathcal{A}_1 \times \mathcal{A}_2$ ).
- Com definim  $p$  a  $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \times \mathcal{A}_2)$ ?  
Voldríem que  $p(A_1 \times A_2) = p(A_1) \cdot p(A_2) \quad \forall A_i \in \mathcal{A}_i$ . Per definir  $p$  necessitarem el següent teorema:

**Teorema** (d'extensió o de Carathéodory) (1.4.1)

Sigui  $p_0$  una funció de probabilitat en una àlgebra  $\mathcal{A}_0$ . Sigui  $\mathcal{A} = \sigma(\mathcal{A}_0)$ .

Aleshores  $p_0$  es pot estendre a una funció de probabilitat  $p$  sobre  $\mathcal{A}$  que coincideix amb  $p_0$  en  $\mathcal{A}_0$ . A més,  $p$  és única.

En el nostre cas, quina àlgebra agafem?

$$\underbrace{(\mathcal{A}_1 \times \mathcal{A}_2)^*}_{\mathcal{A}_0} = \text{totes les unions finites d'elements de } \mathcal{A}_1 \times \mathcal{A}_2.$$

Es comprova que  $\sigma((\mathcal{A}_1 \times \mathcal{A}_2)^*) = \sigma(\mathcal{A}_1 \times \mathcal{A}_2)$  i apliquem el teorema.

## 1.5 Lema de Borel-Cantelli

Sigui  $(\Omega, \mathcal{A}, p)$  un espai de probabilitat.

### Definició 1.5.1

Donada  $\{A_n\}_{n \geq 1}$  un successió de successos,

$$\limsup A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$$

$$\liminf A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k$$

**Observació 1.5.2** Tant  $\limsup A_n$  com  $\liminf A_n$  són successos.

### Definició 1.5.3

Si  $\{A_n\}_{n \geq 1}$  és una successió de successos amb  $\limsup A_n = \liminf A_n$ , aleshores  $\{A_n\}_{n \geq 1}$  té **límit** i l'escriurem  $\lim A_n = \limsup A_n$ .

**Propietats 1.5.4**

Sigui  $\{A_n\}_{n \geq 1}$  una successió de successos:

- (i)  $\limsup A_n = \{\omega \in \Omega : \omega \text{ pertany a un nombre infinit de } A_i's\}$
- (ii)  $\liminf A_n = \{\omega \in \Omega : \exists m = m(\omega) \text{ tal que } \omega \in A_r \text{ si } r \geq m(\omega)\}$
- (iii)  $\liminf A_n \subseteq \limsup A_n$

Si ara  $\limsup A_n = \liminf A_n = \lim A_n$ , podem intercanviar  $\lim$  amb  $p$ :

**Proposició 1.5.5**

Sigui  $\{A_n\}_{n \geq 1}$  una successió de successos amb  $\lim A_n = A$ . Aleshores,  $p(A) = \lim p(A_n)$ .

El següent lema ens dóna condicions suficients molt senzilles per a calcular  $p(\limsup A_n)$ .

**Lema 1.5.6** (Borel-Cantelli)

Sigui  $(\Omega, \mathcal{A}, p)$  un espai de probabilitat i  $\{A_n\}_{n \geq 1}$  una successió de successos. Aleshores:

- (i) Si  $\sum_{n \geq 1} p(A_n) < \infty \implies p(\limsup A_n) = 0$
- (ii) Si els  $\{A_n\}_{n \geq 1}$  són successos independents i  $\sum_{n \geq 1} p(A_n) = \infty \implies p(\limsup A_n) = 1$

**Observació 1.5.7** No podem treure la condició d'independència de (ii).

Si prenem  $A_n = E$  succés amb  $0 < p(E) < 1$ , aleshores  $\sum_{n \geq 1} p(A_n) = \sum_{n \geq 1} p(E)$ , els  $A_n$  no són independents i  $p(\limsup A_n) = p(E) < 1$ .



## 2 Variables Aleatòries

### 2.1 Definició de variable aleatòria. Llei d'una v.a.

Sigui  $(\Omega, \mathcal{A}, \beta)$  un espai de probabilitat. Volem estudiar funcions de  $\Omega$  amb imatge en  $\mathbb{R}$ .

#### Definició 2.1.1

Una **variable aleatòria** és una funció  $X: \Omega \rightarrow \mathbb{R}$  tal que per tot borelià  $B \in \mathcal{B}$ ,  $X^{-1}(B) \in \mathcal{A}$ .

Per tant, una variable aleatòria és una funció mesurable entre els espais de mesura  $(\Omega, \mathcal{A}, p)$  i  $(\mathbb{R}, \mathcal{B}, \lambda)$ .

#### Exemple 2.1.2

(1) Les funcions constants són variables aleatòries:

$$\begin{array}{ll} X: \Omega \rightarrow \mathbb{R} \\ \omega \mapsto c \end{array} \quad \text{Si prenem } B \in \mathcal{B}, \quad X^{-1}(B) = \begin{cases} \emptyset & \text{si } c \notin B \\ \Omega & \text{si } c \in B \end{cases}$$

(2) **Variables aleatòries indicadores:**

$$\text{Sigui } A \in \mathcal{A}, \text{ definim } \mathbb{1}_A: \Omega \rightarrow \mathbb{R} \text{ on } \mathbb{1}_A(\omega) = \begin{cases} 0 & \text{si } \omega \notin A \\ 1 & \text{si } \omega \in A \end{cases}$$

$$\text{Aleshores, } B \in \mathcal{B}, \mathbb{1}_A^{-1}(B) = \begin{cases} \emptyset & \text{si } \{0, 1\} \not\subseteq B \\ A & \text{si } 1 \in B, \quad 0 \notin B \\ \bar{A} & \text{si } 1 \notin B, \quad 0 \in B \\ \Omega & \text{si } \{0, 1\} \subseteq B \end{cases}$$

(3) Si  $X$  i  $Y$  són v.a., aleshores  $X + Y$ ,  $X \cdot Y$ ,  $|X|$ , etc. són v.a.

En general, si  $g: \mathbb{R}^2 \rightarrow \mathbb{R}$  és una funció mesurable, aleshores  $g(X, Y)$  és una v.a.

Estem dient que  $\forall B \in \mathcal{B}$ ,  $\{\omega \in \Omega: X(\omega) \in B\}$  és un succés i, per tant, podem calcular  $P(\{\omega \in \Omega: X(\omega) \in B\}) \equiv P(X \in B)$ .

#### Exemple 2.1.3

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

Les v.a. permeten traslladar l'estructura d'espai de probabilitat de  $(\Omega, \mathcal{A}, p)$  en  $(\mathbb{R}, \mathcal{B})$ , donant lloc a mesures que no provenen de la mesura de Lebesgue.

**Definició 2.1.4**

Siguin  $(\Omega, \mathcal{A}, p)$  un espai de probabilitat i  $X$  una v.a.

La **mesura de probabilitat induïda** per  $X$  és una mesura de probabilitat sobre  $(\mathbb{R}, \mathcal{B})$  definida per

$$\begin{aligned} p_X: \mathcal{B} &\rightarrow \mathbb{R} \\ B &\mapsto p_X = P(\{\omega \in \Omega: X(\omega) \in B\}) \end{aligned}$$

**Observació 2.1.5**  $(\mathbb{R}, \mathcal{B}, p_X)$  és un espai de probabilitat.

De teoria de la mesura, és equivalent veure que  $[\forall B \in \mathcal{B}, \quad X^{-1}(B) \text{ és de } \mathcal{A}]$  a veure que *[l'antiimatge de qualsevol interval  $\in \mathcal{A}$ ].*

Per tant, per saber si una funció és una v.a. només cal veure si l'antiimatge dels intervals són de  $\mathcal{A}$ .

La següent definició dóna una funció en  $\mathbb{R}$  que codifica molta informació de  $X$ :

**Definició 2.1.6**

Donada  $X$  v.a., la **funció de distribució de probabilitat** de  $X$  és:

$$\begin{aligned} F_X: \mathbb{R} &\rightarrow [0, 1] \\ x &\mapsto P(X \leq x) \end{aligned}$$

**Propietats 2.1.7**

(i) Si  $x_1 \leq x_2 \implies F_X(x_1) \leq F_X(x_2)$

(ii)  $\lim_{x \rightarrow -\infty} F_X(x) = 0$ ,  $\lim_{x \rightarrow +\infty} F_X(x) = 1$

(iii)  $F_X(x)$  és contínua per la dreta:  $\forall x, \lim_{h \rightarrow 0^+} F_X(x+h) = F_X(x)$

**Observació 2.1.8**

- $P(X > x) = 1 - P(X \leq x) = 1 - F_X(x)$
- $P(x_1 < X \leq x_2) = P(X \leq x_2) - P(X \leq x_1) = F_X(x_2) - F_X(x_1)$

**Observació 2.1.9** Les propietats (i), (ii), (iii) de  $F_X(x)$  són de fet suficients.

Si una funció  $F(x)$  satisfà (i), (ii), (iii), aleshores és funció de probabilitat d'una variable aleatòria.

## 2.2 Moments d'una v.a. Desigualtats de Markov i Chebyshev

Siguin  $(\Omega, \mathcal{A}, p)$  uns espai de probabilitat i  $X$  una v.a.

### Definició 2.2.1

L'esperança de  $X$  és:

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

Més en general, si  $f: \mathbb{R} \rightarrow \mathbb{R}$  és una funció mesurable,

$$\mathbb{E}[f(x)] = \int_{\Omega} f(x) dp = \int_{\mathbb{R}} f(x) dp_X$$

**Observació 2.2.2** De teoria de la mesura, cal recordar que una funció  $g$  és integrable sii  $|g|$  ho és (En general,  $\mathbb{E}[f(x)]$  està definida sii  $\mathbb{E}[|f(x)|] < +\infty$ ).

Si particularitzem  $f$ :

### Definició 2.2.3

$f(x) = X^r \implies \mathbb{E}[X^r]$  és el **moment r-èssim**.

### Definició 2.2.4

Si  $\mathbb{E}[X] = p < +\infty$ ,  $\mathbb{E}[(X - p)^r]$  és el **moment normalitzat r-èssim**.

En particular, si  $r = 2$ ,  $\mathbb{E}[(X - p)^2] = \mathbb{V}ar[X]$  és la **variància** de  $X$ .

### Definició 2.2.5

Si  $f(x) = x(x-1) \dots (x-r+1) \implies \mathbb{E}[f(x)] = \mathbb{E}[(X)_r]$  és el **moment factorial r-èssim**.

**Proposició 2.2.6** (Propietats de l'esperança i la variància)

- Si  $c$  és la v.a. constant,  $\mathbb{E}[c] = c$  i  $\mathbb{V}ar[c] = 0$
- Linealitat: si  $a, b \in \mathbb{R}$  i  $X, Y$  v.a.,  $\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$
- $A \in \mathcal{A}$ ,  $X = \mathbb{1}_A$ ,  $\mathbb{E}[\mathbb{1}_A] = P(A)$
- $|\mathbb{E}[X]| \leq \mathbb{E}[|X|]$
- $\mathbb{V}ar[c \cdot X] = c^2 \cdot \mathbb{V}ar[X]$
- $\mathbb{V}ar[c + X] = \mathbb{V}ar[X]$
- $\mathbb{V}ar[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$

**Observació 2.2.7** Si  $\mathbb{E}[|X|^p] < +\infty$ , aleshores podem utilitzar tots els resultats de teoria dels espais  $L_p$ . Així doncs tenim les següents conseqüències:

- Hölder:  $p, q > 0, \frac{1}{p} + \frac{1}{q} = 1, \mathbb{E}[|X|^p] < +\infty, \mathbb{E}[|Y|^q] < +\infty$   
 $\implies \mathbb{E}[|XY|] \leq \mathbb{E}[|X|^p]^{\frac{1}{p}} \cdot \mathbb{E}[|Y|^q]^{\frac{1}{q}} \quad (\mathbb{E}[|XY|]^{pq} \leq \mathbb{E}[|X|^p]^q \cdot \mathbb{E}[|Y|^q]^p)$
- Cauchy-Schwartz: si  $\mathbb{E}[X^2], \mathbb{E}[Y^2] < +\infty$ , aleshores  $\mathbb{E}[XY]^2 \leq \mathbb{E}[X^2] \cdot \mathbb{E}[Y^2]$
- Minkowski: si  $\mathbb{E}[|X|^p], \mathbb{E}[|Y|^p] < +\infty \implies \mathbb{E}[|X + Y|^p]^{\frac{1}{p}} \leq \mathbb{E}[|X|^p]^{\frac{1}{p}} + \mathbb{E}[|Y|^p]^{\frac{1}{p}}$

**Teorema** (Desigualtat de Markov) (2.2.8)

Sigui  $X$  un v.a. que pren valors positius i  $a > 0$ . Aleshores:

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

El següent resultat dóna estimacions quantitatives de quant es dispersa una v.a. en relació a la seva esperança:

**Teorema** (Desigualtat de Chebyshev) (2.2.9)

Sigui  $X$  una v.a. en  $(\Omega, \mathcal{A}, p)$  amb  $\mathbb{E}[X], \mathbb{V}ar[X] < +\infty$ . Aleshores,  $\forall k > 0$

$$P(|X - \mathbb{E}[X]| \geq k \cdot \mathbb{V}ar[X]^{\frac{1}{2}}) \leq \frac{1}{k^2}$$

També es pot escriure:

$$P(|X - \mathbb{E}[X]| \geq k) \leq \frac{\mathbb{V}ar[X]}{k^2}$$

## 2.3 Vectors de variables aleatòries. Independència de v.a.

Donat un espai de probabilitat  $(\Omega, \mathcal{A}, p)$  considerem les v.a.  $X_1, \dots, X_n$ . Cadascuna d'elles defineix una distribució de probabilitat sobre  $\mathbb{R}$ .

Aleshores podem considerar el vector  $(X_1, \dots, X_n): \Omega \rightarrow \mathbb{R}^n$ .

### Definició 2.3.1

Un vector  $(X_1, \dots, X_n): \Omega \rightarrow \mathbb{R}^n$  és un **vector de variables aleatòries** (o una v.a. multi-dimensional), si per tot  $B \in \mathcal{B}_n$  (Borelians en  $\mathbb{R}^n$ ),  $(X_1, \dots, X_n)^{-1}(B) \in \mathcal{A}$ .

Com  $\Pi_i: \mathbb{R}^n \rightarrow \mathbb{R}$  (projecció en la  $i$ -èssima component) és una funció mesurable, aleshores

$$\begin{aligned} \Pi_i(X_1, \dots, X_n): \Omega &\xrightarrow{(X_1, \dots, X_n)} \mathbb{R}^n \xrightarrow{\Pi_i} \mathbb{R} \\ \omega &\longmapsto (X_1, \dots, X_n)(\omega) \longmapsto X_i(\omega) \end{aligned}$$

$\equiv X_i(\omega)$  és una v.a. (en el sentit unidimensional).

De la mateixa manera que vam fer per les v.a. unidimensionals, podem considerar les antiimatges només en intervals.

**Definició 2.3.2**

Donat un espai de probabilitat  $(\Omega, \mathcal{A}, p)$ , i un vector de v.a.  $(X_1, \dots, X_n) = \vec{X}$ , aleshores la **funció de distribució de probabilitat** de  $\vec{X}$  és  $F_{\vec{X}}(x_1, \dots, x_n)$  definida per:

$$F_{\vec{X}}: \mathbb{R}^n \rightarrow [0, 1] \subseteq \mathbb{R}$$

$$(x_1, \dots, x_n) \mapsto P\left((X_1 \leq x_1) \cap (X_2 \leq x_2) \cap \dots \cap (X_n \leq x_n)\right) = P\left(\bigcap_{i=1}^n X_i \leq x_i\right)$$

Vegem propietats de la funció de distribució pel cas  $n = 2$  (Per  $n > 2$ , és idèntic):

**Lema 2.3.3**

(i) Si  $x'_1 \geq x_1, x'_2 \geq x_2 \implies F_{\vec{X}}(x'_1, x'_2) \geq F_{\vec{X}}(x_1, x_2)$

(ii)  $\lim_{(x_1, x_2) \rightarrow (+\infty, +\infty)} F_{\vec{X}}(x_1, x_2) = 1$   $\lim_{(x_1, x_2) \rightarrow (-\infty, -\infty)} F_{\vec{X}}(x_1, x_2) = 0$

(iii)  $\lim_{(h_1, h_2) \rightarrow (0^+, 0^+)} F_{\vec{X}}(x_1 + h_1, x_2 + h_2) = F_{\vec{X}}(x_1, x_2)$  (contínua "per dalt")

**Observació 2.3.4** Aquestes 3 condicions són necessàries i suficients per a definir una v.a. multidimensional.

**Observació 2.3.5**

- Si tenim  $F_{\vec{X}}(x_1, x_2)$  associada a  $\vec{X} = (X_1, X_2)$ , aleshores

$$\lim_{x_2 \rightarrow +\infty} F_{\vec{X}}(x_1, x_2) = \lim_{x_2 \rightarrow +\infty} P\left((X_1 \leq x_1) \cap (X_2 \leq x_2)\right) = P(X_1 \leq x_1) = F_{\vec{X}}(x_1)$$

A aquesta funció  $\left(\lim_{x_1 \rightarrow +\infty} F_{\vec{X}}(x_1, x_2)\right)$  se l'anomena **funció de distribució marginal**.

- Prenem un rectangle en  $\mathbb{R}^2$  i  $\vec{X} = (X, Y)$  v.a. multidimensional:

$$P(a < X \leq b, c < Y \leq d) = F_{\vec{X}}(b, d) - F_{\vec{X}}(a, d) - F_{\vec{X}}(b, c) + F_{\vec{X}}(a, c)$$

**Definició 2.3.6**

Sigui  $(\Omega, \mathcal{A}, p)$  un espai de probabilitat i  $\{x_i\}_{i \in I}$  un conjunt de v.a. Direm que  $\{x_i\}_{i \in I}$  són **independents** si:  $\forall k, \forall i_1, \dots, i_k \subseteq I, \forall B_1, \dots, B_k \in \mathcal{B}$

$$P(X_{i_1} \in B_1, X_{i_2} \in B_2, \dots, X_{i_k} \in B_k) = \prod_{j=1}^k P(X_{i_j} \in B_j)$$

Si ara prenem  $X_1, \dots, X_k$  v.a., aleshores si són independents,

$$F_{X_1, \dots, X_k}(x_1, \dots, x_k) = P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_k \leq x_k) = \prod_{j=1}^k P(X_j \leq x_j) = \prod_{j=1}^k F_{X_j}(x_j)$$

**Observació 2.3.7** Si  $X$  i  $Y$  són v.a. independents i  $f, g$  funcions mesurables, aleshores  $f(X)$  i  $g(Y)$  són també independents.

**Observació 2.3.8** Si  $F_{X_1, \dots, X_k}(x_1, \dots, x_k) = \prod_{j=1}^k F_{X_j}(x_j)$ , aleshores les v.a.  $X_1, \dots, X_k$  són independents.

En quant al càlcul de moments, tenim el següent resultat:

**Teorema (2.3.9)**

Siguin  $X_1, \dots, X_k$  v.a. independents. Aleshores, si  $\mathbb{E}[X_i] < +\infty$ , es compleix que

$$\mathbb{E}\left[\prod_{i=1}^k X_i\right] = \prod_{i=1}^k \mathbb{E}[X_i]$$

Vam veure que l'operador esperança és lineal, però això no és cert en general per la variància. De fet, en el segon cas obtenim un terme connector anomenat covariància.

**Definició 2.3.10**

Donades dues v.a.  $X, Y$ , la **covariància** de  $X$  i  $Y$  ( $Cov(X, Y)$ ) és:

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

Si ara desenvolupem aquesta expressió, obtenim:  $Cov(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$ .

Observem que si  $X$  i  $Y$  són independents, aleshores  $X - \mathbb{E}[X]$  i  $Y - \mathbb{E}[Y]$  també ho són i, per tant:

$$Cov(X, Y) = \mathbb{E}\left[(X - \mathbb{E}[X]) \cdot (Y - \mathbb{E}[Y])\right] = \mathbb{E}\left[X - \mathbb{E}[X]\right] \cdot \mathbb{E}\left[Y - \mathbb{E}[Y]\right] = (\mathbb{E}[X] - \mathbb{E}[X]) \cdot (\mathbb{E}[Y] - \mathbb{E}[Y]) = 0$$

**Observació 2.3.11**  $Cov(X, Y) = 0 \not\Rightarrow X$  i  $Y$  independents.

Si ara calculem la variància d'una suma, obtenim:

$$\mathbb{V}ar[X + Y] = \mathbb{V}ar[X] + \mathbb{V}ar[Y] + 2 \cdot Cov(X, Y)$$

**Observació 2.3.12** Si  $X$  i  $Y$  són independents,  $Cov(X, Y) = 0$  i per tant,  $\mathbb{V}ar[X + Y] = \mathbb{V}ar[X] + \mathbb{V}ar[Y]$ .

En general, això és cert per  $n$  v.a. independents:

$$\mathbb{V}ar\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n \mathbb{V}ar[X_i] \quad \text{si } \{X_i\}_{i=1}^n \text{ són independents.}$$

Vegem ara propietats de la covariància:

**Propietats 2.3.13**

- (i)  $Cov(c, X) = 0$  ( $c$  és una constant)
- (ii) Si  $c$  és una constant,  $Cov(c + X, Y) = Cov(X, Y)$
- (iii)  $Cov(X, X) = \mathbb{V}ar[X]$
- (iv)  $Cov(X, Y) = Cov(Y, X)$
- (v)  $Cov(aX + bY, Z) = a \cdot Cov(X, Z) + b \cdot Cov(Y, Z)$

**Definició 2.3.14**

Anomenem **coeficient de correlació de Pearson** ( $\rho(X, Y)$ ) a:

$$\rho(X, Y) = \frac{Cov(X, Y)}{\mathbb{V}ar[X]^{\frac{1}{2}} \cdot \mathbb{V}ar[Y]^{\frac{1}{2}}} \in [-1, 1]$$

**Observació 2.3.15**

Sabem que la igualtat en *Cauchy-Schwartz* es dona quan  $\exists a$  tal que  $a \cdot X = Y$ .  
A més,  $Cov(aX + b, Y) = Cov(aX, Y)$ . Per tant,

$$Y = aX + b \iff \rho(X, Y) \in \{\pm 1\}$$

Així doncs, com més proper sigui  $\rho(X, Y)$  a  $\pm 1$ , millor serà una aproximació lineal de  $Y$  usant  $X$ .

### 3 Variables Aleatòries Discretes

#### 3.1 Definicions i conceptes relacionats. Funció generadora de probabilitat

Sigui  $(\Omega, \mathcal{A}, p)$  un espai de probabilitat i  $X$  una variable aleatòria.

##### Definició 3.1.1

$X$  és una **variable aleatòria discreta** si  $Im(X)$  és numerable.

Si  $X$  és una variable aleatòria discreta,  $Im(X) = \{x_i\}_{i \geq 1}$ .

$X$  ve completament determinada pels valors  $p(X = x_i) = p_i$ .

##### Definició 3.1.2

Anomenem **funció de distribució** de  $X$  a:

$$F_X(x) = p(X \leq x) = \sum_{x_i \leq x} p_i$$

##### Definició 3.1.3

Sigui  $A \in \mathcal{B}$ , la **mesura de probabilitat induïda sobre  $\mathbb{R}$**  és:

$$p_X(A) = \sum_{\substack{x_i \in A \\ x_i \in Im(X)}} p_i$$

En particular, si  $Im(X) \cap A = \emptyset$  aleshores  $p_X(A) = 0 \implies$  obtenim una probabilitat puntual (hi ha punts de  $\mathbb{R}$  amb probabilitat  $> 0$ , a diferència de la mesura de Lebesgue).

##### Definició 3.1.4

Definim l'**esperança matemàtica** com:

$$\mathbb{E}[X] = \int_{\Omega} X dp \stackrel{\text{def. de } \int_{\Omega}}{=} \sum_{i \geq 1} x_i \cdot p(X = x_i) = \sum_{i \geq 1} x_i \cdot p_i$$

Més en general, si  $g(x)$  és una funció mesurable,

$$\mathbb{E}[g(X)] = \sum_{i \geq 1} g(x_i) \cdot p(X = x_i)$$

En particular,

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

$$\text{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \sum_{i \geq 1} x_i^2 \cdot p_i - \left( \sum_{i \geq 1} x_i \cdot p_i \right)^2$$

##### Definició 3.1.5

Prenem  $X, Y$  variables aleatòries discretes amb  $Im(X), Im(Y)$  numerables.

El **vector de variables aleatòries**  $(X, Y)$  ve completament caracteritzat per:

$$p_{(X,Y)}(x_i, y_j) = p(X = x_i, Y = y_j) \quad (x_i \in Im(X); y_j \in Im(Y))$$



Aleshores, tenim les següents propietats:

- (i)  $X$  i  $Y$  són independents  $\iff \forall x_i \in \text{Im}(X), \forall y_j \in \text{Im}(Y), p_{(X,Y)}(x_i, y_j) = p_X(x_i) \cdot p_Y(y_j)$
- (ii) Si  $X, Y$  són independents,  $\mathbb{E}[X \cdot Y] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$

Sigui  $X$  una variable aleatòria discreta amb  $\text{Im}(X) \subseteq \mathbb{N}_{\geq 0}$ . En aquesta situació, tenim la següent definició:

**Definició 3.1.6**

La funció generadora de probabilitat associada a  $X$  és:

$$G_X(z) = \sum_{i \geq 0} p(X = i) \cdot z^i = \sum_{i \geq 0} p_i \cdot z^i$$

Una funció generadora de probabilitat és un objecte formal. En particular,  $G_X(z) = \mathbb{E}[z^X]$ .

Si ens mirem les funcions generadores de probabilitat com funcions (en  $\mathbb{C}$ ), aleshores compleixen el següent:

**Proposició 3.1.7**

1.  $G_X(0) = p(X = 0)$
2.  $G_X(1) = 1$ . A més, si  $|z| \leq 1$  ( $z \in \mathbb{C}$ )  $\implies |G_X(z)| \leq 1$   
Per tant,  $G_X(z)$  (com a sèrie de potències) té radi de convergència  $\geq 1$ .
3.  $\mathbb{E}[X] = \frac{d}{dz} G_X(z) \Big|_{z=1}$
4. Més en general,  $\mathbb{E}[(X)_k] = \frac{d^k}{dz^k} G_X(z) \Big|_{z=1}$   
En particular,  $\text{Var}[X] = G_X''(1) + G_X'(1) - (G_X'(1))^2$
5. Si  $z \in \mathbb{R}$ ,  $z \in [0, 1]$ , aleshores  $G_X(z)$  és una funció creixent.

La propietat fonamental de les funcions generadores de probabilitat és que permeten estudiar fàcilment sumes de variables aleatòries independents.

**Proposició 3.1.8**

Siguin  $X, Y$  variables aleatòries discretes independents amb imatge en  $\mathbb{N}$ , i amb funcions generadores de probabilitat  $G_X(z), G_Y(z)$ . Aleshores:

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

**Corol·lari 3.1.9**

Si  $X_1, \dots, X_N$  són variables aleatòries discretes independents amb imatge en  $\mathbb{N}$ , aleshores:

$$G_{X_1 + \dots + X_N}(z) = \prod_{i=1}^N G_{X_i}(z)$$

### 3.2 Models de variables aleatòries discretes

Seguidament veurem famílies importants de variables aleatòries discretes:

1) Variable aleatòria Bernoulli:  $X \sim B(p)$  (un sol paràmetre)

Modela el llançament d'una moneda amb probabilitat d'èxit igual a  $p$ :

$$p(X = 1) = p, \quad p(X = 0) = 1 - p$$

- $G_X(z) = p \cdot z + (1 - p)$  (és una funció entera)
- $\mathbb{E}[X] = p$
- $\mathbb{V}ar[X] = p \cdot (1 - p)$

2) Binomial:  $X \sim Bin(p, n)$  (o bé  $B(p, n)$ )

Nombre d'èxits al tirar una moneda  $n$  vegades. L'èxit individual té probabilitat  $p$ , i cada tirada de la moneda és independent de la resta.

$$\implies X = Y_1 + \dots + Y_n \text{ on } Y_i \sim B(p)$$

$$p(X = i) = \binom{n}{i} \cdot p^i \cdot (1 - p)^{n-i} \quad i = 0, \dots, n$$

- $G_X(z) = (p \cdot z + (1 - p))^n$  (és una funció entera)
- $\mathbb{E}[X] = n \cdot p$
- $\mathbb{V}ar[X] = n \cdot p \cdot (1 - p)$

3) Poisson:  $X \sim Po(\lambda)$

$$p(X = k) = \frac{1}{k!} \cdot \lambda^k \cdot e^{-\lambda} \quad (k \in \mathbb{N}_{\geq 0})$$

- $G_X(z) = e^{\lambda(z-1)}$  (és una funció entera)
- $\mathbb{E}[X] = \lambda$
- $\mathbb{V}ar[X] = \lambda$

4) Uniforme:  $X \sim U(N)$  (on  $Im(X) = \{1, 2, \dots, N\}$ )

$$p(X = k) = \frac{1}{n}$$

- $G_X = \frac{1}{N} \cdot \frac{1 - z^N}{1 - z} \cdot z$  (sing. evitable en  $z = 1$ )
- $\mathbb{E}[X] = \frac{N + 1}{2}$
- $\mathbb{V}ar[X] = \frac{N^2 - 1}{12}$

5) Geomètrica:  $X \sim Geom(p)$  ( $p \in (0, 1)$ )

Nombre de tirades d'una moneda (amb probabilitat d'èxit =  $p$ ) fins aconseguir el primer èxit.

$$p(X = k) = (1 - p)^{k-1} \cdot p \quad k = 1, 2, \dots$$

- $G_X(z) = \frac{p \cdot z}{1 - (1 - p) \cdot z}$  (té un pol en  $z = \frac{1}{1 - p}$ )

- $\mathbb{E}[X] = \frac{1}{p}$

- $\mathbb{V}ar[X] = \frac{1 - p}{p^2}$

6) Binomial negativa:  $X \sim BinN(r, p)$

Variable aleatòria que compta el nombre de tirades necessàries per aconseguir  $r$  èxits.  
( $Im(X) = \{r, r + 1, r + 2, \dots\}$ )

$$p(X = k) = \binom{k-1}{r-1} \cdot p^r \cdot (1 - p)^{k-r}$$

Podem interpretar una binomial negativa com una suma de geomètriques independents.

- $G_X(z) = \left( \frac{p \cdot z}{1 - (1 - p) \cdot z} \right)^r$

- $\mathbb{E}[X] = \frac{r}{p}$

- $\mathbb{V}ar[X] = r \cdot \frac{1 - p}{p^2}$

### 3.3 Distributions condicionades i esperança condicionada

Siguin  $X, Y$  dues variables aleatòries discretes. Volem definir la noció de condicionar una variable aleatòria a l'altra ("  $Y \mid X$  ").

#### Definició 3.3.1

Donat  $x$  amb  $p(X = x) > 0$ , diem que la **funció de probabilitat condicionada** de  $Y$  amb  $X$  és, per aquest valor de  $x$ :

$$p_{Y|X}(y, x) = p(Y = y \mid X = x)$$

#### Definició 3.3.2

Amb les mateixes condicions que abans, la **funció de distribució de probabilitat condicionada** és:

$$F_{Y|X}(y, x) = p(Y \leq y \mid X \leq x)$$

**Observació 3.3.3** Si  $X$  i  $Y$  són independents,  $X \mid Y = y \sim X$

#### Definició 3.3.4

L'**esperança condicionada** de  $Y$  a  $X = x$  és:

$$\psi(x) = \mathbb{E}[Y \mid X = x] = \sum_{y \in \text{Im}(Y)} (y \cdot p_{Y|X})(y, x)$$

Amb aquesta definició estem definint una variable aleatòria  $E[Y \mid X]$  que pren el valor  $\psi(x)$  amb probabilitat  $p(X = x)$ .

#### Proposició 3.3.5

$$\mathbb{E}[\mathbb{E}[Y \mid X]] = \mathbb{E}[Y]$$

**Observació 3.3.6**  $\mathbb{E}[Y] = \sum_{x \in \text{Im}(X)} \mathbb{E}[Y \mid X = x] \cdot p(X = x)$

**Observació 3.3.7**  $\mathbb{E}[Y \mid X]$  és la millor aproximació de  $Y$  com a funció de  $X$ .

### 3.4 Suma de variables aleatòries discretes (Aplicació 1)

Ja hem vist que si  $X$  i  $Y$  són variables aleatòries discretes amb imatge en  $\mathbb{N}$ , aleshores si són independents,  $G_{X+Y}(z) = G_X(z) \cdot G_Y(z)$

Si fem  $Z = X + Y$ :

$$\begin{aligned} G_Z(z) &= \sum_{n \geq 0} p(Z = n) \cdot z^n = \sum_{r \geq 0} p(X = r) \cdot z^r \cdot \sum_{s \geq 0} p(Y = s) \cdot z^s \\ \implies p(Z = n) &= \sum_{r+s=n} p(X = r) \cdot p(Y = s) = \underbrace{\sum_{m=0}^n p(X = m) \cdot p(Y = n - m)}_{\text{Convolució}} \end{aligned}$$

Aplicant aquest principi de l'ús de funcions generadores, tenim que:

1.  $X_1, X_2, \dots, X_n \sim \text{Ber}(p)$  independents  $\implies X_1 + X_2 + \dots + X_n \sim \text{Bin}(n, p)$
2.  $X_1, X_2, \dots, X_n \sim \text{Geom}(p)$  independents  $\implies X_1 + X_2 + \dots + X_n \sim \text{BinN}(n, p)$
3.  $X_1 \sim \text{Po}(\lambda_1), X_2 \sim \text{Po}(\lambda_2), X_1, X_2$  independents

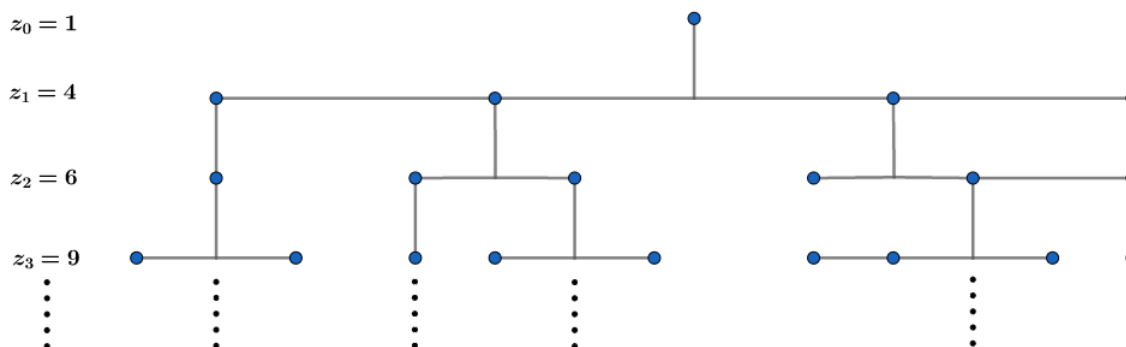
$$\left. \begin{array}{l} G_{X_1}(z) = e^{-\lambda_1} \cdot e^{\lambda_1 \cdot z} \\ G_{X_2}(z) = e^{-\lambda_2} \cdot e^{\lambda_2 \cdot z} \end{array} \right\} \implies G_{X_1+X_2}(z) = e^{-(\lambda_1+\lambda_2)} \cdot e^{(\lambda_1+\lambda_2) \cdot z}$$

$$\implies X_1 + X_2 \sim \text{Po}(\lambda_1 + \lambda_2)$$

### 3.5 Arbres de Galton-Watson (Aplicació 2)

Sigui  $X$  una variable aleatòria discreta amb  $\text{Im}(X) \subseteq \mathbb{N}_{\geq 0}$ .

Considerem el següent procés estocàstic:



Arbre de Galton-Watson

$$Z_0 = 1$$

En les següents generacions, el nombre de fills ve donat per la v.a.  $X$ .

$Z_1 = \text{n}^\circ$  de fills del primer individu (generació 0), seguint  $X$ .

$Z_2 = \text{n}^\circ$  de fills dels individus de la generació 1.

.

.

.

$Z_n = \text{n}^\circ$  de descendents en la generació n-èsima.

Propietat fonamental: el nombre de descendents d'un individu és independent del nombre de descendents de qualsevol altre individu.

**Lema 3.5.1** Sigui  $N$  una variable aleatòria amb imatge en  $\{1, 2, 3, \dots\}$ .

$X_1, X_2, X_3, \dots$  variables aleatòries independents (i independents amb  $N$ ), amb  $X_i \sim X$ .

Prenem  $Y = X_1 + X_2 + \dots + X_N$

$$\implies G_Y(z) = G_N(G_X(z))$$

**Proposició 3.5.2**

$G_{Z_{n+m}}(z) = G_{Z_n}(G_{Z_m}(z))$ . En particular  $G_{Z_n}(z) = G_X \circ \dots \circ G_X(z)$

**Observació 3.5.3** Podem calcular  $\mathbb{E}[Z_n]$  i  $\text{Var}[Z_n]$  a partir de la seva funció generadora de probabilitat:

$$\begin{aligned} \mathbb{E}[Z_n] &= \mathbb{E}[X]^n \\ \text{Var}[Z_n] &= \begin{cases} n \cdot \text{Var}[X] & \text{si } \mathbb{E}[X] = 1 \\ \frac{\text{Var}[X] \cdot (\mathbb{E}[X]^n - 1)}{\mathbb{E}[X] - 1} & \text{si } \mathbb{E}[X] \neq 1 \end{cases} \end{aligned}$$

Ara, la pregunta que ens fem és: quan hi haurà extinció?

Si fem  $A_n = \{Z_n = 0\} \implies \text{Extinció} = \bigcup_{n \geq 1} A_n$

$$\implies p(\text{Extinció}) = p\left(\bigcup_{n \geq 1} A_n\right)$$

$$A_1 \subseteq A_2 \subseteq A_3 \subseteq \dots$$

$$\left. \begin{array}{l} \text{Extingit a la primera gen} \implies \text{Extingit a la segona gen.} \implies \dots \\ \text{Extingit a la segona gen.} \not\Rightarrow \text{Extingit a la primera} \dots \end{array} \right\} \implies$$

$\implies$  Tenim una successió creixent de successos.

$$A_1 \subseteq A_2 \subseteq A_3 \subseteq \dots \text{ és creixent } \implies p(\text{Extinció}) = \lim_{n \rightarrow \infty} p(A_n) = \lim_{n \rightarrow \infty} p(Z_n = 0)$$

**Teorema (Galton) (3.5.4)**

Si  $\mathbb{E}[X] = \mu$ ;  $p(X = 0) \geq 0$

I sigui  $\alpha$  la solució més petita positiva de l'equació  $G_X(z) = z$  (Aleshores  $p(\text{Extinció}) = \alpha$ )

1. Si  $\mu < 1 \implies \alpha = 1$  (règim subcrític)
2. Si  $\mu = 1 \implies \alpha = 1$  (sempre que  $\text{Var}[X] \neq 0$ ) (règim crític)
3. Si  $\mu > 1 \implies 0 < \alpha < 1$  (règim supercrític)

## 4 Variables Aleatòries Contínues

### 4.1 Mesures de probabilitat absolutament contínues. Funció de densitat

Sigui  $(\Omega, \mathcal{A}, p)$  un espai de probabilitat, amb probabilitat induïda en  $(\mathbb{R}, \mathcal{B}, p_X)$  (quan prenem la variable aleatòria  $X$ ).

#### Definició 4.1.1

Siguin  $\mu_1, \mu_2$  mesures sobre un espai de mesura  $(X, \mathcal{A})$ , diem que  $\mu_1$  és **absolutament contínua** respecte  $\mu_2$  ( $\mu_1 \ll \mu_2$ ) si  $\forall A \in \mathcal{A}$ ,

$$\mu_2(A) = 0 \implies \mu_1(A) = 0$$

#### Definició 4.1.2

Una variable aleatòria  $X$  és **absolutament contínua** (o contínua per abreujar), si  $p_X \ll \lambda$  ( $\lambda$  és la mesura de Lebesgue).

**Observació 4.1.3** Les variables aleatòries discretes no són absolutament contínues:

Si  $p(X = a) = P_a > 0$  i prenem  $B = \{a\}$ , tenim  $\lambda(\{a\}) = 0$ , però  $p_X(\{a\}) = P_a > 0$

El teorema fonamental que ens permet traduir  $p_X$  (si  $X$  és absolutament contínua) a càlculs usant la mesura  $\lambda$ , és el següent:

#### Teorema (Radon-Nikodym) (4.1.4)

Sigui  $(X, \mathcal{A})$  un espai mesurable i  $\mu_1, \mu_2$  mesures sobre  $(X, \mathcal{A})$  amb  $\mu_1 \ll \mu_2$ . Aleshores, existeix una funció  $f_{\mu_1, \mu_2}$ -mesurable tal que

$$\forall A \in \mathcal{A}, \mu_1(A) = \int_A f_{\mu_1} d\mu_2$$

A més,  $f_{\mu_1}$  és única  $\mu_2$ -gairebé arreu

#### Definició 4.1.5

La funció  $f_{\mu_1}$  és la **funció de densitat** de la mesura  $\mu_1$  respecte a  $\mu_2$ .

En el nostre context,  $X$  és una variable aleatòria absolutament contínua, i

$$\left. \begin{array}{l} (\mathbb{R}, \mathcal{B}, \underbrace{\lambda}_{\mu_2}) \\ (\mathbb{R}, \mathcal{B}, \underbrace{p_X}_{\mu_1}) \end{array} \right\} \xRightarrow{X \text{ abs. cont.}} p_X \ll \lambda \xRightarrow{R-N} \forall B \in \mathcal{B}, \boxed{p_X(B) = \int_B f_X d\lambda}$$

#### Definició 4.1.6

La funció  $f_X$  s'anomena **funció de densitat de probabilitat** de  $X$ .

**Observació 4.1.7** En la literatura, s'escriu  $f_{\mu_1} = \frac{d\mu_1}{d\mu_2}$  (Derivada de Radon-Nikodym).

**Proposició 4.1.8**

Si  $X$  és una variable aleatòria absolutament contínua, amb funció de densitat de probabilitat  $f_X(x)$ :

1.  $f_X(x) \geq 0$   $\lambda$ -gairebé arreu
2.  $F_X(x) = \int_{(-\infty, x)} f_X d\lambda$ ;  $\int_{\mathbb{R}} f_X d\lambda = 1$
3.  $p(X = x) = \int_{\{x\}} f_X d\lambda = 0 \forall x \in \mathbb{R}$
4. Si  $f_X$  és integrable Riemann,  $F_X(x) = \int_{-\infty}^x f_X(x) dx$  i  $\frac{dF_X(x)}{dx} = f_X(x)$

**Observació 4.1.9** Tota funció mesurable  $f(x)$  que compleixi:

$$\int_{\mathbb{R}} f d\lambda = 1, \quad f(x) \geq 0 \quad \lambda\text{-gairebé arreu.}$$

Aleshores  $\exists X$  variable aleatòria absolutament contínua per la qual  $f(x) = f_X(x)$ .  
En aquest context, tenim:

$$\mathbb{E}[X] = \int_{\Omega} X dp = \int_{\mathbb{R}} x dp_X \stackrel{R-N}{=} \int_{\mathbb{R}} x \underbrace{f_X(x) d\lambda}_{dp_X} \stackrel{\text{si } f_X \text{ int. R.}}{=} \int_{-\infty}^{+\infty} x f_X(x) dx$$

En particular,

$$\mathbb{V}ar[X] = \int_{-\infty}^{+\infty} x^2 f_X(x) dx - \left( \int_{-\infty}^{+\infty} x f_X(x) dx \right)^2$$

**4.2 Models de variables aleatòries absolutament contínues**

1. Uniforme:  $X \sim U([a, b])$  (Tria un nombre uniformement a l'atzar en l'interval  $[a, b]$ ).

$$f_X(x) = \frac{1}{b-a} \mathbb{1}_{[a,b]}(x)$$

$$\mathbb{E}[X] = \frac{b+a}{2}$$

$$\mathbb{V}ar[X] = \frac{(b-a)^2}{12}$$

2. Exponencial:  $X \sim Exp(\lambda)$  (S'usa per modelar el tipus de vida d'un aparell. És l'anàleg continu de la Poisson).

$$f_X(x) = \lambda \cdot e^{-\lambda \cdot x} \cdot \mathbb{1}_{[0,+\infty)}(x)$$

$$\mathbb{E}[X] = \frac{1}{\lambda}$$

$$\mathbb{V}ar[X] = \frac{1}{\lambda^2}$$



3. Normal:  $X \sim N(\mu, \sigma^2)$  ( $\mu \in \mathbb{R}, \sigma > 0$ )

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$\mathbb{E}[X] = \mu$$

$$\mathbb{V}ar[X] = \sigma^2$$

4. Gamma:  $X \sim \Gamma(\lambda, \tau)$

$$f_X(x) = \frac{1}{\Gamma(\tau)} \cdot \lambda^\tau \cdot x^{\tau-1} \cdot e^{-\lambda x} \cdot \mathbb{1}_{[0,+\infty)}(x)$$

$$\mathbb{E}[X] = \frac{\tau}{\lambda}$$

$$\mathbb{V}ar[X] = \frac{\tau}{\lambda^2}$$

5. Weibull:  $X \sim Weib(\alpha, \beta)$

$$f_X(x) = \alpha \cdot \beta \cdot x^{\beta-1} \cdot e^{-\alpha \cdot x^\beta} \cdot \mathbb{1}_{[0,+\infty)}(x)$$

$$\mathbb{E}[X] = \alpha^{-\frac{1}{\beta}} \cdot \Gamma(1 + \frac{1}{\beta})$$

$$\mathbb{V}ar[X] = \alpha^{-\frac{2}{\beta}} (\Gamma(1 + \frac{2}{\beta}) - \Gamma^2(1 + \frac{1}{\beta}))$$

6. Beta:  $X \sim \beta(a, b)$

$$f_X(x) = \frac{\Gamma(a+b)}{\Gamma(a) \cdot \Gamma(b)} \cdot x^{a-1} \cdot (1-x)^{b-1} \cdot \mathbb{1}_{[0,1]}(x)$$

$$\mathbb{E}[X] = \frac{a}{a+b}$$

$$\mathbb{V}ar[X] = \frac{a \cdot b}{(a+b)^2 \cdot (a+b+1)}$$

7. Cauchy:  $X \sim Cauchy$

$$f_X(x) = \frac{1}{\pi \cdot (1+x^2)} \quad (x \geq 0)$$

Cap dels  $\mathbb{E}[X^k]$  és finit.

### 4.3 Distributions conjuntes i marginals. Independència i distribucions condicionades

Ara farem el mateix que per una variable en el cas de tenir vectors de variables aleatòries. Ho farem per vectors  $(X_1, X_2)$ , però és fàcilment generalitzable a vectors de dimensió  $> 2$ . Sigui  $(X, Y)$  un vector de variables aleatòries.  $(X, Y)$  induïx  $p_{(X,Y)}$  mesura de probabilitat en  $\mathbb{R}^2$ .

#### Definició 4.3.1

$(X, Y)$  és un **vector absolutament continu** si  $p_{(X,Y)} \ll \lambda_{\mathbb{R}^2}$ .  
Per *Radon – Nikodym*,  $(X, Y)$  té una **funció de densitat**:

$$B \in \mathcal{B}_{\mathbb{R}^2}, p_{(X,Y)}(B) = \iint_B f_{(X,Y)} d\lambda_{\mathbb{R}^2}$$

#### Definició 4.3.2

$f_{X,Y}(x, y)$  és la **funció de densitat conjunta** de  $X$  i  $Y$ .  
Si ara  $B = (a_1, a_2) \times (b_1, b_2)$  i  $f_{(X,Y)}(x, y)$  és integrable Riemann,

$$p_{(X,Y)}(B) = \iint_B f_{(X,Y)} d\lambda_{\mathbb{R}^2} = \int_{a_1}^{a_2} \int_{b_1}^{b_2} f_{(X,Y)}(x, y) dy dx$$

A partir de  $f_{(X,Y)}(x, y)$  podem extreure la funció de densitat de  $X$  i de  $Y$  integrant:

#### Definició 4.3.3

Donat un vector de variables aleatòries  $(X, Y)$  amb funció de densitat conjunta  $f_{(X,Y)}(x, y)$ , les **funcions de densitat marginals** són:

$$f_X(x) = \int_{-\infty}^{+\infty} f_{(X,Y)}(x, y) dy$$
$$f_Y(y) = \int_{-\infty}^{+\infty} f_{(X,Y)}(x, y) dx$$

A partir de la funció de densitat conjunta podem definir la **funció de distribució**:

$$F_{(X,Y)}(x, y) = P(X \leq x, Y \leq y) \stackrel{\text{int. Riem.}}{=} \int_{-\infty}^x \int_{-\infty}^y f_{(X,Y)}(x, y) dy dx$$

**Observació 4.3.4** Si  $f_{X,Y}(x, y)$  és integrable Riemann  $\implies F_{(X,Y)}(x, y)$  és  $\mathcal{C}^2(\mathbb{R}^2)$  i

$$\frac{\partial^2 F_{(X,Y)}}{\partial x \partial y}(x, y) = \frac{\partial^2 F_{(X,Y)}}{\partial y \partial x}(x, y) = f_{(X,Y)}(x, y)$$

**Observació 4.3.5** Si  $g : \mathbb{R}^2 \rightarrow \mathbb{R}$  és una funció mesurable,  $g(X, Y)$  és una variable aleatòria i es compleix:

$$\mathbb{E}[g(X, Y)] = \iint_{\mathbb{R}^2} g(x, y) \cdot f_{(X,Y)}(x, y) dx dy$$

### 4.3.1 Independència

Si  $(X, Y)$  és un vector aleatori i  $X, Y$  són independents, aleshores  $F_{X,Y}(x, y) = F_X(x) \cdot F_Y(y)$ . Ara si  $F_{X,Y}$  és  $\mathcal{C}^2$  i les  $F_X, F_Y$  són derivables (això sempre passa en el cas absolutament continuu), llavors

$$f_{(X,Y)}(x, y) = \frac{\partial^2 F_{(X,Y)}}{\partial x \partial y} = \frac{\partial F_X}{\partial x} \cdot \frac{\partial F_Y}{\partial y} = f_X(x) \cdot f_Y(y)$$

Per tant,  $X$  i  $Y$  variables aleatòries absolutament contínues són independents si i

$$f_{(X,Y)}(x, y) = f_X(x) \cdot f_Y(y)$$

Notació: si  $(X, Y)$  és un vector aleatori,

$$\mathbb{E}[(X, Y)] = (\mathbb{E}[X], \mathbb{E}[Y])$$

$$\mathbb{V}ar[(X, Y)] = \begin{pmatrix} \mathbb{V}ar[X] & \text{Cov}(X, Y) \\ \text{Cov}(X, Y) & \mathbb{V}ar[Y] \end{pmatrix} \underset{\text{si } X, Y \text{ ind}}{=} \begin{pmatrix} \mathbb{V}ar[X] & 0 \\ 0 & \mathbb{V}ar[Y] \end{pmatrix}$$

### 4.3.2 Distribucions condicionades

Siguin  $X, Y$  variables aleatòries absolutament contínues, amb funció de densitat conjunta  $f_{(X,Y)}(x, y)$  i marginals  $f_X(x), f_Y(y)$ . Sigui  $x$  tal que  $f_X(x) > 0$ .

#### Definició 4.3.6

La variable aleatòria  $Y \mid X = x$  és una variable aleatòria que té com a funció de distribució:

$$F_{Y|X}(y, x) = \frac{1}{f_X(x)} \cdot \int_{-\infty}^y f_{(X,Y)}(x, u) du$$

Si  $f_{(X,Y)}(x, y)$  és integrable respecte a  $y$ , aleshores,  $F_{Y|X}(y, x)$  és derivable respecte a  $y$ , i es compleix:

$$f_{Y|X}(y, x) = \frac{\partial F_{Y|X}(y, x)}{\partial y} = \frac{f_{(X,Y)}(x, y)}{f_X(x)}$$

### 4.3.3 Esperança condicionada

En les condicions d'abans, definim:

$$\mathbb{E}[Y \mid X = x] = \int_{-\infty}^{+\infty} u \cdot f_{Y|X}(u, x) du = \psi(x) \implies \mathbb{E}[Y \mid X] = \psi(X)$$

Totes les propietats que vam veure en el cas discret per l'esperança condicionada s'apliquen aquí de la mateixa manera. En particular:

#### Proposició 4.3.7

$$\mathbb{E}[\mathbb{E}[Y \mid X]] = \mathbb{E}[Y] \quad (\text{si } (X, Y) \text{ és abs. cont.})$$

## 4.4 Funcions de v.a. absolutament contínues i aplicacions

Sigui  $\vec{X} = (X_1, \dots, X_n)$  un vector aleatori absolutament continu i  $G : \mathbb{R}^n \rightarrow \mathbb{R}^n$  una funció bijectiva.

Com relacionem  $f_{(X_1, \dots, X_n)}(x_1, \dots, x_n)$  amb  $f_{(Y_1, \dots, Y_n)}(y_1, \dots, y_n)$ ?

### 4.4.1 Cas univariat

Sigui  $X$  una variable aleatòria absolutament contínua amb funció de densitat  $f_X(x)$ . Sigui  $g : \mathbb{R} \rightarrow \mathbb{R}$  una funció bijectiva, derivable i estrictament creixent. Considerem ara  $g(X) = Y$  (i  $g^{-1} = h$ ). Aleshores,

$$f_Y(u) \text{ i } f_X(h(u)) \cdot h'(u) \text{ són iguals } \lambda - \text{gairebé per tot}$$

**Observació 4.4.1** Si  $g$  no és bijectiva o no és estrictament creixent, en general l'anàlisi és més complicat.

### 4.4.2 Cas multivariat

$G(\vec{X}) = \vec{Y}$  on  $G$  és bijectiva i  $\mathcal{C}^1(\mathbb{R}^n)$ , aleshores, de manera similar al cas anterior, tenim:

$$f_{(Y_1, \dots, Y_n)}(y_1, \dots, y_n) = f_{(X_1, \dots, X_n)}(G^{-1}(y_1, \dots, y_n)) \cdot |Jac G^{-1}(y_1, \dots, y_n)|$$

## 4.5 Distribució normal, multivariant i distribucions associades

Ja vam veure que la distribució normal ve donada per la seva esperança i la seva variància:

$$X \sim N(\mu, \sigma^2)$$

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{\frac{-1}{2\sigma^2} \cdot (x-\mu)^2}$$

La suma de normals independents és normal.

**Teorema** (Moivre-Laplace) (4.5.1)

Sigui  $X \sim Bin(n, p)$ . Aleshores

$$p\left(a \leq \frac{X - np}{\sqrt{np(1-p)}} \leq b\right) \xrightarrow{n \rightarrow \infty} \int_b^a \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}} dx$$

#### 4.5.1 Distributions associades a la normal

- $X_i \sim N(0, 1)$ ;  $X_1, X_2, \dots$  independents.

$$X_1^2 + X_2^2 + \dots + X_n^2 = \chi_n^2 \quad \equiv \textbf{Chi quadrat} \text{ (amb } n \text{ graus de llibertat)}$$

- Si tenim  $\chi_{d_1}^2$  i  $\chi_{d_2}^2$ , aleshores

$$\frac{\chi_{d_1}^2/d_1}{\chi_{d_2}^2/d_2} \equiv \textbf{F-Fisher-Snedecor} \text{ (amb paràmetres } d_1 \text{ i } d_2)$$

- $X \sim N(0, 1)$ ,  $\chi_k^2$  independent de  $X$ .

$$\frac{X}{\sqrt{\chi_k^2/k}} \equiv \textbf{t - de Student} \text{ (amb } k \text{ graus de llibertat)}$$

#### 4.5.2 Normal multivariant

##### Definició 4.5.2

Sigui  $\vec{X} = (X_1, \dots, X_n)$  un vector aleatori, direm que  $\vec{X}$  és un **vector de variables aleatòries normal multivariant** si la seva funció de densitat conjunta és:

$$f_{\vec{X}}(x_1, \dots, x_n) = f_{\vec{X}}(\vec{x}) = \frac{1}{\sqrt{(2\pi)^n \cdot |\det(\Sigma)|}} \cdot e^{-\frac{1}{2}(\vec{x} - \vec{\mu})^T \cdot \Sigma^{-1} \cdot (\vec{x} - \vec{\mu})}$$

on

- $\vec{\mu}$  és un vector de  $\mathbb{R}^n$  (vector d'esperances)
- $\Sigma$  és una matriu  $n \times n$  simètrica i definida positiva (matriu de covariàncies)

$$\begin{aligned} \vec{\mu} &= \mathbb{E}[\vec{x}] \\ \Sigma &= \left( Cov(x_i, x_j) \right)_{i,j} \end{aligned}$$

Cas particular: Si  $\Sigma = Id$ ,  $\vec{\mu} = \vec{0}$ , aleshores escriurem  $\vec{X} = \vec{U}$  i es compleix que

$$f_{\vec{U}}(\vec{x}) = \frac{1}{\sqrt{(2\pi)^n}} \cdot e^{-\frac{1}{2}(\vec{x} - \vec{\mu})^T \cdot (\vec{x} - \vec{\mu})} = \frac{1}{\sqrt{(2\pi)^n}} \cdot e^{-\frac{1}{2} \cdot \sum_{i=1}^n x_i^2} = \prod_{i=1}^n f_{x_i}(x_i)$$

Per tant, les marginals són normals univariades i independents.

Ara veurem que tot vector normal multivariant s'obté de fet com una transformació lineal de  $\vec{U}$ .

##### Teorema (4.5.3)

$\vec{X} = (X_1, \dots, X_n)$  segueix una llei normal multivariant  $\iff \vec{X} = A \cdot \vec{X} + \vec{b}$  on  $A$  és una matriu no singular ( $\Sigma = A^T \cdot diag \cdot A$ ).

**Teorema (4.5.4)**

Sigui  $\vec{X} \sim N(\vec{\mu}, \Sigma)$  de dimensió  $n$ , i  $M$  una matriu  $m \times n$  de rang màxim ( $m \leq n$ ). Aleshores  $M \cdot \vec{X}$  és un vector de variables aleatòries normal multivariant  $N(M \cdot \vec{\mu}, M \cdot \Sigma \cdot M^T)$ .

En el cas de  $m = 1$  tenim el següent corol·lari:

**Corol·lari 4.5.5**

Si  $a_1, \dots, a_n$  són nombres tals que  $\sum a_i^2 > 0$  (la matriu  $(a_1, \dots, a_n)$  té rang 1) i  $\vec{X} \sim N(\vec{\mu}, \Sigma)$ ,  $\vec{X} = (X_1, \dots, X_n)$

$$M\vec{X} = \sum_{i=1}^n a_i x_i \sim N\left(\sum_{i=1}^n a_i \mu_i, \sum_{i=1}^n a_i^2 \sigma_i^2 + 2 \sum_{i < j} a_i a_j \sigma_{ij}\right) \quad \text{on} \quad \begin{cases} \vec{\mu} = (\mu_1, \dots, \mu_n) \\ \sigma_i^2 = \text{Var}[X_i] \\ \sigma_{ij} = \text{Cov}(X_i, X_j) \end{cases}$$

Seguidament veurem els estimadors i el teorema de Fisher:

**Definició 4.5.6**

Siguin  $X_1, X_2, \dots, X_n$  variables aleatòries idènticament distribuïdes. L'**esperança mostral** i la **variància mostral** són:

$$\bar{X} = \frac{X_1 + \dots + X_n}{n}$$

$$S^2 = \frac{1}{n-1} \cdot \sum_{i=1}^n (x_i - \bar{X})^2$$

En particular,  $\mathbb{E}[\bar{X}] = \mathbb{E}[X_1]$ ,  $\text{Var}[\bar{X}] = \frac{\text{Var}[X_1]}{n}$  i  $\mathbb{E}[S^2] = \text{Var}[X_1]$

Si, a més,  $X_1, \dots, X_n$  són  $N(\mu, \sigma^2)$ , tenim el següent teorema:

**Teorema (Fisher) (4.5.7)**

Si  $X_1, \dots, X_n$  són independents i  $N(\mu, \sigma^2)$ , aleshores  $\bar{X}$  i  $S^2$  són independents.

A més,  $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$  i  $S^2 \sim \chi_{n-1}^2$

## 5 Funcions característiques i famílies exponencials

### 5.1 Funció generadora de moments i funció característica

#### Definició 5.1.1

Sigui  $(\Omega, \mathcal{A}, p)$  un espai de probabilitat i  $X$  una variable aleatòria, la **funció generadora de moments** de  $X$  és  $M_X(s)$ , definida per:

$$M_X(s) = \mathbb{E}[e^{sX}] \quad \text{on } s \in \mathbb{C} \quad (\text{No té per què estar definida})$$

**Propietats 5.1.2** 1.  $M_X(0) = 1$

2. Si  $\mathbb{E}[X^k] < \infty \forall k \geq 1$ , aleshores  $\mathbb{E}[X^k] = M_X^{(k)}(0)$

3. Si  $Y = aX + b$ ,

$$M_Y(s) = \mathbb{E}[e^{s(aX+b)}] = e^{sb} \mathbb{E}[e^{saX}] = e^{sb} \cdot M_X(as)$$

4. Si  $X, Y$  són indep.  $\implies M_{X+Y}(s) = \mathbb{E}[e^{s(X+Y)}] = \mathbb{E}[e^{sX}] \cdot \mathbb{E}[e^{sY}] = M_X(s) \cdot M_Y(s)$

Aquesta definició particularitza en el cas discret i en el cas continuu de la següent manera:

- Cas discret:  $\mathbb{E}[e^{sX}] = \sum_{x \in \text{Im}(X)} e^{sx} p(X=x)$

- Cas continuu:  $\mathbb{E}[e^{sx}] = \int_{\mathbb{R}} e^{sx} f_X(x) dx$

#### Exemple 5.1.3

- $X \sim \text{Ber}(p) \implies M_X(s) = e^{s \cdot 0} \cdot (1-p) + e^{s \cdot 1} \cdot p = p \cdot e^s + (1-p)$

$M_X(s)$  és una funció entera  $\implies$  podem calcular tots els moments:  $\mathbb{E}[X^k] = p$

- Si  $Y \sim \text{Bin}(N, p) \implies Y = X_1 + \dots + X_N$  on  $X_i \sim \text{Ber}(p)$ ,  $\{X_i\}_{i=1}^N$  independents

$$M_Y(s) = (p \cdot e^s + (1-p))^N$$

- $X \sim \text{Exp}(\lambda) \quad (\lambda > 0)$

$$M_X(s) = \int_{\mathbb{R}} e^{sx} \cdot \lambda e^{-\lambda x} dx = \lambda \int_{-\infty}^{+\infty} e^{(s-\lambda)x} dx = \lambda \cdot \left. \frac{e^{(s-\lambda)x}}{s-\lambda} \right|_{-\infty}^{+\infty}$$

(No està definida per alguns valors de  $s$ )

- $X \sim \text{Cauchy}$ .  $X$  té funció de densitat  $\frac{1}{\pi(1+x^2)}$

$$\mathbb{E}[e^{sX}] = \int_{\mathbb{R}} \frac{e^{sx}}{\pi(1+x^2)} dx \quad (\text{No té sentit si } \text{Re}(s) > 0)$$

Estem veient que la funció generadora de moments no convergeix en la majoria de casos. Això és degut a que  $s \in \mathbb{C}$ .

Per a solucionar aquest problema, prenem  $s = it$  ( $t \in \mathbb{R}$ ).

### Definició 5.1.4

La funció  $\mathbb{E}[e^{itX}] = M_X(it) = \Phi_X(t)$  és la **funció característica** de  $X$ .

**Observació 5.1.5** Totes les propietats de la funció generadora de moments es compleixen per les funcions característiques:

1.  $\Phi_X(0) = 1$
2.  $\mathbb{E}[X^k] = M_X^{(k)}(0) = \frac{d^k}{ds^k} M_X(s) \Big|_{s=0} = \frac{1}{(i)^k} \cdot \Phi_X^{(k)}(0)$
3.  $Y = aX + b \implies \Phi_Y(t) = e^{ibt} \cdot \Phi_X(at)$
4.  $X, Y$  independents  $\implies \Phi_{X+Y}(t) = \Phi_X(t) \cdot \Phi_Y(t)$

### Exemple 5.1.6

- $X \sim \text{Exp}(\lambda)$  (amb  $\lambda > 0$ ). Té funció de densitat  $f_X(x) = \lambda e^{-\lambda x} \cdot \mathbb{1}_{[0, \infty)}(x)$

$$\Phi_X(t) = \int_0^\infty e^{itx} \cdot \lambda e^{-\lambda x} dx = \lambda \int_0^\infty e^{(it-\lambda)x} dx = \lambda \frac{e^{(it-\lambda)x}}{it-\lambda} \Big|_0^\infty = \frac{-\lambda}{it-\lambda}$$

(Està definida per tot valor de  $t$ )

- $X \sim \text{Cauchy}$

$$\Phi_X(t) = \int_{-\infty}^{+\infty} \frac{e^{itx}}{\pi(1+x^2)} dx = e^{-|t|}$$

S'observa que la variable aleatòria *Cauchy* no té moment, però sí que té funció característica (No es pot derivar  $\Phi_X(t)$  en  $t = 0$ )

- $X \sim N(0, 1) : \Phi_X(t) = e^{-\frac{t^2}{2}}$

Com a conseqüència, si  $X \sim N(\mu, \sigma^2)$ , aleshores  $X = \sigma Y + \mu$  (on  $Y \sim N(0, 1)$ )

$$\implies \Phi_X(t) = e^{i\mu t - \frac{1}{2}\sigma^2 t^2}$$

- $X \sim \text{Geom}(p): \Phi_X(t) = \frac{p}{e^{it} - (1-p)}$

- $X \sim \text{Pois}(\lambda): \Phi_X(t) = e^{\lambda(e^{it}-1)}$

- $X \sim U(0, 1): \Phi_X(t) = \frac{e^{it} - 1}{it}$  (Té una singularitat evitable en  $t=0$ )



### 5.1.1 Propietats generals de les funcions característiques

#### Proposició 5.1.7

Sigui  $X$  una variable aleatòria amb funció característica  $\Phi_X(t)$ . Aleshores:

1.  $\Phi_X(0) = 1$  i  $|\Phi_X(t)| \leq 1 \forall t \in \mathbb{R}$
2.  $\Phi_X(t)$  és uniformement contínua en  $\mathbb{R}$ .
3.  $\forall t_1, \dots, t_n \in \mathbb{R}, \forall z_1, \dots, z_n \in \mathbb{C}, \sum_{j,k} \Phi_X(t_j - t_k) z_j \cdot \bar{z}_k \geq 0$

#### Teorema (d'inversió) (5.1.8)

Sigui  $X$  una variable aleatòria que indueix una probabilitat  $p_X$  sobre  $\mathbb{R}$  i funció característica  $\Phi_X(t)$ . Aleshores  $\forall a, b \in \mathbb{R} (a < b)$ ,

$$\underbrace{p_X((a, b))}_{p(a < X < b)} + \frac{1}{2}(p_X(\{a\}) + p_X(\{b\})) = \lim_{T \rightarrow +\infty} \frac{1}{2\pi} \int_{-T}^T \frac{e^{-ita} - e^{-itb}}{it} \cdot \Phi_X(t) dt$$

#### Lema 5.1.9

Donada una variable aleatòria  $X$ , el nombre de discontinuïtats de  $F_X(x)$  és un conjunt numerable.

#### Corol·lari 5.1.10

$\Phi_X(t)$  caracteritza completament  $X$ .

## 5.2 Famílies exponencials

Ara veurem que podem tractar de manera molt general famílies de variables aleatòries usant la noció de família exponencial.

#### Definició 5.2.1

Una família de variables aleatòries és **exponencial** amb paràmetres  $\vec{\theta} = (\theta_1, \dots, \theta_n)$  si la funció de probabilitat (cas discret) o la funció de densitat (cas continuu) té la forma:

$$p(x | \vec{\theta}) = f(x, \vec{\theta}) = g(x) \cdot \exp\left(\sum_{i=1}^n \theta_i t_i(x) - C(\vec{\theta})\right) \quad (\text{on } \{t_i(x)\}_{i=1}^n \text{ són funcions l.i.})$$

La família es diu **natural** si  $\exists k$  tal que  $t_k(x) = x$ .

#### Exemple 5.2.2

1.  $X \sim \text{Ber}(p)$

$$p(x | p) = p^x (1-p)^{1-x} = \exp\left(x \cdot \log(p) + (1-x) \cdot \log(1-p)\right), \quad x \in \{0, 1\}$$

Agafem l'exponent:

$$x \cdot \log(p) + (1-x) \cdot \log(1-p) = x \left( \log(p) - \log(1-p) \right) = x \underbrace{\left( \log\left(\frac{p}{1-p}\right) \right)}_{\theta} + \log(1-p)$$

$$\text{Si fem } \theta = \log\left(\frac{p}{1-p}\right) \implies e^\theta = \frac{p}{1-p} \implies e^\theta - pe^\theta = p \implies p = \frac{e^\theta}{1+e^\theta}$$

$$\implies \log(1-p) = \log\left(1 - \frac{e^\theta}{1+e^\theta}\right) = -\log(1+e^\theta)$$

Per tant,

$$p(x | \theta) = \exp\left(\underbrace{x \cdot \theta - \log(1+e^\theta)}_{C(\theta)}\right) \quad \left(\theta = \log \frac{p}{1-p}\right)$$

és família exponencial natural.

2.  $X \sim \text{Bin}(N, p)$

$$p(x | N, p) = \binom{N}{x} \cdot p^x (1-p)^{N-x} \quad x \in \{0, 1, \dots, N\}$$

$N$  no pot ser un paràmetre!  $\implies N$  ha de ser fix.

En aquesta situació:

$$\begin{aligned} p(x | p) &= \underbrace{\binom{N}{x}}_{g(x)} p^x (1-p)^{N-x} = g_N(x) \cdot \exp\left(x \cdot \log(p) + (N-x) \cdot \log(1-p)\right) = \\ &= g_N(x) \cdot \exp\left(x \cdot \log\left(\frac{p}{1-p}\right) + N \cdot \log(1-p)\right) \stackrel{\theta = \log \frac{p}{1-p}}{=} g_N(x) \cdot \exp\left(x \cdot \theta - N \cdot \underbrace{\log(1+e^\theta)}_{C(\theta)}\right) \end{aligned}$$

3.  $X \sim N(\mu, \sigma^2)$

$$p(x | \mu, \theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right) = \frac{1}{\sqrt{2\pi}} \cdot \exp\left(\frac{-(x-\mu)^2}{2\sigma^2} - \log(\sigma)\right)$$

Desenvolupem l'exponencial:

$$\begin{aligned} \frac{-(x-\mu)^2}{2\sigma^2} - \log(\sigma) &= \frac{-x^2}{2\sigma^2} + \frac{2x\mu}{2\sigma^2} - \frac{\mu^2}{2\sigma^2} - \log(\sigma) = \\ &= \frac{-x^2}{2\sigma^2} + \frac{\mu}{\sigma^2}x - \left(\frac{\mu^2}{2\sigma^2} + \log(\sigma)\right) = \\ &= \underbrace{\frac{\mu}{\sigma^2}}_{\theta_1} x - \underbrace{\frac{1}{2\sigma^2}}_{\theta_2} x^2 - \left(\frac{\mu^2}{2\sigma^2} + \log(\sigma)\right) = \\ &= \theta_1 \cdot \underbrace{x}_{t_1(x)} + \theta_2 \cdot \underbrace{x^2}_{t_2(x)} - \underbrace{\left(\frac{-\theta_1^2}{2\theta_2} + \frac{1}{2} \cdot \log\left(\frac{-1}{2\theta_2}\right)\right)}_{C(\theta_1, \theta_2)} \end{aligned}$$

Estudiant la fórmula  $p(x \mid \vec{\theta})$  directament, podem trobar resultats genèrics de manera senzilla.

**Proposició 5.2.3**

$$\mathbb{E}[t_i(Z)] = \frac{\partial}{\partial \theta_i} C(\vec{\theta})$$

**Proposició 5.2.4**

Si  $X$  pertany a una família exponencial natural amb  $t_1(x) = x$ , aleshores:

$$\Phi_X(t) = e^{C(\theta_1 + it, \theta_2, \dots, \theta_r) - C(\vec{\theta})}$$

## 6 Convergència de variables aleatòries

### 6.1 Modes de convergència i equivalències

Sigui  $\{X_n\}_{n \geq 1}$  una seqüència de variables aleatòries. Volem donar sentit a la noció de convergència de  $\{X_n\}_{n \geq 1}$  cap a una variable aleatòria  $X$  donada.

Veurem que això ho podem fer de diverses maneres.

#### Definició 6.1.1

$\{X_n\}_{n \geq 1}$  **convergeix quasi-segurament** cap a  $X$ , i ho escriurem  $X_n \xrightarrow{qs} X$  si

$$p\left(\underbrace{\left\{\omega \in \Omega : X_n(\omega) \xrightarrow[n]{} X(\omega)\right\}}_A\right) = 1$$

La definició té sentit perquè  $A$  és un succés. Vegem-ho:

$$A_n(m) = \left\{\omega \in \Omega : |X_n(\omega) - X(\omega)| < \frac{1}{m}\right\} \text{ és un succés.}$$

$$\implies A(m) = \liminf_n A_n(m) = \{\omega \in \Omega : \omega \in A_n(m) \quad \forall n \geq n_0(\omega)\} \text{ (és un succés)}$$

Finalment,

$$A = \bigcap_{m \geq 1} A(m) = \left\{\omega \in \Omega : \forall m \exists n_0(\omega) \text{ tal que } |X_n(\omega) - X(\omega)| < \frac{1}{m} \text{ si } n \geq n_0(\omega)\right\} \text{ és un succés.}$$

#### Definició 6.1.2

Direm que  $\{X_n\}_{n \geq 1}$  **convergeix en mitjana d'ordre  $r$**  cap a  $X$  ( $r \geq 1$ ) i ho escriurem  $X_n \xrightarrow{r} X$  si

$$\mathbb{E}[|X_n - X|^r] \xrightarrow[n]{} 0$$

#### Definició 6.1.3

Direm que  $\{X_n\}_{n \geq 1}$  **convergeix en probabilitat** cap a  $X$  i ho escriurem  $X_n \xrightarrow{p} X$  si  $\forall \epsilon > 0$

$$p(|X_n - X| > \epsilon) \xrightarrow[n]{} 0$$

#### Definició 6.1.4

Direm que  $\{X_n\}_{n \geq 1}$  **convergeix en distribució** cap a  $X$  i ho escriurem  $X_n \xrightarrow{d} X$  si

$$F_{X_n}(x) \xrightarrow[n]{} F(x) \text{ en els punts de continuïtat de } F(x)$$

**Observació 6.1.5** Ens cal que  $x$  sigui un punt de continuïtat de  $F$  per a que famílies de variables aleatòries convergeixin de manera natural en distribució.

### Exemple 6.1.6

Sigui  $X$  una variable aleatòria fixada,  $X_n = X + \frac{1}{n}$ . Aleshores volem que  $X_n \xrightarrow{d} X$ . Si això passa,

$$F_{X_n}(x) = p(X_n \leq x) = p(X \leq x - \frac{1}{n}) \xrightarrow{n \rightarrow \infty} p(X < x) = F_X(x^-) \quad (\text{límit per l'esquerra})$$

$\implies$  per tal que  $F_X(x)$  sigui  $F_X(x^-)$  ens cal suposar que  $x$  és un punt de continuïtat de  $F_X$ .

El que veurem ara són implicacions entre els diversos modes de convergència. També veurem que les implicacions contràries mai seran certes.

#### 6.1.1 Diagrama modes de convergència

$$\begin{array}{ccccccc} X_n & \xrightarrow{qs} & X & \xlongequal{\text{IV}} & X_n & \xrightarrow{p} & X \xlongequal{\text{I}} X_n \xrightarrow{d} X \\ & & & & \nearrow \text{II} & & \\ X_n & \xrightarrow{r} & X & \xlongequal[r>s \geq 1]{\text{III}} & X_n & \xrightarrow{s} & X \end{array}$$

#### Proposició 6.1.7 (I)

$X_n \xrightarrow{p} X \implies X_n \xrightarrow{d} X$ , i el recíproc no és cert.

Vegem un contraexemple de que el recíproc no és cert:

$X \sim \text{Ber}(\frac{1}{2})$ ,  $X_n = X$ ,  $Y = 1 - X$  (En particular,  $Y \sim \text{Ber}(\frac{1}{2})$ ).

És clar que  $X_n \xrightarrow{d} Y$ , ja que  $X_n \xrightarrow{d} X$  i  $X$  i  $Y$  tenen la mateixa funció de distribució. Per altra banda,  $|X_n - Y| = 1$  (ja que si una val 1, l'altra val 0)

$$\implies p(|X_n - Y| > \varepsilon) = 1 \quad \text{si } \varepsilon \text{ és prou petit!!}$$

#### Proposició 6.1.8 (II)

$X_n \xrightarrow{1} X \implies X_n \xrightarrow{p} X$ , i el recíproc no és cert.

Vegem un contraexemple del recíproc:

$$X_n = \begin{cases} n^3 & \text{amb probabilitat } \frac{1}{n^2} \\ 0 & \text{amb probabilitat } 1 - \frac{1}{n^2} \end{cases}$$

El candidat a límit és  $X = 0$ :

$$\forall \varepsilon > 0, p(|X_n - X| > \varepsilon) = p(|X_n| > \varepsilon) = \frac{1}{n^2} \xrightarrow{n \rightarrow \infty} 0$$

Amb això,  $X_n \xrightarrow{p} 0 = X$ . Ara bé,

$$\mathbb{E}[|X_n - X|] = \mathbb{E}[|X_n|] = n^3 \cdot \frac{1}{n^2} = n \xrightarrow{n \rightarrow \infty} \infty \implies X_n \not\xrightarrow{1} 0$$

**Proposició 6.1.9 (III)**

Si  $r \geq s \geq 1$  i  $X_n \xrightarrow{r} X \implies X_n \xrightarrow{s} X$ .

El recíproc no és cert:

$$X_n = \begin{cases} n & \text{amb probabilitat } n^{-\frac{(r+s)}{2}} \\ 0 & \text{amb probabilitat } 1 - n^{-\frac{(r+s)}{2}} \end{cases}$$

Prendrem com a candidat a límit  $X = 0$ .

$$\mathbb{E}[|X_n|^s] = n^s \cdot n^{-\frac{(r+s)}{2}} + 0 \cdot (1 - n^{-\frac{(r+s)}{2}}) = n^{\frac{s-r}{2}} \xrightarrow{n \rightarrow \infty} 0 \text{ (perquè } s \leq r)$$

$$\mathbb{E}[|X_n|^r] = n^r \cdot n^{-\frac{(r+s)}{2}} + 0 \cdot (1 - n^{-\frac{(r+s)}{2}}) = n^{\frac{r-s}{2}} \xrightarrow{n \rightarrow \infty} +\infty \neq 0 !!$$

**Lema 6.1.10**

Sigui  $\varepsilon > 0$ . Definim

$$A_n(\varepsilon) = \{\omega \in \Omega : |X_n(\omega) - X(\omega)| > \varepsilon\}$$

i

$$B_n(\varepsilon) = \bigcup_{m \geq n} A_m(\varepsilon)$$

Aleshores,

$$X_n \xrightarrow{q.s.} X \iff \forall \varepsilon > 0, \lim_n p(B_n(\varepsilon)) = 0$$

Com a conseqüència, tenim el següent corol·lari:

**Corol·lari 6.1.11**

Amb la notació anterior,

$$\forall \varepsilon > 0, \sum_{n \geq 1} p(A_n(\varepsilon)) \implies X_n \xrightarrow{q.s.} X$$

**Proposició 6.1.12 (IV)**

$$X_n \xrightarrow{q.s.} X \implies X_n \xrightarrow{p} X.$$

El recíproc no és cert:

$$X_n = \begin{cases} 1 & \text{amb probabilitat } \frac{1}{n} \\ 0 & \text{amb probabilitat } 1 - \frac{1}{n} \end{cases}$$

Candidat a límit:  $X = 0$

Convergeix en probabilitat:

$$\forall \varepsilon > 0, p(|X_n - X| > \varepsilon) = p(|X_n| > \varepsilon) \underset{\text{si } \varepsilon < 1}{=} \frac{1}{n} \xrightarrow{n \rightarrow \infty} 0$$

Vegem ara que no es compleix que  $X_n \xrightarrow{q.s.} 0$ . Ho veurem calculant  $\lim_n p(B_n(\varepsilon))$ .

Calculem-ho:

Segui  $A_n(\varepsilon) = \{\omega \in \Omega : |X_n(\omega) - X(\omega)| > \varepsilon\}$

$$\begin{aligned} p(B_n(\varepsilon)) &= p\left(\bigcup_{m \geq n} A_m(\varepsilon)\right) = 1 - p\left(\bigcap_{m \geq n} \overline{A_m(\varepsilon)}\right) \underset{\substack{A_n(\varepsilon) \\ \text{indep.}}}{=} 1 - \prod_{m \geq n} p\left(\overline{A_m(\varepsilon)}\right) \underset{\text{si } \varepsilon < 1}{=} \\ &= 1 - \prod_{m \geq n} \left(1 - \frac{1}{m}\right) = 0 \end{aligned}$$

Vegem ara que  $\lim_{n \rightarrow \infty} \prod_{m \geq n} \left(1 - \frac{1}{m}\right) = 0$

$$0 \leq \prod_{m \geq n} \left(1 - \frac{1}{m}\right) \underset{1-x \leq e^{-x}}{\leq} \exp\left(\sum_{m \geq n} \frac{-1}{m}\right) \underset{\forall n}{=} 0$$

Per tant,  $\forall n, \prod_{m \geq n} \left(1 - \frac{1}{m}\right) = a_n = 0 \implies \lim_{n \rightarrow \infty} a_n = 0 \implies p(B_n(\varepsilon)) \xrightarrow{n} 1$

Finalment, veurem una implicació en sentit contrari:

**Proposició 6.1.13**

Si  $X_n \xrightarrow{p} X$ , aleshores existeix una subsuccessió  $\{n_i\}_{i \geq 1}$  per la qual  $Z_{n_i} \xrightarrow{q.s.} X$ .

## 6.2 Convergència quasi-segura i la llei forta dels grans nombres

Comencem amb les propietats bàsiques de la convergència quasi-segura:

### Proposició 6.2.1

Si  $X_n \xrightarrow{q.s.} X$  i  $Y_n \xrightarrow{q.s.} Y$ , aleshores:

- a)  $X_n + Y_n \xrightarrow{q.s.} X + Y$
- b)  $X - n \cdot Y_n \xrightarrow{q.s.} X \cdot Y$
- c)  $c \cdot X_n \xrightarrow{q.s.} c \cdot X$  ( $c \in \mathbb{R}$ )
- d) Si  $g$  contínua,  $g(X_n) \xrightarrow{q.s.} g(X)$

El resultat més important en aquest mode de convergència és la llei forta dels grans nombres:

### Teorema (Llei forta dels grans nombres) (6.2.2)

Sigui  $\{X_n\}_{n \geq 1}$  una successió de variables aleatòries independents i idènticament distribuïdes (i.i.d.) amb  $\bar{X}$ .

A més,  $\mathbb{E}[|X|] < +\infty$ ,  $\mathbb{E}[X] = \mu$ . Aleshores, si  $S_n = X_1 + \dots + X_n$ :

$$\frac{S_n}{n} \xrightarrow{q.s.} \mu$$

**Observació 6.2.3** Sota les mateixes condicions,  $\frac{S_n}{n} \xrightarrow{p} \mathbb{E}[X] \equiv$  *Llei feble dels grans nombres*

**Observació 6.2.4** Sota les mateixes condicions (i també  $\text{Var}[X] < +\infty$ ), podem afirmar convergència en mitjana d'ordre 2 (o mitjana quadràtica):  $\frac{S_n}{n} \xrightarrow{2} \mu$

## 6.3 Convergència en distribució i el teorema del límit central