TEORIA DE LA PROBABILITAT

Apunts basats en el curs impartit pel Juanjo Rué

ApuntsFME

Barcelona, Gener 2019

Autors: Jordi Castellví, Ferran López, Miquel Ortega, Èric Sierra.

Revisors: Iñaki Garrido.

Última modificació: 6 d'octubre de 2020.

Índex

1 Espai de probabilitat **1**

1.1 Definició axiomàtica de probabilitat . . . 1

Desigualtats de Bonferroni . . . 2

1.2 Probabilitat condicionada i independència . . . 4

1.3 Espai producte . . . 5

1.4 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 Txebixov . . . 13

Desigualtat de Markov . . . 15

Desigualtat de Txebixov . . . 15

2.3 Vectors de variables aleatòries. Independència de variables aleatòries . . . 16

3 Variables aleatòries discretes **21**

3.1 Definició i objectes relacionats . . . 21

3.2 Funció generadora de probabilitat . . . 23

3.3 Models de variables aleatòries discretes . . . 25

3.4 Distribucions condicionades i esperança condicionada . . . 29

3.5 Arbres de Galton-Watson . . . 31

4 Variables aleatòries contínues **35**

4.1 Mesures de probabilitat absolutament contínues. Funció de densitat . . . 35

Teorema de Radon-Nikodym . . . 35

4.2 Models de variables aleatòries contínues . . . 37

4.3 Conceptes de variables aleatòries contínues . . . 38

4.3.1 Distribució conjunta i marginals . . . 38

4.3.2 Independència . . . 40

4.3.3 Distribucions condicionades . . . 40

4.4 Distribució normal, multivariant i distribucions associades . . . 41

Teorema de Moivre-Laplace . . . 41

4.4.1 Distribució normal multivariant . . . 42

4.4.2 Estimadors i teorema de Fisher . . . 43

Teorema de Fisher . . . 44

4.5 Funcions de variables aleatòries absolutament contínues . . . 44

4.5.1 Aplicacions . . . 45

5 Funcions característiques i famílies exponencials **47**
5.1 Funció generadora de moments. Propietats i aplicació: fites de Chernoff . 47
5.1.1 Aplicació. Teorema de Chernoff 48
Teorema de Chernoff 48
5.2 Funcions característiques 50
Teorema d'inversió 51
5.2.1 Propietats addicionals de la funció característica 53
5.3 Famílies exponencials: exemples 54

6 Convergència de variables aleatòries **57**
6.1 Modes de convergència. Implicacions 57
6.2 Convergència quasi-segura. Llei forta dels grans nombres 61
Llei forta dels grans nombres 62
6.3 Convergència en distribució. Teorema del límit central 64
Teorema de representació de Skorokhod 65
Teorema de Lévy 67
Llei forta dels grans nombres (en distribució) 67
Teorema del límit central 67
6.4 Aplicacions 68
6.4.1 Conjunts lliures de sumes 68
6.4.2 Funcions de representació en els naturals 69

Índex alfabètic **71**

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

i) $\varnothing \in \mathcal{A}$,

ii) $A \in \mathcal{A} \iff \overline{A} \in \mathcal{A}$,

iii) Si $\{A_i\}_{i\in\mathbb{N}} \subseteq \mathcal{A}$, aleshores $\bigcup\limits_{i\in\mathbb{N}} A_i \in \mathcal{A}$.

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

i) $\mu(\varnothing) = 0$,

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 $\forall i \neq j$, $A_i \cap A_j = \varnothing$, aleshores

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

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

i) Si $A_1, \ldots, A_r \in \mathcal{A}$ t. q. $\forall i \neq j$, $A_i \cap A_j = \varnothing$, aleshores $p\left(\bigcup\limits_{i=1}^{r} A_i\right) = \sum\limits_{i=1}^{r} p\left(A_i\right)$.

ii) $A \in \mathcal{A} \implies p(\overline{A}) = 1 - p(A)$.

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 $p\left(\bigcup_{i\in\mathbb{N}} A_i\right) = \lim_{i\to\infty} p(A_i)$.

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

Demostració.

i) Conseqüència directa de la σ-additivitat.

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

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

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

v)

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

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

$$= \lim_{N\to\infty} \sum_{i=0}^{N} p(B_i) = \lim_{N\to\infty} p\left(\bigcup_{i=0}^{N} B_i\right) = \lim_{N\to\infty} p(A_N).$$

b) Anàleg al cas anterior.

□

Observem que l'apartat v) només es pot aplicar en casos molt particulars. En general, si tenim $A_i, \ldots, A_r$ succcessos, 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 $A_1, \ldots, A_r \in \mathcal{A}$, i per $I \subseteq \{1, \ldots, r\}$ sigui $A_I = \bigcap_{i\in I} A_i$. Definim

$$S_k = \sum_{\forall I | \#I = k} p(A_I),$$

això és, $S_1 = \sum p(A_i)$, $S_2 = \sum_{i\neq j} p\left(A_i \cap A_j\right), \ldots$ 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 $(\Omega, \mathcal{A})$ 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

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

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

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

   Sigui $\Omega = [a, b]$ i prenem $\mathcal{A} = B \cap [a, b]$ amb $B \in \mathcal{B}$ 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

   $$\begin{gathered} 00010001110110\ldots \\ 01001110101101\ldots \\ 10010111110010\ldots \end{gathered}$$

   sent 0 creu i 1 cara. Aquest conjunt és no numerable, cosa 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

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

   on $0.a_1a_2\ldots$ 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 són tots racionals, per això 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 $\left\{\varphi^{-1}(A)\right\}_{A\subseteq \mathcal{B}\cup[0,1]}$. Similarment ho fem amb la mesura.

1.2 Probabilitat condicionada i independència

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

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

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

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

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

1) $p(A_i) > 0$,

2) $i \neq j \implies A_i \cap A_j = \varnothing$,

3) $\bigcup\limits_{i\in I} A_i = \Omega$.

Aleshores,

i) Probabilitat total:

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

ii) Fórmula de Bayes:

$$p(A_i \mid B) = \frac{p(B \mid A_i)\, p(A_i)}{\sum\limits_{j\in I} p\left(B \mid A_j\right) p\left(A_j\right)}, \quad \forall B \in \mathcal{A} \text{ amb } p(B) > 0.$$

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

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

2)

$$p(A_i|B) \sum_{j\in I} p(B|A_j)p(A_j) \overset{i)}{=} p(A_i|B)p(B) =$$
$$\frac{p(B \cap A_i)}{p(B)}p(B) = p(B \cap A_i) = p(B \mid A_i)\, p(A_i).$$

$\square$

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,

$$p(A_k) = p(A_k|B)p(B) + p(A_k|\overline{B})p(\overline{B}) = \frac{p(A_k|B)}{2} + \frac{p(A_k|\overline{B})}{2}$$
$$\implies 2p(A_k) = p(A_{k-1}) + p(A_{k+1}) \implies p(A_k) - p(A_{k-1}) = p(A_{k+1}) - p(A_k) = C,$$

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

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

Definició 1.2.5. Sigui $(\Omega, \mathcal{A}, p)$ 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

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

Exemple 1.2.6.

1. $\varnothing, \Omega$ 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 interseccións dos a dos compleixen que la probabilitat de la intersecció és el producte de probabilitats.
4. A i B independents $\iff$ A i $\overline{B}$ independents, ja que

$$\begin{aligned} p(A \cap \overline{B}) &= p(A) - p(A \cap B) = p(A) - p(A)\, p(B) \\ &= p(A)\,(1 - p(B)) = p(A)\, p(\overline{B}) \end{aligned}$$

1.3 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 1.3.1. 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 1.3.2. 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 1.3.3. Podem estendre λ (la mesura de Lebesgue) a $\mathbb{R}^2$ de la següent forma. Sabem que $\left([0,1], \mathcal{B} \cap [0,1], \lambda_{[0,1]}\right)$ és un espai de probabilitat. Aleshores

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

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

Problema 1.3.4. *Agulla de Buffon.* Considerem el pla $\mathbb{R}^2$ tesel·lat amb linies 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 linies.

Solució. Considerarem dues variables: x com la distància del centre de l'agulla a la linia més propera i θ com l'angle de l'agulla amb la direcció de les lines. Tenim que $x \in \left[0, \frac{L}{2}\right]$ i $\theta \in [0, \pi)$ i per tant, $\Omega = \left[0, \frac{L}{2}\right] \times [0, \pi)$, $\mathcal{A}$ 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 $w \in A \iff x \leq \frac{l}{2} \sin\theta$. Per tant,

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

1.4 Lema de Borel-Cantelli

Siguin $(\Omega, \mathcal{A}, p)$ 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 1.4.1. Sigui $(\Omega, \mathcal{A}, p)$ 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

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

Observació 1.4.2. Els dos límits pertanyen a $\mathcal{A}$ ja que són unió i intersecció numerable de sucessos.

Proposició 1.4.3. Sigui $(\Omega, \mathcal{A}, p)$ un espai de probabilitat i siguin $\{A_n\}_{n\geq 1} \subseteq \mathcal{A}$. Aleshores,

i) $\liminf_{n\to\infty} A_n = \{\omega \in \Omega \colon \exists m \equiv m(\omega) \text{ amb } \omega \in A_r \ \forall r \geq m(\omega)\}$,

ii) $\limsup_{n\to\infty} A_n = \{\omega \in \Omega \colon \omega \text{ pertany a un nombre infinit dels } A_n\}$,

iii) $\liminf\limits_{n\to\infty} A_n \subseteq \limsup\limits_{n\to\infty} A_n$.

Demostració.

i) $\omega \in \liminf A_n \iff \omega \in \bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty} A_k \iff \exists m \equiv m(\omega)$ t. q. $\omega \in \bigcap_{k=m(\omega)}^{\infty} A_k$, és a dir, si $\exists m \equiv m(\omega)$ t. q. $\omega \in A_r \quad \forall r \geq m(\omega)$.

ii) $\omega \in \limsup A_n \iff \omega \in \bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} A_k \iff \omega \in \bigcup_{k=n}^{\infty} A_k \quad \forall n \iff \forall n, \exists n_0 \geq n$ t. q. $\omega \in A_{n_0} \iff \omega$ 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 1.4.4. Sigui $(\Omega, \mathcal{A}, p)$ 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 $B_n = \bigcup\limits_{k\geq n} A_k$ i $C_n = \bigcap\limits_{k\geq n} A_k$. Observem que $\{B_n\}_{n\geq 1}$ és decreixent i que $\{C_n\}_{n\geq 1}$ és creixent. Naturalment tenim que, $C_n \subseteq A_n \subseteq B_n$, $\limsup A_n = \bigcap\limits_{n\geq 1} B_n$ i $\liminf A_n = \bigcup\limits_{n\geq 1} C_n$.
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 $p\left(\bigcap\limits_{k\geq n} A_k\right) \leq p(A_n)$. 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 1.4.5. *Lema de Borel-Cantelli.*
Sigui $(\Omega, \mathcal{A}, p)$ un espai de probabilitat i siguin $\{A_n\}_{n\geq 1} \subseteq \mathcal{A}$. Aleshores,

i) $\sum\limits_{n\geq 1} p(A_n) < \infty \implies p(\limsup A_n) = 0$.

ii) Si $\{A_n\}_{n\geq 1}$ és independent, $\sum\limits_{n\geq 1} p(A_n) = \infty \implies p(\limsup A_n) = 1$.

Demostració. Anomenem $A = \limsup A_n = \bigcap\limits_{n\geq 1}\bigcup\limits_{k\geq n} A_k$.

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

ii) Observem primer que $\overline{A} = \overline{\bigcap_{n\geq 1} \bigcup_{k\geq n} A_k} = \bigcup_{n\geq 1} \bigcap_{k\geq n} \overline{A}_k = \liminf \overline{A}_n$. Veurem que $p(\overline{A}) = 0$. Calculem $p\left(\bigcap_{m\geq n} \overline{A}_m\right)$.

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

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

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

i concloem que $p\left(A\right) = 1$.

$\square$

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 $X\colon \Omega_1 \to \Omega_2$ és una variable aleatòria si

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

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

Exemple 2.1.2.

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

$$\begin{aligned} X\colon \Omega &\to \mathbb{R} \\ \omega &\mapsto c \end{aligned}$$

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

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

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 $g\colon \mathbb{R}^2 \to \mathbb{R}$ és una funció mesurable.

3. Sigui $(\Omega, \mathcal{A})$ un espai de mesura i sigui $A \in \mathcal{A}$. Definim la variable aleatòria indicadora d'A com

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

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

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

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

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

Exemple 2.1.4. Per a considerar la probabilitat de que X sigui més petit o igual que 2 escriurem

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

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

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

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

i) $P_X(\varnothing) = p\left(\left\{\omega \in \Omega \mid \omega \in X^{-1}(\varnothing)\right\}\right) = p(\varnothing) = 0$, atès que p és una funció de probabilitat.

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

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

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

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

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

Observació 2.1.6. Sigui $(\Omega, \mathcal{A})$ un espai mesurable. Recordem que $X\colon \Omega \to \mathbb{R}$ 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 $(\Omega, \mathcal{A}, p)$ un espai de probabilitat i sigui X una variable aleatòria. Anomenem funció de distribució de probabilitat d'X a l'aplicació

$$\begin{aligned} F_X\colon \mathbb{R} &\to [0,1] \\ x &\mapsto F_X(x) = p(X \le x) = P_X((-\infty, x]) . \end{aligned}$$

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

i) $x_1 \le x_2 \implies F_X(x_1) \le F_X(x_2)$.

ii) $\lim\limits_{x\to-\infty} F_X(x) = 0$ i $\lim\limits_{x\to\infty} F_X(x) = 1$.

iii) F_X és contínua per la dreta, és a dir, $\lim\limits_{h\to 0^+} F_X(x+h) = F_X(x)$.

Demostració.

i) $F_X(x_1) = p\Big(\{\omega \in \Omega \mid X(\omega) \le x_1\}\Big) \le p\Big(\{\omega \in \Omega \mid X(\omega) \le x_2\}\Big) = F_X(x_2)$, atès que $\{\omega \in \Omega \mid X(\omega) \le x_1\} \subseteq \{\omega \in \Omega \mid X(\omega) \le x_2\}$ i que p és una mesura.

ii) Vegem que $\forall \{x_n\}_{n\in\mathbb{N}}$ tal que $\lim\limits_{n\to\infty} x_n = -\infty$, es té que $\lim\limits_{n\to\infty} F_X(x_n) = 0$. Definim $A_n = \{\omega \in \Omega \mid X(\omega) \le x_n\}$. Tenim que $\varnothing \subseteq \liminf A_n \subseteq \limsup A_n$. A més, $\limsup A_n = \varnothing$ perquè, altrament, hi hauria un nombre infinit de conjunts A_n contenint un $\omega \in \Omega$ determinat. Per tant,

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

Anàlogament, es demostra que $\lim\limits_{x\to\infty} F_X(x) = 1$.

iii) Fixat x, volem veure que $\lim\limits_{h\to 0^+} F_X(x+h) = F_X(x)$.

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

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

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

$\square$

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

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

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

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

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

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 $F \colon \mathbb{R} \to [0,1]$ una funció de probabilitat tal que

i) $x_1 \leq x_2 \implies F(x_1) \leq F(x_2)$.

ii) $\lim_{x \to -\infty} F(x) = 0$ i $\lim_{x \to \infty} F(x) = 1$.

iii) F és contínua per la dreta, és a dir, $\lim_{h \to 0^+} F(x+h) = F(x)$.

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

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

$$\begin{aligned} X \colon\ [0,1] &\to [0,1] \\ \omega &\mapsto X(\omega) = \sup\{y \in \mathbb{R} \mid F(y) \leq \omega\}. \end{aligned}$$

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

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

i observem que

$$\begin{aligned} p(A) &= p(X \leq x) = F_X(x), \\ p(B) &= \lambda\Big([0, F(x)]\Big) = F(x). \end{aligned}$$

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

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

Definició 2.2.1. Sigui $(\Omega, \mathcal{A}, p)$ un espai de probabilitat i sigui $X\colon \Omega \to \mathbb{R}$ 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

$$\mathbb{E}[X] = \int_\Omega X \,\mathrm{d}p = \int_\mathbb{R} x \,\mathrm{d}P_X,$$

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 $|\int_\Omega f \,\mathrm{d}p| \leq \int_\Omega |f| \,\mathrm{d}p$, sovint demanarem que $\mathbb{E}[|X|] \leq +\infty$ per poder afirmar que $\mathbb{E}[X] \leq +\infty$.

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

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

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 matex.

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

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

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

$$\mathbb{E}[f(X)] = \int_\Omega f(X) \,\mathrm{d}p = \int_\mathbb{R} x \,\mathrm{d}P_{f(X)}.$$

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

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

- Moment d'ordre r d'X:
$$\mathbb{E}\left[X^r\right],$$
on $r \in \mathbb{R}$ i hem suposat que $\mathbb{E}\left[|X|^r\right] < +\infty$.

- Moment factorial d'ordre r d'X:
$$\mathbb{E}\left[(X)_r\right] = \mathbb{E}\left[X(X-1)\cdots(X-r+1)\right],$$
on $r \in \mathbb{N}$.

- Variància d'X:
$$\mathbb{V}\mathrm{ar}\left[X\right] = \mathbb{E}\left[(X - \mathbb{E}\left[X\right])^2\right].$$

- Desviació típica d'X:
$$\sigma(X) = \sqrt{\mathbb{V}\mathrm{ar}\left[X\right]}.$$

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.

i) $\mathbb{E}\left[a\right] = a$,

ii) $\mathbb{E}\left[aX + bY\right] = a\mathbb{E}\left[X\right] + b\mathbb{E}\left[Y\right]$,

iii) $\mathbb{E}\left[\mathbb{I}_A\right] = p(A)$,

iv) $|\mathbb{E}\left[X\right]| \leq \mathbb{E}\left[|X|\right]$.

Demostració. La demostració d'aquestes propietats es deixa com a exercici ja que es dedueixen directament de la definició o d'altres propietats bàsiques. □

Proposició 2.2.8. Siguin X, Y variables aleatòries, siguin $a, b \in \mathbb{R}$ i sigui $A \in \mathcal{A}$. Es tenen les següents propietats de la variància.

i) $\mathbb{V}\mathrm{ar}\left[X\right] = \mathbb{E}\left[(X - \mathbb{E}\left[X\right])^2\right] = \mathbb{E}\left[X^2 + \mathbb{E}\left[X\right]^2 - 2X\mathbb{E}\left[X\right]\right] = \mathbb{E}\left[X^2\right] + \mathbb{E}\left[X\right]^2 - 2\mathbb{E}\left[X\right]^2 = \mathbb{E}\left[X^2\right] - \mathbb{E}\left[X\right]^2$,

ii) $\mathbb{V}\mathrm{ar}\left[a\right] = 0$,

iii) $\mathbb{V}\mathrm{ar}\left[a + X\right] = \mathbb{V}\mathrm{ar}\left[X\right]$,

iv) $\mathbb{V}\mathrm{ar}\left[aX\right] = a^2\,\mathbb{V}\mathrm{ar}\left[X\right]$.

Demostració. La demostració d'aquestes propietats es deixa com a exercici ja que es dedueixen directament de la definició o d'altres propietats bàsiques. □

Proposició 2.2.9. *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 $\mathbb{E}\left[|X|^p\right], \mathbb{E}\left[|Y|^q\right] < +\infty$, aleshores

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

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

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

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

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

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

Observació 2.2.10. La desigualtat de Cauchy-Schwarz és el cas particular $p = q = 2$ de la desigualtat de Holder.

Teorema 2.2.11. *Desigualtat de Markov.*
Sigui $(\Omega, \mathcal{A}, p)$ un espai de probabilitat, sigui $X \colon \Omega \to \mathbb{R}$ una variable aleatòria amb $X > 0$ i sigui $a \in \mathbb{R}^+$. Aleshores,

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

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,

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

□

Teorema 2.2.12. *Desigualtat de Txebixov.*
Sigui X una variable aleatòria amb $\mathbb{E}\left[X\right] < +\infty$, $\mathbb{V}\text{ar}\left[X\right] < +\infty$ i $\mathbb{V}\text{ar}\left[X\right] \neq 0$. Aleshores, per tot $k > 0$, es té que

$$p\left(|X - \mathbb{E}\left[X\right]| \geq k\,\mathbb{V}\text{ar}\left[X\right]^{\frac{1}{2}}\right) \leq \frac{1}{k^2}.$$

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

$$\begin{aligned} p\left(Y \geq a\right) &= p\left(|X - \mathbb{E}\left[X\right]| \geq a\right) = \\ &= p\left(\left(X - \mathbb{E}\left[X\right]\right)^2 \geq a^2\right) \leq \\ &\leq \frac{\mathbb{E}\left[\left(X - \mathbb{E}\left[X\right]\right)^2\right]}{a^2} = \\ &= \frac{\mathbb{V}\text{ar}\left[X\right]}{a^2}, \end{aligned}$$

on hem aplicat la desigualtat de Markov. Prenent ara $a = k\,\mathbb{V}\text{ar}\left[X\right]^{\frac{1}{2}}$, deduïm el resultat volgut. □

Observació 2.2.13. Quan $\mathbb{V}\mathrm{ar}\left[X\right]=0$, seguim tenint que, per tot $a>0$,

$$0\leq p\Big(\left|X-\mathbb{E}\left[X\right]\right|\geq a\Big)\leq\frac{\mathbb{V}\mathrm{ar}\left[X\right]}{a^2}=0,$$

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

2.3 Vectors de variables aleatòries. Independència de variables aleatòries

Definició 2.3.1. Sigui $(\Omega,\mathcal{A},p)$ un espai de probabilitat. Anomenem un vector de variables aleatòries o una variable aleatòria multidimensional de dimensió n a una funció mesurable $\vec{X}\colon\Omega\to\mathbb{R}^n$.

Observació 2.3.2. La funció

$$\begin{aligned}\pi_i\colon\mathbb{R}^n&\to\mathbb{R}\\(x_1,\cdots,x_n)&\mapsto x_i\end{aligned}$$

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

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

$$F_{\vec{X}}\left(x_1,\ldots,x_n\right)=p\left(X_1\leq x_1,\ldots,X_n\leq x_n\right).$$

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

i) $\displaystyle\lim_{(x,y)\to(-\infty,-\infty)}F_{\vec{X}}\left(x,y\right)=0$ i $\displaystyle\lim_{(x,y)\to(+\infty,+\infty)}F_{\vec{X}}\left(x,y\right)=1$.

ii) $x_1\leq x_2\implies F_{\vec{X}}\left(x_1,y\right)\leq F_{\vec{X}}\left(x_2,y\right),\ \forall y\in\mathbb{R}$.

iii) $\displaystyle\lim_{(x,y)\to\left(x_0^+,y_0^+\right)}F_{\vec{X}}\left(x,y\right)=F_{\vec{X}}\left(x_0,y_0\right)$.

iv) $\displaystyle\lim_{y\to+\infty}F_{\vec{X}}\left(x,y\right)=F_X\left(x\right)$ i les anomenem distribucions marginals.

v) $p\left(a<x\leq b,c<y\leq d\right)=F_{\vec{X}}\left(b,d\right)-F_{\vec{X}}\left(b,c\right)-F_{\vec{X}}\left(a,d\right)+F_{\vec{X}}\left(a,c\right)=\Delta_RF$,

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. □

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

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

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

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

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

En general, no podem conèixer la distribució de $X_1, \ldots, X_n$ 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 independents $\{X_i\}_{i=1}^n$ (i posant $\vec{X} = (X_1, \ldots, X_n)$), tenim que

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

Demostració. Directament a partir de la definició tenim que

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

□

Observació 2.3.8. Es pot demostrar que la implicació 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 $f_1, \ldots, f_n \colon \mathbb{R} \to \mathbb{R}$ funcions mesurables. Aleshores, $\{f_i(X_i)\}_{i=1}^n$ també són independents.

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

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

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

$$\mathbb{C}\mathrm{ov}(X, Y) = \mathbb{E}\left[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])\right]$$

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

$$\mathbb{C}\mathrm{ov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

Demostració. Directament de la definició de covariància i usant les propietats de l'esperança

$$\begin{aligned} \mathbb{C}\mathrm{ov}(X, Y) &= \mathbb{E}\left[XY - \mathbb{E}[X]Y - \mathbb{E}[Y]X + \mathbb{E}[X]\mathbb{E}[Y]\right] = \\ &= \mathbb{E}[XY] - 2\mathbb{E}[X]\mathbb{E}[Y] + \mathbb{E}[X]\mathbb{E}[Y] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]. \end{aligned}$$

□

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

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

Demostració. Usant la definició de variància tenim que

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

□

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

i) $\mathbb{C}\text{ov}(C,X)=0$.

ii) $\mathbb{C}\text{ov}(X,X)=\mathbb{V}\text{ar}[X]$.

iii) $\mathbb{C}\text{ov}(X,Y)=\mathbb{C}\text{ov}(Y,X)$.

iv) $\mathbb{C}\text{ov}(aX+bY,Z)=a\,\mathbb{C}\text{ov}(X,Z)+b\,\mathbb{C}\text{ov}(Y,Z)$.

v) $\mathbb{C}\text{ov}(X,Y)^2\le\mathbb{V}\text{ar}[X]\,\mathbb{V}\text{ar}[Y]$.

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

$$|\mathbb{C}\text{ov}(X,Y)|=\left|\mathbb{E}\left[(X-\mathbb{E}[X])\,(Y-\mathbb{E}[Y])\right]\right|\le\mathbb{E}\left[\left|(X-\mathbb{E}[X])\,(Y-\mathbb{E}[Y])\right|\right]$$

que, aplicant la desigualtat de Cauchy-Schwarz, és menor que

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

És a dir, $\mathbb{C}\text{ov}(X,Y)^2\le\mathbb{V}\text{ar}[X]\,\mathbb{V}\text{ar}[Y]$. □

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. $\mathbb{C}\text{ov}(X,Y)=0\;\not\Longrightarrow\;X,Y$ independents.

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

$$X=\begin{cases}1 & \text{amb probabilitat } \frac12\\ -1 & \text{amb probabilitat } \frac12\end{cases},\; Y=\begin{cases}0 & \text{si } X=-1\\ \pm1 & \text{amb probabilitat } \frac12 \text{ si } X=1\end{cases}.$$

En aquest cas,

$$\begin{aligned}\mathbb{E}[XY]&=\int_\Omega XY\,\mathrm{d}p=\\ &=0\cdot p(X=-1)+1\cdot p(X=1,Y=1)-1\cdot p(X=1,Y=-1)=\\ &=\frac14-\frac14=0,\end{aligned}$$

i també tenim que

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

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

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

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

Tema 3

Variables aleatòries discretes

3.1 Definició i objectes relacionats

Definició 3.1.1. Sigui $(\Omega, \mathcal{A}, p)$ un espai de probabilitat i sigui $X \colon \Omega \to \mathbb{R}$ una variable aleatòria. Diem que X és discreta si $\operatorname{Im}(X)$ és numerable.

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

Observació 3.1.3. Sigui X una variable aleatòria discreta i sigui $A \in \mathcal{B}$. Aleshores,

$$P_X(A) = p(X \in A) = \sum_{x_i \in A \cap \operatorname{Im}(X)} p_i.$$

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 $\{x_i\}$ tal que $p(X = x_i) = p_i \; \forall\, i$.

Observació 3.1.5. La funció de distribució, amb $|\operatorname{Im}(X)| < +\infty$, 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 una variable aleatòria discreta. Aleshores,

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

ii) Si $g \colon \mathbb{R} \to \mathbb{R}$ és una funció mesurable, $\mathbb{E}[g(x)] = \sum_{i \geq 1} g(x_i) p_i$.

Demostració.

i) Fent servir la definició d'esperança tenim que

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

perquè $P_X(\mathbb{R} \setminus \{x_i\}_{i \geq 1}) = 0$.

ii) Directe a partir del cas anterior i del fet que g és mesurable.

□

Observació 3.1.7. Sigui X una variable aleatòria. Aleshores,

$$\mathbb{V}\mathrm{ar}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \sum_{i\geq 1} x_i^2 p_i - \left(\sum_{i\geq 1} x_i p_i\right)^2.$$

Proposició 3.1.8. Siguin X, Y variables aleatòries discretes, amb $\mathrm{Im}(X) = \{x_i\}_{i\geq 1}$ i $\mathrm{Im}(Y) = \{y_i\}_{i\geq 1}$, i sigui $g\colon \mathbb{R}^2 \to \mathbb{R}$ mesurable. Aleshores

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

Demostració. Directe a partir de la definició i del fet que g és mesurable. □

Proposició 3.1.9. Siguin X, Y variables aleatòries discretes. Són independents si i només si $\forall x \in \mathrm{Im}(X)$ i $\forall y \in \mathrm{Im}(Y)$,

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

Demostració. Exercici. □

Definició 3.1.10. Sigui $(X,Y)\colon \Omega \to \mathbb{R}^2$ un vector de variables aleatòries. Direm que és discret si $\mathrm{Im}((X,Y))$ és numerable.

Observació 3.1.11. Sigui (X,Y) vector de variables aleatòries. Aleshores és discret si i només si X i Y són discretes.

Definició 3.1.12. Sigui (X,Y) vector de variables aleatòries discret. Definim

$$\begin{aligned} P_{(X,Y)}\colon \mathbb{R}^2 &\to \mathbb{R} \\ (x,y) &\mapsto P_{(X,Y)}(x,y) \equiv p(X = x)\, p(Y = y). \end{aligned}$$

Si X, Y són independents, $P_{(X,Y)}(x,y) = p(X = x, Y = y)$.

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

$$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y].$$

Demostració.

$$\begin{aligned} \mathbb{E}[XY] &= \sum_{a\in \mathrm{Im}(X)\cup \mathrm{Im}(Y)} ap\,(XY = a) = \\ &= \sum_{a\in \mathrm{Im}(X)\cup \mathrm{Im}(Y)} a \sum_{b\in \mathrm{Im}(X)\setminus\{0\}} p\left(X = b, Y = \frac{a}{b}\right) = \\ &= \sum_{a\in \mathrm{Im}(X)\cup \mathrm{Im}(Y)} a \sum_{b\in \mathrm{Im}(X)\setminus\{0\}} p\,(X = b)\, p\left(Y = \frac{a}{b}\right) = \\ &= \sum_{b\in \mathrm{Im}(X)\setminus\{0\}} \sum_{a\in \mathrm{Im}(X)\cup \mathrm{Im}(Y)} ap\left(Y = \frac{a}{b}\right) p\,(X = b) = \\ &= \sum_{b\in \mathrm{Im}(X)\setminus\{0\}} bp\,(X = b) \sum_{a\in \mathrm{Im}(X)\cup \mathrm{Im}(Y)} \frac{a}{b} p\left(Y = \frac{a}{b}\right) = \\ &= \sum_{b\in \mathrm{Im}(X)\setminus\{0\}} bp\,(X = b)\, \mathbb{E}\,[Y] = \\ &= \mathbb{E}\,[X]\, \mathbb{E}\,[Y]\,. \end{aligned}$$

□

3.2 Funció generadora de probabilitat

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

Definició 3.2.1. Definim la funció generadora de probabilitat d'X com la sèrie formal de potències

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

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

Demostració.

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

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

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) Directe a partir de la definició.

iii) Si derivem terme a terme obtenim

$$\frac{\mathrm{d}^k G_X(z)}{\mathrm{d}z} = \frac{\mathrm{d}^k}{\mathrm{d}z}\left(\sum_{n\geq 0} p(X=n) z^n\right) = \sum_{n\geq 0} n(n-1)\dots(n-k+1) p(X=n) z^{n-k},$$

que avaluat en $z=1$ és

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

$\square$

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,

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

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 $\mathrm{Im}(X)=\mathrm{Im}(Y)=\mathbb{N}_{\geq 0}$. 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 $\mathrm{Im}(X)=\mathrm{Im}(Y)=\mathbb{N}_{\geq 0}$ aleshores:

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

Demostració.

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

I, com X i 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 els 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 i Y són variables aleatòries discretes independents amb $\mathrm{Im}(X) = \mathrm{Im}(Y) = \mathbb{N}_{\geq 0}$ aleshores

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

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

$$G_{X_1+\cdots+X_n}(z) = \prod_{i=1}^{n} G_{X_i}(z).$$

3.3 Models de variables aleatòries discretes

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

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

Distribució de Bernoulli

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

Definició 3.3.2. Sigui X una variable aleatòria. Direm que X segueix una distribució de Bernoulli

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

També es pot escriure $\mathrm{Be}(p)$.

Proposició 3.3.3. Sigui X una variable aleatòria que segueix una distribució de Bernoulli. Aleshores,

i) $G_X(z) = (1-p)z^0 + pz^1 = (1-p) + pz$

ii) $\mathbb{E}[X] = p$

iii) $\mathbb{V}\mathrm{ar}[X] = p(1-p)$

Demostració.

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

□

Distribució binomial

Modela el nombre d'èxits en fer N experiments independents, on cadascun és $\mathrm{B}(p)$.

Definició 3.3.4. Sigui X una variable aleatòria. Direm que X segueix una distribució binomial

$$X \sim \mathrm{Bin}(N,p) \iff X = X_1 + \cdots + X_N,$$

on $\{X_i\}_{i=1}^N$ són independents i $X_i \sim \mathrm{B}(p)\ \forall i$.

Proposició 3.3.5. Sigui X una variable aleatòria que segueix una distribució binomial. Aleshores,

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) $\mathbb{V}\mathrm{ar}[X]=N\,\mathbb{V}\mathrm{ar}[X_1]=Np(1-p)$

Demostració.

ii) Per ser X_i independents,

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

□

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

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

Distribució uniforme

Definició 3.3.7. Sigui X una variable aleatòria. Direm que X segueix una distribució uniforme

$$X\sim \mathrm{U}[1,N] \iff p(X=i)=\frac{1}{N} \text{ per } i=1,\dots,N.$$

Proposició 3.3.8. Sigui X una variable aleatòria que segueix una distribució uniforme. Aleshores,

i) $G_X(z)=\frac{1}{N}\frac{z(z^N-1)}{z-1}$

ii) $\mathbb{E}[X]=\frac{N+1}{2}$

iii) $\mathbb{V}\mathrm{ar}[X]=\frac{N^2-1}{12}$

Demostració. Directament de la definició de distribució uniforme tenim que

i) $G_X(z)=\sum_{n=1}^N \frac{1}{N}z^n=\frac{1}{N}(z+z^2+\cdots+z^N)=\frac{1}{N}\frac{z(z^N-1)}{z-1}$.

ii) $\mathbb{E}[X]=\sum_{i=0}^N i\frac{1}{N}=\frac{1}{N}\frac{N(N+1)}{2}=\frac{N+1}{2}$.

iii) Tenim que $\mathbb{V}\text{ar}[X] = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2$, i fent servir l'apartat anterior ens queda que $\mathbb{V}\text{ar}[X] = \left(\sum_{i=0}^{N} i^2 \frac{1}{N}\right) - \left(\frac{N+1}{2}\right)^2$. Fent servir la suma de quadrats i desenvolupant una mica finalment queda que

$$\mathbb{V}\text{ar}[X] = \frac{4N^2 + 6N + 2 - \left(3N^2 + 6N + 3\right)}{12} = \frac{N^2 - 1}{12}.$$

□

Distribució de Poisson

S'usa per modelar successos "estranys" (persones en una cua, emissió de partícules, etc).

Definició 3.3.9. Sigui X una variable aleatòria. Direm que X segueix una distribució de Poisson

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

També es pot escriure $\mathrm{Po}(\lambda)$.

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. Sigui X una variable aleatòria que segueix una distribució de Poisson. Aleshores,

i) $G_X(z) = e^{\lambda(z-1)}$

ii) $\mathbb{E}[X] = \lambda$

iii) $\mathbb{V}\text{ar}[X] = \lambda$

Demostració.

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

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

iii) $\mathbb{E}[X(X-1)] = \lambda^2 e^{\lambda(z-1)}|_{z=1} = \lambda^2 \implies \mathbb{E}[X^2] = \lambda^2 + \lambda$, i per tant, $\mathbb{V}\text{ar}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$.

□

Observació 3.3.12. La suma de dues variables independents amb distribució de Poisson també té distribució de Poisson:

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

Distribució geomètrica

La distribució geomètrica representa el nombre d'experiments necessaris abans d'obtenir el primer èxit en un succés binari.

Definició 3.3.13. Sigui X una variable aleatòria. Direm que X segueix una distribució geomètrica
$$X \sim \mathrm{Geom}(p) \iff p(X = i) = (1-p)^{i-1} p,\ \forall i \geq 1.$$

Observació 3.3.14. La distribució està ben definida, atès que
$$\sum_{i\geq 1} p\,(1-p)^{i-1} = p \sum_{i\geq 1} (1-p)^{i-1} = p\frac{1}{1-(1-p)} = 1.$$

Proposició 3.3.15. Sigui X una variable aleatòria que segueix una distribució geomètrica. Aleshores,

i) $G_X(z) = \frac{pz}{1-(1-p)z}$

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

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

Demostració. Directament de les definicions tenim que

i) $G_X(z) = \sum_{i\geq 1} (1-p)^{i-1} pz^i = \frac{p}{1-p} \sum_{i\geq 1} ((1-p)\,z)^i = \frac{p}{1-p} \cdot \frac{(1-p)z}{1-(1-p)z} = \frac{pz}{1-(1-p)z}$.

ii)
$$\begin{aligned}
\mathbb{E}\,[X] &= \left(\frac{\mathrm{d}G_X\,(X)}{\mathrm{d}z}\right)(1) = \\
&= \left(\frac{p\,(1-(1-p)\,z) + p\,(1-p)\,z}{(1-(1-p)\,z)^2}\right)(1) = \\
&= \frac{p^2 + p - p^2}{p^2} = \\
&= \frac{1}{p}.
\end{aligned}$$

iii) Es pot fer igual que en la distribució anterior: calculant $\mathbb{E}\,[X\,(X-1)]$ amb la segona derivada de $G_X\,(z)$ i utilitzant que $\mathbb{V}\mathrm{ar}\,[X] = \mathbb{E}\left[X^2\right] - \mathbb{E}\,[X]^2$, però les expressions són farragoses i no ho farem.

□

Distribució binomial negativa

Modela el nombre d'experiments necessaris per aconseguir un nombre d'èxits r donat.

Definició 3.3.16. Sigui X una variable aleatòria. Direm que X segueix una distribució binomial negativa
$$X \sim \mathrm{BinN}(r, p) \iff X = X_1 + \cdots + X_r,$$
on $\{X_i\}_{i=1}^r$ són independents i $X_i \sim \mathrm{Geom}(p)\ \forall i$.

Observació 3.3.17. Una distribució binomial negativa és equivalent a r distribucions geomètriques seguides.

Proposició 3.3.18. Sigui X una variable aleatòria que segueix una distribució binomial negativa. Aleshores,

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

ii) $p(X=k) = \begin{cases} 0 & \text{si } k < r \\ \binom{k-1}{r-1} p^r (1-p)^{k-r} & \text{si } k \geq r \end{cases}$

iii) $\mathbb{E}[X] = r\mathbb{E}[X_1] = \frac{r}{p}$.

iv) $\mathbb{V}\text{ar}[X] = r\,\mathbb{V}\text{ar}[X_1] = r\frac{1-p}{p^2}$.

Demostració.

i) $G_X(z) = (G_{X_1}(z))^r = \left(\frac{pz}{1-(1-p)z}\right)^r$.

ii) La probabilitat que es necessitin exactament k experiments per obtenir r èxits és la probabilitat d'obtenir $r-1$ èxits i $k-r$ fracassos als primers $k-1$ experiments per la probabilitat d'obtenir èxit al k-èssim experiment. Si $k < r$, això val 0, com és natural, i si $k \geq r$, només cal comptar el nombre total de combinacions que donen lloc a aquest esdeveniment:

$$p(X=k) = \left(\binom{k-1}{r-1} p^{r-1}(1-p)^{k-r}\right) p = \binom{k-1}{r-1} p^r (1-p)^{k-r}.$$

iii)

$$\begin{aligned} \mathbb{E}[X] &= \left(\frac{\mathrm{d}(G_X(X_1))^r}{\mathrm{d}z}\right)(1) = \\ &= \mathbb{E}[X_1]\left(r\left(\frac{pz}{1-(1-p)z}\right)^{r-1}\right)(1) = \\ &= r\mathbb{E}[X_1]. \end{aligned}$$

iv) Es resol de forma anàloga als casos anteriors.

$\square$

3.4 Distribucions condicionades i esperança condicionada

Definició 3.4.1. Siguin X i Y variables aleatòries discretes i sigui $x \in \mathbb{R}$ tal que $p(X=x) > 0$. Aleshores definim

1) $F_{Y|X}(y,x) = p(Y \leq y \mid X = x)$ com la funció de distribució condicionada de Y amb $X = x$,

2) $P_{Y|X}(y,x) = p(Y = y \mid X = x)$ com la funció de probabilitat condicionada.

Observació 3.4.2. Una definició anàloga consisteix a prendre $A \in \mathcal{A}$ enlloc de $X = x$ sempre que $p(A) > 0$. Aleshores tenim

1) $F_{Y|A}(y) = p(Y \leq y \mid A)$

2) $P_{Y|A}(y) = p(Y = y \mid A)$

Exemple 3.4.3. Siguin $\{Y_r\}_{r\geq 1}$ variables aleatòries independents tals que $Y_r \sim \text{Be}(p)\ \forall i$. Sigui X una variable aleatòria tal que $X = i$ si $Y_1 = Y_2 = \cdots = Y_{i-1} = 0$ i $Y_i = 1$. Observem que $X \mid \{Y_1 = 1\} = X|_{Y_1=1} = 1$, $X|_{Y_1=0} = 1 + \text{Geom}(p)$ i $X|_{Y_1=0,Y_2=0} = 2 + \text{Geom}(p)$.

Definició 3.4.4. Siguin X, Y variables aleatòries discretes, i sigui x tal que $p(X = x) > 0$. L'esperança condicionada de Y a $X = x$ és

$$\mathbb{E}[Y|X = x] = \sum_{y\in\text{Im}(Y)} y \cdot p(Y = y|X = x) = \phi(x).$$

Observem que el valor de $\mathbb{E}[Y|X = x]$ por variar depenent de l'elecció d'x.

Definició 3.4.5. Siguin X, Y variables aleatòries discretes i sigui $y \in \mathbb{R}$. Definim l'aplicació

$$\begin{aligned} \mathbb{E}[Y|X] :\ & \text{Im}(X) \to \mathbb{R} \\ & x \mapsto \mathbb{E}[Y|X = x] = \phi(x). \end{aligned}$$

Observació 3.4.6. Observem que $\mathbb{E}[Y|X]$ és una variable aleatòria i que

$$p(\mathbb{E}[Y|X] = y) = \sum_{x\in\phi^{-1}(y)} p(X = x).$$

Exemple 3.4.7.

$$\left.\begin{aligned} p(X = 1) = \tfrac{1}{4}, \phi(1) = 2 \\ p(X = 2) = \tfrac{1}{4}, \phi(2) = 2 \end{aligned}\right\} \implies p(\mathbb{E}[Y|X] = 2) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}.$$

Observació 3.4.8. $\phi(x) = \mathbb{E}[Y|X = x]$ és una funció sobre x amb valors reals i, en canvi, $\mathbb{E}[Y|X]$ és una variable aleatòria.

Proposició 3.4.9. Sigun X i Y variables aleatòries discretes. Aleshores,

$$\mathbb{E}[\mathbb{E}[Y|X]] = \mathbb{E}[Y].$$

Demostració.

$$\begin{aligned} \mathbb{E}[\mathbb{E}[Y|X]] &= \sum_{x\in\text{Im}(X)} \phi(x)p(X = x) = \\ &= \sum_{x\in\text{Im}(X)} \sum_{y\in\text{Im}(Y)} yp(Y = y|X = x)p(X = x) = \\ &= \sum_{y\in\text{Im}(Y)} \sum_{x\in\text{Im}(X)} yp(Y = y|X = x)p(X = x) = \\ &= \sum_{y\in\text{Im}(Y)} y \sum_{x\in\text{Im}(X)} p(Y = y, X = x) = \\ &= \sum_{y\in\text{Im}(Y)} yp(Y = y) = \mathbb{E}[Y]. \end{aligned}$$

□

Observació 3.4.10. Siguin X, Y variables aleatòries discretes independents. Aleshores, per qualsevol $x \in \mathrm{Im}(X)$,

$$\mathbb{E}[Y|X=x] = \sum_{y \in \mathrm{Im}(Y)} yp(Y=y|X=x) = \sum_{y \in \mathrm{Im}(Y)} yp(Y=y) = \mathbb{E}[Y],$$

i per tant $\mathbb{E}[Y|X] = \mathbb{E}[Y]$ amb probabilitat 1.

Observació 3.4.11. Per altra banda, si $Y = f(X)$ amb f bijectiva,

$$\phi(x) = \sum_{y \in \mathrm{Im}(Y)} yp(Y=y|X=x) = \sum_{z \in \mathrm{Im}(X)} f(z)p(f(X)=f(z)|X=x) = f(x).$$

i per tant $\mathbb{E}[Y|X] = f(X) = Y$.

Exemple 3.4.12. $N \sim \mathrm{Po}(\lambda)$, $K \sim \mathrm{Bin}(N,p)$. Per cada n,

$$\begin{aligned}\phi(n) = \mathbb{E}[K|N=n] &= \sum_{i \in \mathrm{Im}(K)} ip(K=i|N=n) = \\ &= \sum_{i=0}^{n} ip(K=i|N=n) = \sum_{i=0}^{n} i\binom{n}{i}p^i(1-p)^{n-i} = np.\end{aligned}$$

Per tant, $\mathbb{E}[K|N] = np$ amb probabilitat $p(N=n) = \frac{1}{n!}\lambda^m e^{-\lambda}$.
Si volguéssim calcular $\mathbb{E}[N|K]$, podríem prendre $k \in \mathrm{Im}(K)$ i fer

$$\begin{aligned}\mathbb{E}[N|K=k] &= \sum_{i \geq 0} ip(N=i|K=k) = \sum_{i \geq 0} i\frac{p(N=i \cap K=k)}{p(K=k)} = \\ &= \sum_{i \geq 0} i\frac{p(K=k|N=i)}{p(K=k)}p(N=i).\end{aligned}$$

Observació 3.4.13. Com podem veure amb l'exemple anterior, si $\mathbb{E}[Y|X=x] = \phi(x)$, aleshores $\mathbb{E}[Y|X] = \phi(X)$.

3.5 Arbres de Galton-Watson

Els arbres de Galton-Watson són una aplicació dels conceptes de distribució condicionada i esperança condicionada.

Sigui X una variable aleatòria discreta amb $\mathrm{Im}\,(X) = \mathbb{N}$. Considerem ara l'evolució de següent procés estocàstic. Sigui Z_n el nombre d'individus de l'n-èssima generació i posem Z_0=1. Cada individu de la generació n té fills seguint la distribució X, i els fills dels individus de la generació n conformen la generació $n+1$. Volem estudiar Z_n i determinar amb quina probabilitat la població s'extingeix en alguna generació.

Lema 3.5.1. Sigui $\{X_i\}_{i \geq 1}$ una seqüència de variables aleatòries discretes amb imatge en $\mathbb{N}$, idènticament distribuïdes i amb funció generadora de probabilitat $G_X\,(z)$. Sigui N una variable aleatòria discreta independent dels X_i amb imatge en $\mathbb{N}$ i funció generadora de probabilitat $G_N\,(z)$. Sigui $S = X_1 + \cdots + X_N$ una variable aleatòria discreta. Aleshores,

$$G_S\,(z) = G_N\,(G_X\,(z))\,.$$

Demostració.

$$\begin{aligned} G_S(z) &= \mathbb{E}\left[z^S\right] = \mathbb{E}\left[\mathbb{E}\left[z^S|N\right]\right] = \sum_{n\geq 1} \mathbb{E}\left[z^{X_1+\cdots+X_n}\right] p(N=n) = \\ &= \sum_{n\geq 1}\prod_{i=1}^{n} \mathbb{E}\left[z^{X_i}\right] p(N=n) = \sum_{n\geq 1} (G_X(z))^n \, p(N=n) = \\ &= G_N(G_X(z)). \end{aligned}$$

□

Teorema 3.5.2. Sigui $G_n(z)$ la funció generadora de probabilitat de Z_n. Aleshores,

$$G_{n+m}(z) = G_n(G_m(z))$$

i, per tant, $G_n(z) = \underbrace{G_X \circ \cdots \circ G_X(z)}_{n \text{ termes}}$.

Demostració. Primer observem que cada individu de la generació $n+m$ té exactament un ascendent en la generació m. Així doncs,

$$Z_{n+m} = \sum_{i=1}^{Z_m} Y_i,$$

on Y_i és el nombre de membres de la generació $n+m$ amb ancestre i a la geneació m. Aquestes Y_i són variables aleatòries discretes independents i amb distribució igual a Z_n. En virtut del lema anterior, concloem el que volíem: $G_{n+m}(z) = G_n(G_m(z))$. □

Observació 3.5.3. En general, és difícil calcular $G_n(z)$, però, en canvi, $\mathbb{E}[Z_n]$ i $\mathbb{V}\mathrm{ar}[Z_n]$ són fàcils de calcular.

$$\begin{aligned} \mathbb{E}[Z_n] &= \mathbb{E}[X]^n, \\ \mathbb{V}\mathrm{ar}[Z_n] &= \begin{cases} n\,\mathbb{V}\mathrm{ar}[X] & \text{si } \mathbb{E}[X]=1, \\ \mathbb{V}\mathrm{ar}[X]\,\mathbb{E}[X]^{n-1}\frac{\mathbb{E}[X]^n-1}{\mathbb{E}[X]-1} & \text{altrament.} \end{cases} \end{aligned}$$

Estudiem ara quan es produeix l'extinció de la població. Denotem $A_n = \{Z_n = 0\}$, és a dir, els successos pels quals la població ja s'ha extingit a la generació n i denotem també $\mathrm{Ex} = \bigcup_{n\geq 1} A_n$, és a dir, els successos pels quals la població s'extingeix en algun moment. Observem que $A_1 \subseteq A_2 \subseteq \cdots$ i que, per tant, $p(\mathrm{Ex}) = \lim_{n\to\infty} p(Z_n = 0)$.

Teorema 3.5.4. Sigui $\mathbb{E}[X] = \mu$ i sigui η la solució positiva més petita de l'equació del punt fix $G_X(z) = z$. Aleshores, $p(\mathrm{Ex}) = \eta$. En particular,

i) $\mu < 1 \implies \eta = 1$. La població s'extingeix amb probabilitat 1 (règim subcrític).

ii) $\mu > 1 \implies \eta < 1$. La població pot no extingir-se (règim supercrític).

iii) $\mu = 1 \implies \eta = 1$, sempre que $\mathbb{V}\mathrm{ar}[X] > 0$. La població s'extingeix amb probabilitat 1.

Demostració. Sigui $\eta_n = p\left(Z_n = 0\right) = G_n\left(0\right) = G_X\left(G_{n-1}\left(0\right)\right) = G_X\left(\eta_{n-1}\right), \forall n > 0$ i amb $\eta_0 = 0$. Aleshores, cal estudiar la solució límit de l'equació $\eta_n = G_X\left(\eta_{n-1}\right)$. Per fer-ho, estudiarem el comportament de la funció G_x a l'interval $[0,1]$. Recordem que aquesta funció té totes les derivades positives en aquest interval (en particular és creixent i convexa) i la seva derivada a $z = 1$ és $\mu = \mathbb{E}[X]$. Farem un esquema de la prova dividint el problema en els tres casos corresponents a l'enunciat, és a dir, segons si $\mu < 1$, $\mu = 1$ o $\mu > 1$. En tots tres casos, observem que la iteració de η_n convergeix al primer punt fix de la funció, i que $z = 1$ sempre és punt fix.

i) Quan $\mu < 1$, el primer punt fix serà a $z = 1$, ja que en cas contrari la convexitat de la funció junt amb el fet que $G_X(1) = 1$ obligaria a la derivada a ser més gran que 1 a $z = 1$.

ii) Quan $\mu = 1$, l'anàlisi és equivalent al cas anterior però cal afegir que si $\mathbb{V}\mathrm{ar}[X] = 0$ aleshores $G_X(0) = 0$ cosa que trenca l'anàlisi anterior. En aquest cas, no hi ha extinció.

iii) Quan $\mu > 1$, G_X és necessàriament menor que la recta z a un entorn de 1, i per tant hi ha un punt fix abans. Així doncs, l'extinció serà amb probabilitat η, amb η primer punt fix de G_x.

$\square$

Tema 4

Variables aleatòries contínues

4.1 Mesures de probabilitat absolutament contínues. Funció de densitat

Sigui $(\Omega, \mathcal{A}, p)$ un espai de probabilitat i sigui X una variable aleatòria. Aleshores X indueix una probabilitat P_X sobre $(\mathbb{R}, \mathcal{B})$. Ara estudiarem variables aleatòries X on P_X és "compatible" amb la mesura de Lebesgue λ.

Definició 4.1.1. Sigui $(\Omega, \mathcal{A})$ un espai de mesurable i μ_1, μ_2 dues mesures sobre $(\Omega, \mathcal{A})$. Diem que μ_1 és absolutament contínua respecte a μ_2 ($\mu_1 \ll \mu_2$) si

$$\forall A \in \mathcal{A}, \quad \mu_2(A) = 0 \implies \mu_1(A) = 0$$

Definició 4.1.2. Una variable aleatòria X és absolutament contínua (també anomenada contínua) si $P_X \ll \lambda$.

Observació 4.1.3. Les variables aleatòries discretes no són contínues. Sigui X una variable aleatòria discreta amb $\operatorname{Im}(X) = \{a_i\}_{i\geq 1}$, i $p(X = a_i) = p_i > 0$. Aleshores $P_X(\{a_i\}) = p_i$, però $\lambda(\{a_i\}) = 0$.

Teorema 4.1.4. *Teorema de Radon-Nikodym.*
Sigui $(\Omega, \mathcal{A})$ un espai mesurable i $\mu_1 \ll \mu_2$ dues mesures. Aleshores existeix una funció $f_{\mu_1} \colon \Omega \to \mathbb{R}$ mesurable en $(\Omega, \mathcal{A})$ tal que

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

És a dir, $\mathrm{d}\mu_1 = f_{\mu_1}\, \mathrm{d}\mu_2$, o "$\frac{\mathrm{d}\mu_1}{\mathrm{d}\mu_2} = f_{\mu_1}$" (derivada de Radon-Nikodym).

Demostració. És una prova complicada i amb eines avançades que no farem. $\square$

Corol·lari 4.1.5. En la nostra situació, $\mu_1 = P_X$ i $\mu_2 = \lambda$, per tant si $P_X \ll \lambda$ tenim

$$\forall A \in \mathcal{B}, \quad P_X(A) = \int_A f_X\, \mathrm{d}\lambda = \int_{\mathbb{R}} f_X \mathbb{I}_A\, \mathrm{d}\lambda$$

Definició 4.1.6. A $f_X = f_{\mu_1}$ l'anomenarem funció de densitat de probabilitat de X.

Observació 4.1.7. El teorema de Radon-Nikodym no afirma la unicitat de f_{μ_1}, de fet si f_{μ_1} i $\overline{f_{\mu_1}}$ satisfan les condicions, aleshores el teorema també afirma que $f_{\mu_1} = \overline{f_{\mu_1}}$ μ_2-g.a.

Proposició 4.1.8. *Propietats de la funció de densitat.* Sigui X una variable aleatòria amb funció de densitat f_X,

i) $f_X \geq 0 \quad \lambda$-g.a.

ii) $\int_{\mathbb{R}} f_X \,\mathrm{d}\lambda = 1$.

iii) Si f_X és integrable Riemann, $\forall A = (a,b)$ interval, $P_X(A) = \int_A f_X \,\mathrm{d}\lambda = \int_a^b f_X(x)\,\mathrm{d}x$.

iv) $p(X = x) = P_X(\{x\}) = \int_{\{x\}} f_X \,\mathrm{d}\lambda = 0, \quad \forall x \in \mathbb{R}$.

Demostració.

i) Sigui $A = \{x \in \mathbb{R} \colon f_X(x) < 0\}$, volem veure que té mesura (de Lebesgue) 0.

$$A_n = \bigcup_{n\geq 1} \left\{ x \in \mathbb{R} \colon f_X(x) < \frac{-1}{n} \right\}$$

i $\{A_n\}_{n\geq 1}$ és creixent.

$$0 \leq P_X(A_n) = \int_{A_n} f_X \,\mathrm{d}\lambda \leq \frac{-1}{n}\lambda(A_n) \leq 0 \implies \lambda(A_n) = 0, \forall n$$

Per tant, $\lambda(A) = \lim_n \lambda\left(A_n\right) = 0$. Les altres tres en són conseqüència. □

Observació 4.1.9. Tota funció f que compleixi les tres primeres propietats defineix una variable aleatòria X:

$$F_X(x) = \int_{-\infty}^{x} f(x)\,\mathrm{d}x$$

De fet quan es defineix informalment una variable aleatòria contínua se sol definir com una funció amb aquestes propietats.

Ara volem calcular l'esperança d'una variable aleatòria X i, en general, calcular $\mathbb{E}[g(x)]$ per $g \colon \mathbb{R} \to \mathbb{R}$ funció mesurable.

Proposició 4.1.10.

i) Per l'esperança d'una variable aleatòria X contínua tenim

$$\mathbb{E}[X] = \int_{\mathbb{R}} x \,\mathrm{d}p_X = \int_{\mathbb{R}} x f_X \,\mathrm{d}\lambda = \int_{-\infty}^{+\infty} x f(x)\,\mathrm{d}x.$$

ii) I per la seva variància

$$\mathbb{V}\mathrm{ar}[X] = \int x^2 f_X(x)\,\mathrm{d}\lambda - \left(\int x f_X \,\mathrm{d}\lambda\right)^2.$$

Observació 4.1.11. Si volem que $\mathbb{E}[g(x)] < +\infty$, és prou amb demanar

$$\int_{\mathbb{R}} |g(x)||f_X(x)|\,\mathrm{d}\lambda < +\infty.$$

4.2 Models de variables aleatòries contínues

Distribució uniforme

Dóna un valor a $[a, b)$ uniformement a l'atzar.

Definició 4.2.1. $\mathrm{U}(a, b)$ és la distribució amb funció de densitat

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

on $\mathbb{I}_{[a,b]}$ és la funció identitat de $[a, b]$ (val 1 si $x \in [a, b]$ i 0 altrament).

Proposició 4.2.2.

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

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

Demostració.

i) $\mathbb{E}[X] = \int_{-\infty}^{+\infty} xf(x)\,\mathrm{d}x = \int_a^b x\frac{1}{b-a}\,\mathrm{d}x = \frac{1}{b-a}\left[\frac{x^2}{2}\right]_a^b = \frac{1}{2}\frac{b^2-a^2}{b-a} = \frac{a+b}{2}$.

ii) $\mathbb{E}[X^2] = \frac{b^2+ab+a^2}{3} \implies \mathbb{V}\mathrm{ar}[X] = \frac{(b-a)^2}{12}$.

□

Distribució exponencial

Es pot entendre com l'extensió de la distribució geomètrica.

Definició 4.2.3. $\mathrm{Exp}(\lambda)$ és la distribució amb funció de densitat

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

Proposició 4.2.4.

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

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

Distribució normal

Es pot entendre com l'extensió de la distribució binomial.

Definició 4.2.5. $\mathrm{N}(\mu, \sigma^2)$ és la distribució amb funció de densitat

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

Altres distribucions contínues

Definició 4.2.6. Hi ha altres distribucions de variables aleatòries contínues:

Nom	Símbol	Funció de densitat $F_X(x)$
Gamma	$\Gamma(\lambda, z)$	$\frac{\lambda^z}{\Gamma(z)} x^{z-1} e^{-\lambda x} \mathbb{I}_{[0,+\infty)}(x)$
Weibull	$\text{Weib}(\alpha, \beta)$	$\alpha\beta \exp\left(-\alpha x^\beta\right) \mathbb{I}_{[0,\alpha)}(x)$
Beta	$\beta(a, b)$	$\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1} \mathbb{I}_{(0,1)}(x)$
Cauchy	$\text{Cauchy}(0, 1)$	$\frac{1}{\pi(1+x^2} \mathbb{I}_{[0,+\infty)}(x)$

Nom	$\mathbb{E}[X]$	$\mathbb{V}\text{ar}[X]$
Gamma	$\frac{z}{\lambda}$	$\frac{z}{\lambda^2}$
Weibull	$\alpha^{\frac{-1}{\beta}} \Gamma(1 + \frac{1}{\beta})$	$\alpha^{\frac{-2}{\beta}} (\Gamma(1 + \frac{2}{\beta}) - \Gamma(1 - \frac{2}{\beta}))^2$
Beta	$\frac{a}{a+b}$	$\frac{ab}{(a+b)^2(a+b+1)}$
Cauchy	$+\infty$	No definida

On $\Gamma(z) = \int_0^{+\infty} x^{z-1} e^{-x} \, \mathrm{d}x$ és la funció gamma.

4.3 Conceptes de variables aleatòries contínues

Farem tota l'anàlisi per dues variables aleatòries, es pot generalitzar immediatament a k variables aleatòries. Prenem (X, Y) una variable aleatòria multidimensional de dimensió 2.

Definició 4.3.1. Direm que (X, Y) és un vector de variables aleatòries absolutament continu si $P_{(X,Y)} \ll \lambda_{\mathbb{R}^2}$.

Observació 4.3.2. En aquesta situació l'extensió del teorema de Radon-Nikodyn a $\mathbb{R}^2$ garanteix que $\exists f_{(X,Y)} \colon \mathbb{R}^2 \to \mathbb{R}$ mesurable en $(\mathbb{R}^2, \mathcal{B}_{\mathbb{R}^2})$ i per la qual

$$B \in \mathcal{B}_{\mathbb{R}^2}, \quad P_{(X,Y)}(B) = \int_B f_{(X,Y)} \, \mathrm{d}\lambda_{\mathbb{R}^2}$$

que, si $f_{(X,Y)}$ és integrable Riemann i $B = (a, b) \times (c, d)$, és $\int_a^b \int_c^d f_{(X,Y)}(x, y) \, \mathrm{d}x \, \mathrm{d}y$

4.3.1 Distribució conjunta i marginals

Definició 4.3.3. $f_{(X,Y)}$ és la funció de densitat de probabilitat del vector (X, Y), o funció de densitat conjunta de X i Y.

Definició 4.3.4. D'aquesta forma podem definir $F_{(X,Y)}(x,y)$ a partir de $f_{(X,Y)}(x,y)$:

$$F_{(X,Y)}(x,y) = p(X \leq x, Y \leq y) = \int_B f_{(X,Y)} \,\mathrm{d}\lambda_{\mathbb{R}^2}$$

on $B = (-\infty, x) \times (-\infty, y)$.

Si $f_{(X,Y)}$ és integrable Riemann, $F_{(X,Y)}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{(X,Y)}(u,v) \,\mathrm{d}u \,\mathrm{d}v$.

Proposició 4.3.5. Si (X,Y) és un vector de variables aleatòries absolutament continu (a $\mathbb{R}^2$), aleshores tant X com Y són variables aleatòries absolutament contínues (respecte $\lambda_{\mathbb{R}}$).

Demostració. Es pot fer per contradicció. $\square$

Observació 4.3.6. Per tant, X i Y tenen funcions de densitat $f_X(x)$ i $f_Y(y)$, respectivament.

Definició 4.3.7. Donat el vector absolutament continu (X,Y) amb funció de densitat conjunta $f_{(X,Y)}$, les funcions de probabilitat marginals són f_X i f_Y.

Proposició 4.3.8. En particular, si són integrables Riemann,

$$f_X(u) = \int_{-\infty}^{+\infty} f_{(X,Y)}(u,v) \,\mathrm{d}v.$$

$$f_Y(v) = \int_{-\infty}^{+\infty} f_{(X,Y)}(u,v) \,\mathrm{d}u.$$

Demostració.

$$F_X(x) = p(X \leq x) = p(X \leq x, Y \in (-\infty, +\infty)) = P_{(X,Y)}((-\infty, x) \times \mathbb{R}).$$

Que, per ser integrable Riemann, val $\int_{-\infty}^{x} \int_{-\infty}^{+\infty} f_{(X,Y)}(u,v) \,\mathrm{d}v \,\mathrm{d}u$. Per tant,

$$\int_{-\infty}^{x} f_X(u) \,\mathrm{d}u = \int_{-\infty}^{x} \int_{-\infty}^{+\infty} f_{(X,Y)}(u,v) \,\mathrm{d}v \,\mathrm{d}u \implies$$
$$0 = \int_{-\infty}^{x} \left[f_X(u) - \int_{-\infty}^{+\infty} f_{(X,Y)}(u,v) \,\mathrm{d}v \right] \mathrm{d}u.$$

Derivant obtenim que

$$0 = f_X(u) - \int_{-\infty}^{+\infty} f_{(X,Y)}(u,v) \,\mathrm{d}v.$$

$\square$

Observació 4.3.9. $f_{(X,Y)}$ conté més informació que f_X i f_Y.

Observació 4.3.10.

$$F_{(X,Y)}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{(X,Y)}(u,v) \,\mathrm{d}u \,\mathrm{d}v \implies \frac{\partial^2 F_{(X,Y)}}{\partial x \partial y}(x,y) = f_{(X,Y)}(x,y).$$

Observació 4.3.11. Si $g\colon \mathbb{R}^2 \to \mathbb{R}$ mesurable:

$$\mathbb{E}[g(X,Y)] = \int_{\mathbb{R}^2} g(x,y) f_{(X,Y)} \,\mathrm{d}\lambda_{\mathbb{R}^2} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x,y) f_{(X,Y)}(x,y) \,\mathrm{d}y \,\mathrm{d}x.$$

4.3.2 Independència

És l'únic cas en què podem obtenir informació de (X, Y) sabent només X i Y.

Proposició 4.3.12. Si X, Y són independents,

$$f_{(X,Y)}(u,v) = f_X(u) f_Y(v)$$

Demostració. $F_{(X,Y)}(x,y) = F_X(x)F_Y(y)$, per tant, $\forall x, y$,

$$\int_{-\infty}^{x}\int_{-\infty}^{y} f_{(X,Y)}(u,v)\,\mathrm{d}v\,\mathrm{d}u = \int_{-\infty}^{x} f_X(u)\,\mathrm{d}u \int_{-\infty}^{y} f_Y(v)\,\mathrm{d}v = \int_{-\infty}^{x}\int_{-\infty}^{y} f_X(u) f_Y(v)\,\mathrm{d}v\,\mathrm{d}u$$

Com que és cert $\forall x, y$, tenim que $f_{(X,Y)}(u,v) = f_X(u)f_Y(v)$. $\square$

Observació 4.3.13. Si X, Y són independents, $\forall B_1, B_2 \in \mathcal{B}$:

$$p(x \in B_1, y \in B_2) = p(x \in B_1)p(y \in B_2) = \int_{B_1} f_X(u)\,\mathrm{d}\lambda \int_{B_2} f_Y(v)\,\mathrm{d}\lambda$$

$$p((x,y) \in B_1 \times B_2) = \int_{B_1 \times B_2} f_{(X,Y)}(u,v)\,\mathrm{d}\lambda_{\mathbb{R}^2}$$

Per Fubini, $\int_{B_1} f_X(u)\,\mathrm{d}\lambda \int_{B_2} f_Y(v)\,\mathrm{d}\lambda = \int_{B_1\times B_2} f_{(X,Y)}(u,v)\,\mathrm{d}\lambda_{\mathbb{R}^2}$

Observació 4.3.14. Notació:

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

$$\mathbb{V}\mathrm{ar}[(X,Y)] = \begin{pmatrix} \mathbb{V}\mathrm{ar}[X] & \mathbb{C}\mathrm{ov}(X,Y) \\ \mathbb{C}\mathrm{ov}(X,Y) & \mathbb{V}\mathrm{ar}[Y] \end{pmatrix}$$

Si X, Y són independents, $\mathbb{V}\mathrm{ar}[(X,Y)]$ és una matriu diagonal $\mathrm{diag}(\mathbb{V}\mathrm{ar}[X], \mathbb{V}\mathrm{ar}[Y])$.

4.3.3 Distribucions condicionades

Definició 4.3.15. Siguin X, Y variables aleatòries absolutament contínues i sigui x pel qual $f_X(x) > 0$. La variable aleatòria Y condicionada a $X = x$ $(Y|X = x)$ és la variable aleatòria amb funció de distribució

$$F_{Y|X}(y,x) = \frac{1}{f_X(x)} \int_{-\infty}^{y} f_{(X,Y)}(x,v)\,\mathrm{d}v.$$

En aquesta situació, la funció de densitat de $Y|X = x$ és $\frac{f_{(X,Y)}(x,y)}{f_X(x)}$.

Definició 4.3.16. Per a x t. q. $f_X(x) > 0$, l'esperança condicionada $\mathbb{E}[Y|X = x]$ és

$$\int_{-\infty}^{+\infty} y f_{Y|X}(y,x)\,\mathrm{d}y = \psi(x).$$

I $\mathbb{E}[Y|X]$ és $\psi(X)$.

Teorema 4.3.17.

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

Demostració.

$$\begin{aligned}
\mathbb{E}[\mathbb{E}[Y|X]] &= \int_{-\infty}^{+\infty} \mathbb{E}[Y|X=x] f_X(x)\,\mathrm{d}x = \\
&= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} y f_{Y|X}(y,x)\,\mathrm{d}y f_X(x)\,\mathrm{d}x = \\
&= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} y \frac{f_{Y,X}(y,x)}{f_X(x)}\,\mathrm{d}y f_X(x)\,\mathrm{d}x = \\
&= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} y f_{Y,X}(y,x)\,\mathrm{d}x\,\mathrm{d}y = \\
&= \int_{-\infty}^{+\infty} y \int_{-\infty}^{+\infty} f_{Y,X}(y,x)\,\mathrm{d}x\,\mathrm{d}y = \\
&= \int_{-\infty}^{+\infty} y f_Y(y)\,\mathrm{d}y \\
&= \mathbb{E}[Y]
\end{aligned}$$

□

4.4 Distribució normal, multivariant i distribucions associades

Definició 4.4.1. La distribució normal ve donada per dos paràmetres μ i σ^2 i té funció de distribució

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

La denotarem per $\mathrm{N}(\mu, \sigma^2)$

Exemple 4.4.2.

$$p(x \in [\mu-\sigma, \mu+\sigma]) \approx 0,68$$
$$p(x \in [\mu-2\sigma, \mu+2\sigma]) \approx 0,955$$
$$p(x \in [\mu-1,95\sigma, \mu+1,95\sigma]) \approx 0,95$$

Teorema 4.4.3. *Teorema de Moivre-Laplace.*
Prenem p constant, $X_n = \mathrm{Bin}(n,p)$ i $a, b \in \mathbb{R}$. Llavors,

$$p\left(a \le \frac{X_n - np}{\sqrt{np(1-p)}} \le b\right) \xrightarrow{n\to\infty} \frac{1}{\sqrt{2\pi}} \int_a^b e^{\frac{-x^2}{2}}\,\mathrm{d}x$$

Observació 4.4.4. Demostrarem més endavant, usant funcions característiques, que si $X \sim \mathrm{N}(\mu_1, \sigma_1^2)$ i $Y \sim \mathrm{N}(\mu_2, \sigma_2^2)$ són independents $X + Y \sim \mathrm{N}\left(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2\right)$.

Ara veurem unes quantes distribucions relacionades amb la normal.

Definició 4.4.5. Donades $X \sim \mathrm{N}(0,1)$, $\{X_i\}_{i=1}^n$ independents, la distribucó χ^2 (chi quadrat) és

$$\chi_n^2 \sim X_1^2 + \cdots X_n^2$$

Definició 4.4.6. Donades $\chi^2_{d_1}, \chi^2_{d_2}$ independents, la distribució de Fisher-Snedecor o distribució F és

$$\mathrm{F} \sim \frac{\chi^2_{d_1}/d_1}{\chi^2_{d_2}/d_2}$$

Definició 4.4.7. Donades $\mathrm{N}(0,1)$ i χ^2_k independents, la distribució t de Student és

$$t \sim \frac{\mathrm{N}(0,1)}{\sqrt{\chi^2_k/k}}$$

4.4.1 Distribució normal multivariant

Es tracta d'un vector de variables aleatòries que generalitza la normal univariant.

Definició 4.4.8. El vector aleatori $\overrightarrow{X} = (X_1, \ldots, X_n)$ té una distribució normal multivariant si $f_{\overrightarrow{X}}(x_1, \ldots, x_n)$ s'escriu com

$$\frac{1}{\sqrt{(2\pi)^n \det\{\Sigma\}}} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu})\right)$$

on $\mathbf{x} = (x_1, \ldots, x_n), \boldsymbol{\mu} = (\mu_1, \ldots, \mu_n) \in \mathbb{R}^n$ i Σ és una matriu $n \times n$ simètrica definida positiva.

Observació 4.4.9.

$$\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f_{\overrightarrow{X}}(x_1, \ldots, x_n)\, \mathrm{d}x_1 \cdots \mathrm{d}x_n = 1$$

Exemple 4.4.10. Si $U_1, \ldots, U_n$ són variables aleatòries $\mathrm{N}(0,1)$ independents, $\overrightarrow{U} = (U_1, \ldots, U_n)$. Aleshores

$$\begin{aligned} f_{\overrightarrow{U}}(x_1, \ldots, x_n) &\overset{\text{indep.}}{=} \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x_i^2}{2}\right) \\ &= \frac{1}{\sqrt{(2\pi)^n \cdot 1}} \exp\left(-\frac{1}{2}\left(x_1^2 + \cdots x_n^2\right)\right) \\ &= \frac{1}{\sqrt{(2\pi)^n \det(\Sigma)}} \exp\left(-\frac{1}{2}\mathbf{x}^T \Sigma^{-1} \mathbf{x}\right) \end{aligned}$$

on $\Sigma = I_n$ és la identitat, que és simètrica i definida positiva.

Observació 4.4.11. De fet, el que veurem ara és que si $\overrightarrow{X}$ és una multivariant, aleshores $\overrightarrow{X}$ es pot construir mitjançant una transformació lineal de $\overrightarrow{U} = (U_1, \ldots, U_n)$ on $U_i \sim \mathrm{N}(0,1)$ i $\{U_i\}_{i=1}^n$ independents.

Teorema 4.4.12. Si $\overrightarrow{X}$ és una normal univariant, aleshores $\exists A$ matriu quadrada no singular, un vector $\mathbf{b}$ pel qual $\overrightarrow{X} = A\overrightarrow{U} + \mathbf{b}$, on $\overrightarrow{U} = (U_1, \ldots, U_n)$ amb $U_i \sim \mathrm{N}(0,1)$ són variables aleatòries independents.

Demostració. Σ és simètrica i definida positiva, per tant Σ^{-1} també, i per tant $\Sigma^{-1} = A^T D A$, on D és una matriu diagonal amb tots els elements de la diagonal positius, i A és una matriu no singular (ve de fer un canvi de base). Així,

$$-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T\Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu}) = -\frac{1}{2}\left[A(\mathbf{x}-\boldsymbol{\mu})\right]^T D\left[A(\mathbf{x}-\boldsymbol{\mu})\right]$$

Si fem $A(\mathbf{x}-\boldsymbol{\mu}) = \mathbf{y}$, la funció de densitat obtinguda és

$$f_{\mathbf{y}}(y_1,\dots,y_n) = \frac{1}{\sqrt{(2\pi)^n \det(D)}}\exp\left(-\frac{1}{2}\mathbf{y}^T D\mathbf{y}\right)$$

Aquesta fórmula és conseqüència de tenir en compte el jacobià de canvi de base. Ara, si $D = \operatorname{diag}(d_1^2,\dots,d_n^2)$ amb els $d_i^2 > 0$,

$$f_{\mathbf{y}}(y_1,\dots,y_n) = \prod_{i=1}^{n}\frac{1}{2\pi d_i^2}\exp\left(-\frac{1}{2}y_i^2 d_i^2\right)$$

Observem que amb això no demostrem el que volíem perquè les components no són $\mathrm{N}(0,1)$. Això ho podem arreglar prenent

$$\Sigma^{-1} = A^T D A = A d^T d A = (dA)^T\cdot I_n(dA)$$

on $d = \operatorname{diag}(d_1,\dots,d_n)$, i ara sí que obtenim $\mathrm{N}(0,1)$. $\square$

Observació 4.4.13. Ho veurem quan fem funcions característiques: $\mathbb{E}[\overrightarrow{X}] = \boldsymbol{\mu}$, $\Sigma_{ij} = \mathbb{C}\mathrm{ov}(X_i, Y_j)$.

Teorema 4.4.14. Sigui $\overrightarrow{X}$ una normal multivariant amb paràmetres $\boldsymbol{\mu}$ i Σ ($\mathrm{N}(\boldsymbol{\mu},\Sigma)$) i sigui M una matriu $m\times n$ amb rang m, $m\leq n$, on n és el nombre de components de $\overrightarrow{X}$. Considerem $\overrightarrow{Y} = M\overrightarrow{X}$. Aleshores

$$\overrightarrow{Y}\sim \mathrm{N}(M\boldsymbol{\mu}, M\Sigma M^T)$$

Exemple 4.4.15. Si prenem $\{a_i\}_{i=1}^n$ tal que $\sum_{i=1}^n a_i^2 > 0$ (algun és no nul), aleshores prenent $M = (a_1,\dots,a_n)$ tenim que $M\overrightarrow{X} = \sum_{i=1}^n a_i X_i$ és una normal. Per tant, encara que les $\{X_i\}_{i=1}^n$ siguin normals no independents (però provenen de la mateixa normal multivariant), la seva suma sí que és una normal. Això no ho podem afirmar en general per qualsevol parella de variables aleatòries normals Y_1, Y_2.

4.4.2 Estimadors i teorema de Fisher

Els resultats que veurem a continuació es veuen amb més detall a l'assignatura d'Estadística.

Definició 4.4.16. Donades les variables aleatòries $\{X_i\}_{i=1}^n$ independents i idènticament distribuïdes, l'esperança mostral és

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

Definició 4.4.17. La variància mostral és

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

Observació 4.4.18. En particular, si $\mathbb{E}[X_i] = \mu, \mathbb{V}\mathrm{ar}[X_i] = \sigma^2$, aleshores $\mathbb{E}[\bar{X}] = \mu, \mathbb{V}\mathrm{ar}[\bar{X}] = \frac{\sigma^2}{n}, \mathbb{E}[S^2] = \sigma^2$.

El teorema de Fisher particularitza aquests estimadors quan $X_i \sim \mathrm{N}(\mu, \sigma^2)$.

Teorema 4.4.19. *Teorema de Fisher.*
Siguin $\{X_i\}_{i=1}^n$ variables aleatòries independents amb $X_i \sim \mathrm{N}(\mu, \sigma^2)$. Aleshores, $\bar{X}$ i S^2 són independents i tenim que

$$\bar{X} \sim \mathrm{N}\left(\mu, \frac{\sigma^2}{n}\right), \qquad S^2 = \frac{n-1}{\sigma^2}\xi^2_{n-1}$$

4.5 Funcions de variables aleatòries absolutament contínues

Volem resoldre el següent problema:

Problema 4.5.1. Sigui $\overrightarrow{X} = (X_1, \ldots, X_n)$ una variable aleatòria multidimensional absolutament contínua amb funció de densitat $f_{\overrightarrow{X}}(x_1, \ldots, x_n)$. Sigui $g\colon \mathbb{R}^n \to \mathbb{R}^n$ bijectiva i $\mathscr{C}^1(\mathbb{R}^n)$. Si $g(\overrightarrow{X}) = \overrightarrow{Y}$, com podem trobar $f_{\overrightarrow{Y}}$?

Solució. Farem primer el cas unidimensional, $n = 1$. Sigui X una variable aleatòria amb funció de densitat $f_X(x)$ i $g\colon \mathbb{R} \to \mathbb{R}$ derivable i bijectiva. Suposem que g és a més monòtona creixent i $g^{-1} = h$. Aleshores, si fem $Y = g(X)$:

$$F_Y(u) = p(Y \leq u) = p(g(X) \leq u) = p(X \leq h(u)) = \int_{-\infty}^{h(u)} f_X(x)\,\mathrm{d}x$$

que fent $x = h(y), \mathrm{d}x = h'(y)\,\mathrm{d}y$ és

$$F_Y(u) = \int_{-\infty}^{u} f_X(h(y))h'(y)\,\mathrm{d}y$$

Fent els càlculs corresponents si g és monòtona decrexent obtenim que

$$f_Y(y) = f_X(h(y))(h'(y))$$

En el cas multivariant, el paper de ser bijectiva i estrictament monòtona ve donat pel fet que $\det(J_h(y_1, \ldots, y_n)) \neq 0$ (J_h és la jacobiana de h) i, procedint com abans, la transformació que fem és

$$f_{\overrightarrow{Y}}(y_1, \ldots, y_n) = f_{\overrightarrow{X}}(h(y_1, \ldots, y_n))\det(J_h(y_1, \ldots, y_n))$$

4.5.1 Aplicacions

Problema 4.5.2. Siguin X_1, X_2 variables aleatòries i $\overrightarrow{X} = (X_1, X_2)$ i volem calcular la funció de densitat de $X_1 + X_2$ en termes de $f_{X_1}(x_1)$, $f_{X_2}(x_2)$ o $f_{\overrightarrow{X}}(x_1, x_2)$.

Solució.

Problema 4.5.3. Siguin X_1, X_2 variables aleatòries, volem calcular la funció de densitat de $X_1 \cdot X_2$.

Solució.

Una altra aplicació n'és la simulació de variables aleatòries. Partint d'una variable aleatòria $\mathrm{U}[0,1]$ construirem qualsevol altra variable aleatòria X.

Lema 4.5.4. *Simulació de variables aleatòries.* Sigui X una variable aleatòria amb funció de distribució $F_X(x)$. Aleshores $F_X(X)$ és una uniforme $\mathrm{U}[0,1]$ (i per tant $F_X^{-1}(U) = X$).

Demostració. Buscarem la funció de densitat de $F_X(X)$. Això ho podem fer perquè $F_X(x)$ és $\mathscr{C}^1(\mathbb{R})$ i estrictament creixent (després veurem que fem si no és estrictament creixent):

$$\begin{aligned} g\colon \mathbb{R} &\to [0,1] \\ y &\mapsto F_X(y) \end{aligned}, \qquad \begin{aligned} h\colon (0,1) &\to \mathbb{R} \\ x &\mapsto F_X^{-1}(x) \end{aligned} \qquad (h = g^{-1})$$

Calculem doncs

$$\begin{aligned} f_{F_X(X)}(x) &= f_{g(X)}(x) \\ &= f_X(F_X^{-1}(x)) \frac{\mathrm{d}}{\mathrm{d}x}(h(x)) \\ &= f_X(F_X^{-1}(x)) \frac{\mathrm{d}}{\mathrm{d}x}(F_X^{-1}(x)) \\ &= f_X(F_X^{-1}(x)) \frac{1}{f_X(F_X^{-1}(x))} = 1 \end{aligned}$$

per tant $F_X(X)$ és una variable aleatòria uniforme.

Aquest argument només és vàlid quan F_X és estrictament creixent. En cas contrari, els punts on $F_X'(x)$ és 0 és un conjunt numerable i argumentant sobre la resta obtenim que $f_{F_X(X)}(x) = 1$ llevat d'un conjunt de mesura 0. □

Tema 5

Funcions característiques i famílies exponencials

5.1 Funció generadora de moments. Propietats i aplicació: fites de Chernoff

Aquesta eina ens permetrà estudiar de manera unificada les variables aleatòries discretes i les absolutament contínues.

Definició 5.1.1. Donada una variable aleatòria X, la seva funció generadora de moments és

$$M_X(s) = \mathbb{E}\left[e^{sX}\right],$$

on $s \in \mathbb{C}$. Si $\forall i \geq 1$, $\mathbb{E}\left[|X|^i\right] < \infty$ aleshores $\mathbb{E}\left[X^i\right] < \infty$ i tenim que

$$e^{sX} = \sum_{i \geq 0} \frac{X^i}{i!} s^i \implies \mathbb{E}\left[e^{sX}\right] = \mathbb{E}\left[\sum_{i \geq 0} \frac{X^i}{i!} s^i\right],$$

que, pel teorema de la convergència dominada, és igual a

$$\sum_{i \geq 0} \mathbb{E}\left[\frac{X^i}{i!} s^i\right] = \sum_{i \geq 0} \frac{\mathbb{E}\left[X^i\right]}{i!} s^i = M_X(s).$$

Aquesta sèrie de potències podria tenir radi de convergència igual a 0. Per exemple, si $\mathbb{E}\left[X^i\right] = i!^2$ aleshores el radi de convergència és 0.

Proposició 5.1.2. La funció generadora de moments té les següents propietats:

i) Si $Y = aX + b$, amb X variable aleatòria amb funció generadora de moments $M_X(s)$ i $a \in \mathbb{R}, b \in \mathbb{R}$, aleshores $M_Y(s) = e^{bs} M_x(as)$.

ii) $\frac{\mathrm{d}^k}{\mathrm{d}s^k} M_X(s)\Big|_{s=0} = \mathbb{E}\left[X^k\right]$.

iii) Si X i Y són variables aleatòries independents, aleshores $M_{X+Y}(s) = M_X(s) M_Y(s)$.

Demostració.

i) $M_Y(s) = \mathbb{E}\left[e^{Ys}\right] = \mathbb{E}\left[e^{(aX+b)s}\right] = e^{bs}\mathbb{E}\left[e^{(as)X}\right] = e^{bs} M_X(as)$.

ii) La prova es deixa com a exercici pel lector.

iii) $M_{X+Y}(s) = \mathbb{E}\left[e^{s(X+Y)}\right] = \mathbb{E}\left[e^{sX}e^{sY}\right] = \mathbb{E}\left[e^{sX}\right]\mathbb{E}\left[e^{sY}\right] = M_X(s)M_Y(s)$.

$\square$

Observació 5.1.3. De la fórmula anterior, si X i Y tenen tots els moments definits, aleshores $X+Y$ també: $\mathbb{E}\left[(X+Y)^k\right] = \sum\limits_{i=0}^{k}\binom{k}{i}\mathbb{E}\left[X^i\right]\mathbb{E}\left[Y^{k-i}\right]$.

Exemple 5.1.4. $X \sim \mathrm{Be}(p)$. Aleshores

$$M_X(s) = \mathbb{E}\left[e^{sX}\right] = e^{s\cdot 1}p + e^{s\cdot 0}(1-p) = e^s p + (1-p)$$

Exemple 5.1.5. $X \sim \mathrm{Exp}(\lambda)$. Aleshores

$$M_X(s) = \int_{-\infty}^{+\infty} e^{sx} f_X(x)\,\mathrm{d}x = \int_0^{\infty} e^{sx}\lambda e^{-\lambda x}\,\mathrm{d}x = \int_0^{\infty}\lambda e^{(s-\lambda)x}\,\mathrm{d}x = \left[\lambda\frac{e^{(s-\lambda)x}}{s-\lambda}\right]_0^{\infty}.$$

Que, si $s-\lambda$ té part real negativa, és igual a

$$\lambda\frac{-1}{s-\lambda} = \frac{\lambda}{\lambda-s}.$$

Observem que el radi de convergència és λ.

Observació 5.1.6. En general, el càlcul de les funcions generadores de moments serà de la forma

i) En el cas discret, $M_X(s) = \sum\limits_{x\in\mathrm{Im}(x)} e^{sx}p(X=x)$.

ii) En el cas absolutament continu, $M_X(s) = \int_{\mathbb{R}} e^{sx} f_X(x)\,\mathrm{d}x$.

5.1.1 Aplicació. Teorema de Chernoff

Observació 5.1.7. Prenem n variables aleatòries $\{X_i\}_{i=1}^n$ independents i amb la mateixa distribució. Sigui $S_n = \sum\limits_{i=1}^{n} X_i$, $\mathbb{E}[X_i] = \mu, \mathbb{V}\mathrm{ar}[X_i] = \sigma^2$. Si apliquem la desigualtat de Txebixov (2.2.12) obtenim que

$$p\left(|S_n - n\mu| > \varepsilon\right) \le \frac{n\sigma^2}{\varepsilon^2} \implies 0 \le p\left(\left|\frac{S_n}{n} - \mu\right| > \varepsilon\right) \le \frac{\sigma^2}{n\varepsilon^2}$$

Que va a 0 quan n tendeix a ∞. Com veurem ara, aquesta fita es pot millorar en casos particulars.

Teorema 5.1.8. *Teorema de Chernoff.*
Siguin $\{X_i\}_{i=1}^n$ variables aleatòries independents $X_i \sim \mathrm{Be}(p_i)$. Sigui $X = \sum\limits_{i=1}^{n} X_i$, $\mu = \mathbb{E}[X] = \sum\limits_{i=1}^{n} p_i$. Aleshores es tenen les següents fites:

i) Cua superior: $p\left(X-\mu \ge \delta\mu\right) = p\left(X \ge (1+\delta)\mu\right) \le e^{-\frac{\delta^2}{2+\delta}\mu}$, amb $\delta \ge 0$.

ii) Cua inferior: $p\left(X-\mu \leq -\delta\mu\right) = p\left(X \leq (1-\delta)\mu\right) \leq e^{-\frac{\delta^2}{2}\mu}$, amb $\delta \in (0,1)$.

Demostració. Demostrarem només la cua superior. Calculem la funció generadora de moments de X:

$$M_X(s) = \prod_{i=1}^{n} M_{X_i}(s) = \prod_{i=1}^{n} \left(p_i e^s + (1-p_i)\right) = \prod_{i=1}^{n} \left(1 + p_i\left(e^s - 1\right)\right).$$

Que, per $s \in \mathbb{R}$ i $p_i \in [0,1]$, fent servir $1+y \leq e^y$ és menor o igual a

$$\prod_{i=1}^{n} e^{p_i(e^s-1)} = e^{\sum\limits_{i=1}^{n} p_i(e^s-1)} = e^{\mu(e^s-1)}$$

Si ara usem Markov, per $a \geq 0$

$$p(X \geq a) \overset{s\in\mathbb{R}^+}{=} p(sx \geq sa) = p(e^{sx} \geq e^{sa}) \overset{\text{Markov}}{\leq} \frac{\mathbb{E}\left[e^{sx}\right]}{e^{sa}} \leq \frac{e^{\mu(e^s-1)}}{e^{as}}$$

Fent ara $a = (1+\delta)\mu$:

$$p\left(X \geq (1+\delta)\mu\right) \leq e^{\left[(e^s-1)\mu - \left((1+\delta)\mu\right)s\right]} = f(s) = e^{g(s)} \tag{5.1}$$

Volem trobar $s \in \mathbb{R}$, $s > 0$ tal que optimitzi el valor de $f(s)$:

$$g'(s) = e^s\mu - (1+\delta)\mu = 0 \implies s = \log(1+\delta).$$

I, com $g''(s) = e^s\mu$,

$$g''(\log(1+\delta)) = e^{\log(1+\delta)}\mu = (1+\delta)\mu > 0.$$

És a dir, $\log(1+\delta)$ és el mínim que buscàvem. Si avaluem ara a la fita 5.1 obtinguda anteriorment, tenim que

$$p\left(X \geq (1+\delta)\mu\right) \leq e^{\delta\mu - (1+\delta)\mu\log(1+\delta)} = e^{\mu\left(\delta - (1+\delta)\log(1+\delta)\right)}.$$

A més, utilitzant que $\log(1+x) \geq \frac{x}{1+\frac{x}{2}}$ per $x \geq 0$

$$e^{\mu\left(\delta-(1+\delta)\log(1+\delta)\right)} \leq e^{\mu\left(\delta - (1+\delta)\left(\frac{\delta}{1+\frac{\delta}{2}}\right)\right)} = e^{\mu\frac{\delta + \frac{\delta^2}{2} - \delta - \delta^2}{1+\frac{\delta}{2}}} = e^{-\mu\frac{\delta^2}{2+\delta}}.$$

$\square$

Observació 5.1.9. Per demostrar la fita per a la cua inferior, cal fer el següent:

$$p(X \leq a) \overset{s\geq 0}{=} p(Xs \leq as) = p(e^{sX} \leq e^a s) = p(e^{-sX} \geq e^{-as}) \leq \frac{\mathbb{E}\left[e^{-sX}\right]}{e^{-as}} \leq \frac{e^{\mu\left(e^{-s}-1\right)}}{e^{-as}}.$$

I optimitzar-ho per $a = (1-\delta)\mu$.

Corol·lari 5.1.10. Amb la notació anterior i amb $\delta \in (0,1)$:

$$p\left(|X-\mu| \geq \delta\mu\right) \leq 2e^{-\frac{\delta^2}{3}\mu}.$$

Que fem notar que és una millor fita que Txebixov.

Demostració. Sigui $0 < \delta < 1$. Pel teorema de Chernoff (5.1.8) sabem que

$$p(X - \mu \geq \delta\mu) = p\left(X \geq (1+\delta\mu)\right) \leq e^{-\frac{\delta^2}{2+\delta}\mu} \overset{\mu>0,\,\delta\in(0,1)}{\leq} e^{-\frac{\delta^2}{3}\mu}$$

i que

$$p(X - \mu \leq -\delta\mu) = p(X \leq (1-\delta\mu) \leq e^{-\frac{\delta^2}{2}\mu} \overset{\mu>0}{\leq} e^{-\frac{\delta^2}{3}\mu}.$$

I, finalment,

$$p\left(|X-\mu| \geq \delta\mu\right) = p\left(X - \mu \geq \delta\mu\right) + p\left(X - \mu \leq -\delta\mu\right) \leq 2e^{-\frac{\delta^2}{3}\mu}.$$

$\square$

Observació 5.1.11. Hem vist que donada X amb moments $\{a_i\}_{i\geq 1}$, podem construir la funció generadora de moments $M_X(s)$. Una pregunta a fer-se és si es pot reconstruir la variable X donats els seus moments. Aquest problema es coneix com el problema de Hamburger i hi ha criteris que garanteixen que es pot fer. Veiem-ne un exemple.

Teorema 5.1.12. *Condició de Carleman.*
Amb la notació de l'observació anterior, si $a_{2n} \neq 0\ \forall n$ i $\sum\limits_{n\geq 1} a_{2n}^{-\frac{1}{2n}} = \infty$, aleshores els moments determinen unívocament la variable aleatòria.

5.2 Funcions característiques

Hem vist que la funció generadora de moments introdueix un problema tècnic (analític) perquè molts cops no la podrem avaluar. Això ho podrem solucionar prenent la variable aleatòria $s = it$, $t \in \mathbb{R}$. En el cas absolutament continu,

$$\mathbb{E}\left[e^{sX}\right] = \int_{\mathbb{R}} e^{sx} f_X(x)\,\mathrm{d}x \overset{s=it}{=} \int_{\mathbb{R}} e^{itx} f_X(x)\,\mathrm{d}x.$$

I

$$\int_{\mathbb{R}} \left|e^{itx} f_X(x)\right| \mathrm{d}x = \int_{\mathbb{R}} |f_X(x)|\,\mathrm{d}x = 1,$$

d'on $E\left[e^{itX}\right]$ està ben definit.

Definició 5.2.1. La funció característica de X és

$$\begin{aligned} \phi_X \colon \mathbb{R} &\to \mathbb{C} \\ t &\mapsto \mathbb{E}\left[e^{itX}\right] \end{aligned}$$

Observació 5.2.2. Les propietats de la funció generadora de moments es tradueixen immediatament a la funció característica. En particular, si X, Y són independents,

$$\phi_{X+Y}(t) = \phi_X(t)\phi_Y(t)$$

Observació 5.2.3. En el cas absolutament continu, $\phi_X(t)$ és la transformada de Fourier de $f_X(t)$.

Exemple 5.2.4.

1. $X \sim \text{Geom}(p) \implies \phi_X(t) = \frac{p}{e^{-it}-(1-p)}$
2. $X \sim \text{Poisson}(\lambda) \implies \phi_X(t) = e^{\lambda\left(e^{it}-1\right)}$
3. $X \sim \text{U}[0,1] \implies \phi_X(t) = \frac{1}{it}\left(e^{it}-1\right)$

Exemple 5.2.5. $X \sim \text{Cauchy}$, $f_X(x) = \frac{1}{\pi(1+x^2)}$. Llavors,

$$M_X(s) = \int_{\mathbb{R}} \frac{e^{sx}}{\pi(1+x^2)}\,\mathrm{d}x = \begin{cases} 1 & \text{si } s = 0 \\ \text{no està definida} & \text{si } s \in \mathbb{R},\ s \neq 0 \end{cases}.$$

D'altra banda, $\phi_X(t) = \int_{-\infty}^{\infty} \frac{e^{itx}}{\pi(1+x^2)}\,\mathrm{d}x = e^{-|t|}$

Exemple 5.2.6. $X \sim \text{N}(0,1)$:

$$\begin{aligned}\phi_X(t) &= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{itx}e^{-\frac{1}{2}x^2}\,\mathrm{d}x \overset{u=it}{=} \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{-\frac{1}{2}(u-x)^2+\frac{1}{2}u^2}\,\mathrm{d}x = \\ &= \frac{1}{\sqrt{2\pi}}e^{\frac{1}{2}u^2}\int_{\mathbb{R}} e^{-\frac{1}{2}(u-x)^2}\,\mathrm{d}x = e^{\frac{1}{2}u^2} = e^{-\frac{1}{2}t^2},\end{aligned}$$

on hem usat que $\int_{\mathbb{R}} e^{-\frac{1}{2}(u-x)^2} = \sqrt{2\pi}$. Veurem com es demostra per un cas més senzill on la demostració és anàloga:

$$\left(\int_{\mathbb{R}} e^{-x^2}\,\mathrm{d}x\right)^2 = \left(\int_{\mathbb{R}} e^{-x^2}\,\mathrm{d}x\right)\left(\int_{\mathbb{R}} e^{-y^2}\,\mathrm{d}y\right) = \int_{\mathbb{R}^2} e^{-x^2-y^2}$$

Que, fent un canvi de variables a polars, és

$$\int_0^{2\pi}\int_0^{\infty} e^{-R^2}\left(-\frac{2R}{2}\right)\mathrm{d}R\,\mathrm{d}\theta = 2\pi\left[-\frac{e^{-R^2}}{2}\right]_0^{\infty} = \frac{2\pi}{2} = \pi \implies \int_{\mathbb{R}} e^{-x^2}\,\mathrm{d}x = \sqrt{\pi}.$$

Per tant, si $Y \sim \text{N}(\mu,\sigma^2)$,

$$\phi_Y(t) = \phi_{\sigma X+\mu}(t) \overset{5.1.2}{=} e^{i\mu t}\phi_X(\sigma t) = e^{i\mu t-\frac{\sigma^2t^2}{2}},$$

i, finalment, si $X \sim \text{N}(\mu_1,\sigma_1^2)$, $Y \sim \text{N}(\mu_2,\sigma_2^2)$ independents,

$$\phi_{X+Y}(t) = e^{i(\mu_1+\mu_2)t-\frac{(\sigma_1^2+\sigma_2^2)t^2}{2}}.$$

La pregunta que un es pot fer ara és si a partir de $\phi_{X+Y}(t)$ podem assegurar que $X+Y$ és normal. Amb el proper resultat veurem que, en general, la funció característica determina unívocament la variable aleatòria X.

Teorema 5.2.7. *Teorema d'inversió.*
Sigui X una variable aleatòria amb probabilitat induïda P_X i funció característica $\phi_X(t)$. Aleshores, si $a < b$

$$P_X((a,b)) + \frac{1}{2}(P_X(\{a\}) + P_X(\{b\})) = \lim_{T\to\infty}\frac{1}{2\pi}\int_{-T}^{T}\frac{e^{-ita}-e^{-itb}}{it}\phi_X(t)\,\mathrm{d}t$$

Demostració. Per a avaluar la integral del teorema, usarem la següent integral:

$$\lim_{T\to\infty}\int_0^T \frac{\sin(tc)}{t}\,\mathrm{d}t = \begin{cases} -\frac{\pi}{2} & c<0 \\ 0 & c=0 \\ \frac{\pi}{2} & c>0 \end{cases}$$

Fem el càlcul de la integral original:

$$\frac{1}{2\pi}\int_{-T}^{T}\frac{e^{-ita}-e^{-itb}}{it}\phi_X(t)\,\mathrm{d}t = \frac{1}{2\pi}\int_{-T}^{T}\frac{e{-ita}-e^{-itb}}{it}\left(\int_{\mathbb{R}} e^{itx}\,\mathrm{d}p_x\right)\mathrm{d}t =$$
$$= \frac{1}{2\pi}\int_{-T}^{T}\int_{\mathbb{R}}\frac{e^{it(x-a)}-e^{it(x-b)}}{it}\,\mathrm{d}p_x\,\mathrm{d}t = \frac{1}{2\pi}\int_{\mathbb{R}}\int_{-T}^{T}\frac{e^{it(x-a)}-e^{it(x-b)}}{it}\,\mathrm{d}t\,\mathrm{d}P_X.$$

Observem ara que

$$\int_{-T}^{T}\frac{e^{itc}}{it}\,\mathrm{d}t = 2\int_0^T\frac{\sin(tc)}{t}\,\mathrm{d}t,$$

perquè $\sin(tc) = \frac{e^{itc}-e^{-itc}}{2i}$. I, substituint a l'expressió anterior, tenim que és igual a

$$= \frac{1}{\pi}\int_{\mathbb{R}}\int_0^T\left(\frac{\sin(t(x-a))}{t} - \frac{\sin(t(x-b))}{t}\right)\mathrm{d}t\,\mathrm{d}P_X.$$

Definim ara $g(x,T,a,B) = \int_0^T\left(\frac{\sin\left(t(x-a)\right)}{t} - \frac{\sin\left(t(x-b)\right)}{t}\right)\mathrm{d}t$. Tenim que

$$\lim_{T\to\infty} g(x,T,a,b) = \begin{cases} 0 & x<a<b \\ \frac{\pi}{2} & x=a<b \\ \pi & a<x<b \\ \frac{\pi}{2} & a<x=b \\ 0 & a<b<x \end{cases}$$

Finalment, calculem el límit de l'integral anterior

$$\lim_{T\to\infty}\int_{\mathbb{R}}\frac{g(x,T,a,b)}{\pi}\,\mathrm{d}P_X \overset{\text{TCD}}{=} \int_{\mathbb{R}}\lim_{T\to\infty}\frac{g(x,T,a,b)}{\pi}\,\mathrm{d}P_X =$$
$$= \frac{1}{2}(P_X(\{a\})+P_X(\{b\})) + P_X((a,b)).$$

□

Corol·lari 5.2.8. $\phi_X(t)$ caracteritza λ-g.a. la variable aleatòria X (la seva distribució), on λ és la mesura de Lebesgue a $\mathbb{R}$.

Demostració. Si $\phi_X(t) = \phi_Y(t)$ per dues variables aleatòries X, Y aleshores, pel teorema d'inversió, $\forall a<b$:

$$\frac{1}{2}(P_X(\{a\})+P_X(\{b\})) + P_X((a,b)) = \frac{1}{2}(P_Y(\{a\})+P_Y(\{b\})) + P_Y((a,b)).$$

I, quan $a\to-\infty$, veiem que el terme $P_X(\{a\})$ va a 0:

$$\lim_{a\to-\infty}P_X(\{a\}) = \lim_{a\to-\infty}P(X=a) = \lim_{a\to-\infty}F_X(a) - \lim_{h\to0^+}F_X(a-h) = 0.$$

D'on, estudiant el límit de l'expressió anterior a $a \to \infty$, tenim

$$\frac{1}{2}P_X(\{b\}) + P_X((-\infty, b)) = \frac{1}{2}P_Y(\{b\}) + P_Y((-\infty, b)).$$

És a dir,

$$F_X(b) + \frac{1}{2}P_X(\{b\}) = F_Y(b) + \frac{1}{2}P_Y(\{b\}).$$

Com $P_X(\{b\})$ i $P_Y(\{b\})$ són diferents de zero en un conjunt numerable (són punts de discontinuïtat de les funcions creixents F_X i F_Y), tenim que $F_X = F_Y$ λ-g.a.. $\square$

5.2.1 Propietats addicionals de la funció característica

Proposició 5.2.9. Sigui X una variable aleatòria amb funció característica $\phi_X(t) = \mathbb{E}\left[e^{itx}\right]$. Aleshores,

i) $\phi_X(0) = 1$ i $|\phi_X(t)| \leq 1$ $\forall t \in \mathbb{R}$.

ii) $\phi_X(t)$ és uniformement contínua.

iii) $\forall t_1, \ldots, t_n \in \mathbb{R}$, $\forall z_1, \ldots, z_n \in \mathbb{C}$, $\sum_{j,k} \phi_X(t_j - t_k) z_j \bar{z}_k \geq 0$ $(\forall n \in \mathbb{N})$.

Demostració.

i) És immediat.

ii) Volem fitar $|\phi_X(x) - \phi_X(y)|$ quan $|x - y| < \varepsilon$. Fem $x = t + h$, $y = t$ i volem fitar $|\phi_X(t + h) - \phi_X(t)|$ en termes de h:

$$|\phi_X(t+h) - \phi_X(t)| = \left|\mathbb{E}\left[e^{i(t+h)X} - e^{itX}\right]\right| = \left|\mathbb{E}\left[e^{itX}\left(e^{ihX} - 1\right)\right]\right| \leq$$

$$\leq \mathbb{E}\left[\left|e^{itX}\right|\left|e^{ihX} - 1\right|\right] \overset{Y(h) = \left|e^{ihX} - 1\right|}{=} \mathbb{E}\left[Y(h)\right]$$

Observem que $0 \leq Y(h) \leq 2$. A més, $\lim_{h \to 0} Y(h) = 0$. Aplicant el teorema de convergència dominada,

$$\lim_{h \to 0} \mathbb{E}\left[Y(h)\right] = \mathbb{E}\left[\lim_{h \to 0} Y(h)\right] = 0.$$

D'on $\lim_{h \to 0} |\phi_X(t + h) - \phi_X(t)| = 0$ amb independència de t.

iii) Reescrivim l'expressió donada:

$$\sum_{j,k} \phi_X(t_j - t_k) z_j \bar{z}_k = \sum_{j,k} \mathbb{E}\left[e^{i(t_j - t_k)X}\right] z_j \bar{z}_k = \mathbb{E}\left[\sum_{j,k} e^{it_jX} z_j e^{-it_kX} \bar{z}_k\right] =$$

$$= \mathbb{E}\left[\left(\sum_j e^{it_jX} z_j\right)\left(\sum_k e^{-it_kX} \bar{z}_k\right)\right] = \mathbb{E}\left[\left(\sum_j e^{it_jX} z_j\right)\overline{\left(\sum_k e^{it_kX} z_k\right)}\right] =$$

$$= \mathbb{E}\left[\left|\sum_j e^{it_jX} z_j\right|^2\right] \geq 0$$

$\square$

Observació 5.2.10. Les propietats anteriors són condicions suficients per assegurar que una funció donada és funció característica d'una variable aleatòria (teorema de Bochner).

Observació 5.2.11. Si (X, Y) és un vector de variables aleatòries, podem definir la seva funció característica com

$$\phi_{(X,Y)}(t_1, t_2) = \mathbb{E}\left[e^{it_1X}e^{it_2Y}\right]$$

5.3 Famílies exponencials: exemples

Veurem ara que podem tractar de manera unificada moltes variables aleatòries que han aparegut fins ara.

Definició 5.3.1. Una família de variables aleatòries és exponencial amb paràmetres $\vec{\theta} = (\theta_1, \dots, \theta_n)$ si la seva funció de densitat (o funció de probabilitat) és de la forma:

$$p(X, \vec{\theta}) = p(X|\vec{\theta}) = f(x, \vec{\theta}) = g(x) \exp\left(\sum_{i=1}^{n} \theta_i t_i(x) - c(\vec{\theta})\right)$$

amb $g(x) \geq 0$ i $\{t_i(x)\}$ linealment independents.

Exemple 5.3.2. Bernoulli per $x \in \{0, 1\}$:

$$p(x, p) = p(x|p) = p^x(1-p)^{1-x} = \exp\left(x \log p + (1-x)\log(1-p)\right) =$$
$$= \exp\left(x \log\left(\frac{p}{1-p}\right) + \log(1-p)\right).$$

És a dir, $\theta_1 = \log\left(\frac{p}{1-p}\right)$, $t_1(x) = x$, $g(x) = 1$, i $c(\theta) = -\log(1-p) = \log\left(\frac{1}{1-p}\right) = \log\left(1 + e^{\theta_1}\right)$.

Exemple 5.3.3. Normal:

$$f(x|\mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) = \frac{1}{\sqrt{2\pi}} \exp\left(-\log\sigma - \frac{x^2}{2\sigma^2} + \frac{x\mu}{\sigma^2} - \frac{\mu^2}{2\sigma^2}\right) =$$
$$= \frac{1}{\sqrt{2\pi}} \exp\left(x\frac{\mu}{\sigma^2} + x^2\left(-\frac{1}{2\sigma^2}\right) - \left(\log\sigma + \frac{\mu^2}{2\sigma^2}\right)\right).$$

És a dir, $\theta_1 = \frac{\mu}{\sigma^2}$, $\theta_2 = -\frac{1}{2\sigma^2}$, $t_1(x) = x$, $t_2(x) = x^2$ i

$$c(\theta_1, \theta_2) = \log\sigma + \frac{\mu^2}{2\sigma^2} = \frac{1}{2}\log\left(\sigma^2\right) + \frac{\mu^2}{2\sigma^2} =$$
$$= \frac{1}{2}\log\left(-\frac{1}{2\theta_2}\right) + \frac{\theta_1^2}{4\theta_2^2}(-\theta_2) = \frac{1}{2}\log\left(-\frac{1}{2\theta_2}\right) - \frac{\theta_1^2}{4\theta_2}$$

Proposició 5.3.4. $\mathbb{E}\left[t_i(X)\right] = \frac{\partial}{\partial\theta_i} c(\theta_1, \dots, \theta_n)$ en una família exponencial.

Demostració. Farem el cas absolutament continu. Observem que $\int_{\mathbb{R}} f(x, \vec{\theta})\, \mathrm{d}x = 1$, de manera que si integrem a tots dos costats de la definició de $f(x, \vec{\theta})$ per famílies exponencials obtenim:

$$\exp c(\vec{\theta}) = \int_{\mathbb{R}} g(x) \exp\left(\sum_{i=1}^{n} \theta_i t_i(x)\right) \mathrm{d}x. \tag{5.2}$$

I, derivant a tots dos costats respecte de θ_i, tenim que:

$$\frac{\partial}{\partial\theta_i}c(\vec{\theta}) = \frac{\frac{\partial}{\partial\theta_i}\int_{\mathbb{R}} g(x)\exp\left(\sum_{i=1}^n \theta_i t_i(x)\right)\mathrm{d}x}{\exp\left(c(\vec{\theta})\right)} \stackrel{\text{TCD}}{=}$$
$$\stackrel{\text{TCD}}{=} \frac{\int_{\mathbb{R}} g(x)t_i(x)\exp\left(\sum_{i=1}^n \theta_i t_i(x)\right)\mathrm{d}x}{\exp\left(c(\vec{\theta})\right)} = \mathbb{E}\left[t_i(x)\right].$$

□

Definició 5.3.5. Una família exponencial es diu natural si $\exists k\colon t_k(x) = x$.

Proposició 5.3.6. Sigui X una variable aleatòria d'una família exponencial natural. Aleshores $\phi_X(t) = \exp\left(c(\theta_1 + it, \dots, \theta_n) - c(\vec{\theta})\right)$

Demostració. Calculem la integral:

$$\phi_X(t) = \mathbb{E}\left[e^{itX}\right] = \int_{\mathbb{R}} e^{itx} g(x)\exp\left(\theta_1 x + \sum_{r=2}^n \theta_r t_r(x) - c(\vec{\theta})\right)\mathrm{d}x =$$
$$\int_{\mathbb{R}} g(x)\exp\left((\theta_1 + it)x + \sum_{r=2}^n \theta_r t_r(x) - c(\vec{\theta})\right)\mathrm{d}x.$$

Que, utilitzant l'equació 5.2 que hem fet servir a la demostració de la proposició anterior, tenim que és igual a

$$\exp\left(c(\theta_1 + it, \dots, \theta_n) - c(\vec{\theta})\right)$$

□

Exemple 5.3.7. En el cas de la normal, $\theta_1 = \frac{\mu}{\sigma^2}$, $\theta_2 = -\frac{1}{2\sigma^2}$. Aleshores

$$\phi_X(t) = \exp\left(c(\theta_1 + it, \theta_2) - c(\theta_1, \theta_2)\right)$$

i

$$c(\theta_1, \theta_2) = \frac{\mu^2}{2\sigma^2} + \log(\sigma) = -\frac{\theta_1^2}{4\theta_2} + h(\theta_2).$$

De manera que

$$c(\theta_1 + it, \theta_2) - c(\theta_1, \theta_2) = -\frac{(\theta_1 + it)^2}{4\theta_2} + \frac{\theta_1^2}{4\theta_2} = -\frac{2\theta_1 it}{4\theta_2} + \frac{t^2}{4\theta_2} = \mu it - \sigma^2\frac{t^2}{2}.$$

Per tant, $\phi_X(t) = \exp\left(it\mu - \sigma^2\frac{t^2}{2}\right)$.

Tema 6

Convergència de variables aleatòries

6.1 Modes de convergència. Implicacions

Sigui $(\Omega, \mathcal{A}, p)$ un espai de probabilitat i $\{X_n\}_{n\geq 1}$ una reunió de variables aleatòries. Volem respondre les següents preguntes:

- Establir la noció de límit d'una successió.
- Podem establir resultats *universals* en quant als límits?

Pel primer punt, establirem essencialment quatre modes diferents de convergència.

Definició 6.1.1. Direm que $\{X_n\}_{n\geq 1}$ convergeix quasi-segurament (almost-surely) cap a X si

$$p\left(\left\{\omega \in \Omega \text{ t.q. } X_n(\omega) \xrightarrow{n} X(\omega)\right\}\right) = 1.$$

Si és així, escriurem $X_n \xrightarrow{\text{q-s}} X$.

Observació 6.1.2. $A = \left\{\omega \in \Omega \text{ t.q. } X_n(\omega) \xrightarrow{n} X(\omega)\right\}$ és efectivament un succés i la definició te sentit:

Definim el succés $A_n(m) = \left\{\omega \in \Omega \text{ t.q. } |X_n(\omega) - X(\omega)| < \frac{1}{m}\right\}$ $\forall n, m \geq 1$. Definim el succés $A(m) = \liminf_n A_n(m) = \{\omega \in \Omega \text{ t.q. } \omega \in A_n(m)\, \forall n \geq n_0\}$. Finalment, $A = \bigcap_{m\geq 1} A(m)$ i, per tant, és un succés.

Definició 6.1.3. Donada $\{X_n\}_{n\geq 1}$ successió de variables aleatòries i X variable aleatòria, direm que $\{X_n\}_{n\geq 1}$ convergeix en mitjana d'ordre r ($r \in \mathbb{R}_{>0}$) si

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

i escriurem $X_n \xrightarrow{r} X$. Si $r = 1$, direm que convergeix en mitjana, i si $r = 2$, en mitjana quadràtica.

Definició 6.1.4. Donada $\{X_n\}_{n\geq 1}$ successió de variables aleatòries i X variable aleatòria, direm que $\{X_n\}_{n\geq 1}$ convergeix en probabilitat cap a X si

$$p(|X_n - X| > \varepsilon) \xrightarrow{n} 0 \quad \forall \varepsilon > 0.$$

Escriurem $X_n \xrightarrow{p} X$.

Definició 6.1.5. Sigui $\{X_n\}_{n\geq 1}$ successió de variables aleatòries i X variable aleatòria. Direm que $\{X_n\}_{n\geq 1}$ convergeix en distribució cap a X si

$$F_{X_n} \xrightarrow{n} F_X(x) \quad \forall x \in \mathbb{R} \text{ t.q. } F_X \text{ és contínua en } x.$$

Escriurem $X_n \xrightarrow{d} X$.

La següent tasca serà demostrar les següents implicacions entre modes de convergència:

$$\begin{tikzcd}
 & X_n \xrightarrow{\text{q-s}} X \arrow[dr, Rightarrow] & & \\
 & & X_n \xrightarrow{p} X \arrow[r, Rightarrow] & X_n \xrightarrow{d} X \\
X_n \xrightarrow{r} X \arrow[r, Rightarrow] & X_n \xrightarrow{s} X \arrow[ur, Rightarrow] & &
\end{tikzcd}$$

On $r \geq s \geq 1$ i cap de les implicacions recíproques és certa.

Proposició 6.1.6. Si $X_n \xrightarrow{p} X$, aleshores $X_n \xrightarrow{d} X$.

Demostració. Com $X_n \xrightarrow{p} X$, sabem que $\forall \varepsilon > 0$ $p\left(|X_n - X| > \varepsilon\right) \xrightarrow{n} 0$ i volem veure que per tot punt $x \in \mathbb{R}$ punt de continuïtat de F_x, aleshores $F_{X_n}(x) \xrightarrow{n} F_X(x)$. Observem que, en general, x no és punt de continuïtat de totes les F_{X_n}. Per una banda,

$$\begin{aligned}
F_{X_n}(x) = p\left(X_n \leq x\right) &\overset{\forall\varepsilon\geq 0}{=} p\left(X_n \leq x,\, X \leq x+\varepsilon\right) + p\left(X_n \leq x,\, X > x+\varepsilon\right) \leq \\
&\leq p\left(X_n \leq x,\, X \leq x+\varepsilon\right) + p\left(|X_n - X| > \varepsilon\right) = F_X(x+\varepsilon) + p\left(|X_n - X| > \varepsilon\right).
\end{aligned}$$

Similarment, per tot $\varepsilon > 0$,

$$\begin{aligned}
F_X(x-\varepsilon) = p\left(X \leq x-\varepsilon\right) &= p\left(X \leq x-\varepsilon,\, X_n \leq x\right) + p\left(X \leq x - \varepsilon \leq x,\, X_n > x\right) \leq \\
&\leq p\left(X_n \leq x\right) + p\left(|X_n - X| > \varepsilon\right) = F_{X_n}(x) + p\left(|X_n - X| > \varepsilon\right).
\end{aligned}$$

Com a priori $\lim\limits_{n\to\infty} F_{X_n}(x)$ pot no existir, cal usar $\liminf$ i $\limsup$. Utilitzant les fites anteriors, $\forall \varepsilon > 0$

$$\begin{aligned}
F_X(x-\varepsilon) &\overset{p\left(|X_n - X| > \varepsilon\right) \xrightarrow{n} 0}{=} \liminf \Big(F_X(x-\varepsilon) - p\left(|X_n - X| > \varepsilon\right)\Big) \leq \liminf F_{X_n}(x) \leq \\
&\leq \limsup F_{X_n}(x) \leq \limsup F_X(x+\varepsilon) + p\left(|X_n - x| > \varepsilon\right) = F_X(x+\varepsilon).
\end{aligned}$$

Per tant, $F_X(x-\varepsilon) \leq \liminf\limits_n F_{X_n}(x) \leq \limsup\limits_n F_{X_n}(x) \leq F_X(x+\varepsilon)$. Finalment, si fem tendir ε a 0, per ser F_X contínua en x $\lim\limits_{\varepsilon\to 0} F_X(x-\varepsilon) = \lim\limits_{\varepsilon\to 0} F_X(x+\varepsilon) = F_X(x)$ i, per tant,

$$F_X(x) \leq \liminf_n F_{X_n}(x) \leq \limsup_n F_{X_n}(x) \leq F_X(x).$$

D'on $\liminf_n F_{X_n}(x) = \limsup_n F_{X_n}(x) = \lim F_{X_n}(x) = F_X(x)$. □

Observació 6.1.7. En general, $X_n \xrightarrow{d} X \not\Longrightarrow X_n \xrightarrow{p} X$. Posem-ne un exemple:

Prenem $Y \sim \text{Bern}\left(\frac{1}{2}\right)$ i $X_n = Y$ $\forall n \geq 1$. Sigui $X = 1 - Y$, que també és una Bernoulli:

- És clar que $F_{X_n} = F_Y(x) = F_X(x)$ i trivialment $X_n \xrightarrow{d} X$.
- Ara bé, $X_n \overset{p}{\not\to} X$ perquè $\forall \varepsilon > 0$

$$p\left(|X_n - X| > \varepsilon\right) = p\left(|Y - (1-Y)| > \varepsilon\right) = p(|2Y-1|) = p(1) > \varepsilon.$$

Que és 1 si $\varepsilon < 1$, cosa que es una contradicció amb la convergència en probabilitat.

Proposició 6.1.8. $X_n \xrightarrow{1} X \implies X_n \xrightarrow{p} X$.

Demostració. Sabem que $\mathbb{E}\left[|X_n - X|\right] \xrightarrow{n} 0$ i volem veure que $\forall \varepsilon > 0$, $p\left(|X_n - X| > \varepsilon\right)$ tendeix a 0. Usant Markov, $\forall \varepsilon > 0$,

$$0 \leq \lim_{n\to\infty} p\left(|X_n - X| > \varepsilon\right) \leq \lim_{n\to\infty} \frac{\mathbb{E}\left[|X_n - X|\right]}{\varepsilon} = 0.$$

De manera que $X_n \xrightarrow{p} X$. □

Observació 6.1.9. En general, $X_n \xrightarrow{p} X \not\Longrightarrow X_n \xrightarrow{1} X$. Vegem-ne un exemple. Prenem $X_n = \begin{cases} n^3 & \text{amb probabilitat } \frac{1}{n^2}. \\ 0 & \text{amb probabilitat } 1 - \frac{1}{n^2}. \end{cases}$

- $X_n \xrightarrow{p} X$: $\forall \varepsilon > 0$, calculem $p\left(|X_n - X| > \varepsilon\right) = p\left(|X_n| > \varepsilon\right) = p(X_n > \varepsilon) \overset{n\geq 1,\, \varepsilon<1}{=} \frac{1}{n^2} \xrightarrow{n} 0$.
- $\mathbb{E}\left[|X_n - X|\right] = \mathbb{E}\left[X_n\right] = \frac{1}{n^2}n^3 + 0\left(1 - \frac{1}{n^2}\right) = n \xrightarrow{n} \infty \implies X_n \overset{1}{\not\to} X$.

Proposició 6.1.10. Si $r \geq s \geq 1$, aleshores $X_n \xrightarrow{r} X \implies X_n \xrightarrow{s} X$.

Demostració. Volem veure que si $\mathbb{E}\left[|X_n - X|^r\right] \xrightarrow{n} 0$, aleshores $\mathbb{E}\left[|X_n - X|^s\right] \xrightarrow{n} 0$. De fet, demostrarem el següent: sigui Z una variable aleatòria tal que $\mathbb{E}\left[|Z|^p\right] < \infty \ \forall p$; aleshores si $r \geq s \geq 1$, $||Z||_r = \mathbb{E}\left[|Z|^r\right]^{\frac{1}{r}} \geq \mathbb{E}\left[|Z|^s\right]^{\frac{1}{s}} = ||Z||_s$. Ho demostrem amb la desigualtat de Hölder: si prenem $p, q \in \mathbb{R}_{>1}$ i $\frac{1}{p} + \frac{1}{q} = 1$

$$\mathbb{E}\left[|XY|\right] \leq \mathbb{E}\left[|X|^p\right]^{\frac{1}{p}} \mathbb{E}\left[|Y|^q\right]^{\frac{1}{q}} \quad \left(\text{amb } \mathbb{E}\left[|X|^p\right], \mathbb{E}\left[|Y|^q\right] < \infty\right).$$

Per tant, fent $Y = 1$, $\mathbb{E}\left[|X|\right] \leq \mathbb{E}\left[|X|^p\right]^{\frac{1}{p}} \overset{X=Z^s}{\Longrightarrow} \mathbb{E}\left[|Z|^s\right] \leq \mathbb{E}\left[|Z|^{sp}\right]^{\frac{1}{p}}$. Triem $p > 1$ tal que $sp = r$. Aleshores $\mathbb{E}\left[|Z|^s\right] \leq \mathbb{E}\left[|Z|^r\right]^{\frac{s}{r}}$ cosa que implica $\mathbb{E}\left[|Z|^s\right]^{\frac{1}{s}} \leq \mathbb{E}\left[|Z|^r\right]^{\frac{1}{r}}$. □

Observació 6.1.11. En general, si $r \geq s > 1$, no és cert que $X_n \xrightarrow{s} X \implies X_n \xrightarrow{r} X$. Posem-ne un exemple:

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

En aquest cas $\mathbb{E}\left[|X_n - X|^s\right] = \mathbb{E}\left[|X_n|^s\right] = n^s n^{-\frac{r+s}{2}} = n^{\frac{s-r}{2}} \xrightarrow{n\to\infty} 0$ i $\mathbb{E}\left[|X_n - X|^r\right] = \mathbb{E}\left[|X_n|^r\right] = n^r n^{-\frac{r+s}{2}} = n^{\frac{r-s}{2}} \xrightarrow{n\to\infty} \infty$.

Falta finalment demostrar l'implicació $X_n \xrightarrow{q-s} X \implies X_n \xrightarrow{p} X$. Comencem amb el següent lema:

Lema 6.1.12. Sigui $\varepsilon > 0$ i sigui $A_n(\varepsilon) = \{\omega \in \Omega \text{ t.q. } |X_n(\omega) - X(\omega)| > \varepsilon\}$. Sigui $B_n(\varepsilon) = \bigcup_{m \geq n} A_m(\varepsilon)$. Aleshores, $X_n \xrightarrow{q-s} X \iff \forall\varepsilon > 0 \lim P(B_n(\varepsilon)) = 0$.

Demostració. Prenem $C = \left\{\omega \in \Omega \text{ t.q. } X_n(\omega) \xrightarrow{n} X(\omega)\right\}$, que és un succés (vist a 6.1.2). Sigui $A(\varepsilon) = \limsup_n A_n(\varepsilon) = \{\omega \in \Omega \text{ t.q. } \omega \text{ pertany a un nombre infinit dels } A_n(\varepsilon)\}$. Observem que $P(C) = 1 \iff P(A(\varepsilon)) = 0\ \forall\varepsilon$ (per les definicions de C i $A(\varepsilon)$). Com ara $B_n(\varepsilon)$ és una successió de successos decreixent, té límit amb:

$$\lim B_n(\varepsilon) = \bigcap_{n\geq 1} B_n(\varepsilon) = \bigcap_{n \geq 1} \bigcup_{m \geq n} A_m(\varepsilon) = \limsup_n A_n(\varepsilon) = A(\varepsilon).$$

Per tant, $X_n \xrightarrow{q-s} X \iff P(C) = 1 \iff P(A(\varepsilon)) = 0\ \forall\varepsilon > 0 \iff P(\lim_n B_n(\varepsilon)) = 0\ \forall\varepsilon > 0 \overset{B_n(\varepsilon) \text{ decreixent}}{\iff} \lim_n P(B_n(\varepsilon)) = 0\ \forall\varepsilon > 0$. ☐

Proposició 6.1.13. $X_n \xrightarrow{q-s} X \implies X_n \xrightarrow{p} X$.

Demostració. Com $X_n \xrightarrow{q-s} X$, aleshores pel lema anterior $\forall\varepsilon > 0$ es té $\lim_n p(B_n(\varepsilon)) = 0$. Finalment, prenent límits, es té per tot $\varepsilon > 0$ que

$$0 \leq \liminf p(A_n(\varepsilon)) \leq \limsup p(A_n(\varepsilon)) \leq \lim p(B_n(\varepsilon)) = 0,$$

I per tant,

$$\liminf p(A_n(\varepsilon)) = \limsup p(A_n(\varepsilon)) = \lim p(A_n(\varepsilon)) = 0.$$

☐

El següent resultat dóna un criteri per a veure la convergència quasi-segura a partir de les probabilitats $p(A_n(\varepsilon))$.

Proposició 6.1.14. $\forall\varepsilon > 0$, sigui $A_n(\varepsilon) = \{\omega \in \Omega \text{ t.q. } |X_n(\omega) - X(\omega)| > \varepsilon\}$. Aleshores

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

Demostració. Sigui $B_n(\varepsilon) = \bigcup_{n \geq m} A_m(\varepsilon)$, aleshores $p(B_n(\varepsilon)) = p\left(\bigcup_{n\geq m} A_m(\varepsilon)\right) \leq \sum_{m \geq n} p(A_m(\varepsilon))$. Com tenim que $\sum_{n\geq 1} p(A_n(\varepsilon)) < \infty\ \forall\varepsilon > 0$, sabem que $\lim_n \sum_{m\geq n} p(A_n(\varepsilon)) = 0$. Per tant, $\lim_n p(B_n(\varepsilon)) = 0 \overset{6.1.12}{\implies} X_n \xrightarrow{q-s} X$. ☐

Observació 6.1.15. No és cert que $X_n \xrightarrow{p} X \implies X_n \xrightarrow{q-s} X$. Prenguem $X_n = \begin{cases} 1 & \text{amb probabilitat } \frac{1}{n} \\ 0 & \text{amb probabilitat } 1 - \frac{1}{n} \end{cases}$ independents. Veurem $X_n \xrightarrow{p} 0$, però $X_n \overset{q-s}{\not\to} 0$. Fixat $\varepsilon > 0$,

$$p(|X_n - X| > \varepsilon) = p(|X_n| > \varepsilon) \overset{\varepsilon \leq 1}{=} \frac{1}{n} \xrightarrow{n} 0.$$

És a dir, les X_n convergeixen en probabilitat a 0. Ara bé, donat $\varepsilon > 0$,

$$p(B_n(\varepsilon)) = p\left(\bigcup_{m\geq n} A_m(\varepsilon)\right) = 1 - p\left(\bigcap_{m\geq n} \overline{A_m(\varepsilon)}\right) = 1 - p\left(\lim_{k\to\infty} \bigcap_{k\geq m\geq n} \overline{A_m(\varepsilon)}\right) =$$

$$= 1 - \lim_{k\to\infty} p\left(\bigcap_{k\geq m\geq n} \overline{A_m(\varepsilon)}\right) \overset{\text{indeps.}}{=} 1 - \lim_{k\to\infty} \prod_{k\geq m\geq n} p\left(\overline{A_m(\varepsilon)}\right) =$$

$$= 1 - \prod_{m\geq n} p\left(\overline{A_m(\varepsilon)}\right) \overset{\varepsilon\leq 1}{=} 1 - \prod_{m\geq n}\left(1 - \frac{1}{m}\right).$$

Si estudiem ara l'últim producte infinit tenim que

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

I per tant,

$$\prod_{m\geq n}\left(1-\frac{1}{m}\right) = 0 \implies p\left(B_n(\varepsilon)\right) = 1\ \forall n\in\mathbb{N},\ \varepsilon < 1.$$

D'on, si $\varepsilon < 1$,

$$\lim p(B_n(\varepsilon)) = 1 \implies X_n \overset{q-s}{\not\to} 0.$$

Malgrat que $X_n \xrightarrow{p} X$ no implica que $X_n \xrightarrow{q-s} X$ veurem ara que sí que és cert per una subsuccessió.

Proposició 6.1.16. Si $X_n \xrightarrow{p} X$ existeix una seqüència determinista $\{n_i\}_{i\geq 1}$ tal que $X_{n_i} \xrightarrow{q-s} X$.

Demostració. Sabem que $\forall\varepsilon > 0$, $p(|X_n - X| > \varepsilon) \xrightarrow{n} 0$. Si prenem ara $\varepsilon = \frac{1}{i}$, $p\left(|X_n - X| > \frac{1}{i}\right) < \frac{1}{i^2} \xrightarrow{n} 0$. Per tant, podem trobar un índex n_i pel qual si $n \geq n_i$ $p\left(|X_n - X| > \frac{1}{i}\right) < \frac{1}{i^2}$. Prenem ara la subsuccessió $\left\{X_{n_i}\right\}_{i\geq 1}$. Estudiem ara la suma $\sum_{i\geq 1} p\left(A_{n_i}(\varepsilon)\right)$. Aleshores,

$$\sum_{i\geq 1} p\left(A_{n_i}(\varepsilon)\right) = \sum_{i<\frac{1}{\varepsilon}} p\left(A_{n_i}(\varepsilon)\right) + \sum_{i\geq\frac{1}{\varepsilon}} p\left(A_{n_i}(\varepsilon)\right).$$

El primer terme és finit i volem veure si el segon també ho és.

$$\sum_{i\geq\frac{1}{\varepsilon}} p\left(A_{n_i}(\varepsilon)\right) = \sum_{i\geq\frac{1}{\varepsilon}} p\left(|X_{n_i} - X| > \varepsilon\right) \leq \sum_{i\geq\frac{1}{\varepsilon}} p\left(|X_{n_i} - X| > \frac{1}{i}\right) \leq \sum_{i\geq\frac{1}{\varepsilon}} \frac{1}{i^2} < \infty.$$

Per tant, $\forall\varepsilon > 0$ $\sum_{i\geq 1} p\left(A_{n_i}(\varepsilon)\right) < \infty \implies X_{n_i} \xrightarrow{q-s} X$. □

6.2 Convergència quasi-segura. Llei forta dels grans nombres

Començarem veient les propietats més habituals dels límits en el marc de la convergència quasi-segura.

Proposició 6.2.1. Sigui $\{X_n\}_{n\geq 1}$ una successó de variables aleatòries i X i Y dues variables aleatòries tal que $X_n \overset{q.s.}{\to} X$ i $X_n \overset{q.s.}{\to} Y$ aleshores $p\,(X=Y)=1$.

Demostració. Siguin $A = \left\{\omega \in \Omega \mid X_n(\omega) \overset{n}{\to} X(\omega)\right\}$ i $B = \left\{\omega \in \Omega \mid X_n(\omega) \overset{n}{\to} Y(\omega)\right\}$. Sabem per l'enunciat que $p\,(A) = p\,(B) = 1$, i a més que $p\,(A\cup B) = 1$. Per tant, $p\,(A\cap B) = 1$, que ens diu que

$$\begin{aligned} 1 = p\,(A\cap B) = p\left(\left\{\omega \in \Omega \mid X_n(\omega) \overset{n}{\to} X(\omega)\,,\, X_n(\omega) \overset{n}{\to} Y(\omega)\right\}\right) \leq \\ \leq p\left(\{\omega \in \Omega \mid X(\omega) = Y(\omega)\}\right) = p\,(X=Y) \leq 1, \end{aligned}$$

ja que $A\cap B \subseteq \{\omega\in\Omega \mid X(\omega) = Y(\omega)\}$ degut a la unicitat del límit a $\mathbb{R}$; i per tant $p\,(X=Y)=1$. □

Lema 6.2.2. Sigui $\{X_n\}_{n\geq 1}$ una successó de variables aletòries i X i Y dues variables aleatòries tal que $X_n \overset{q.s.}{\to} X$ i $Y_n \overset{q.s.}{\to} Y$. Aleshores,

a) $X_n + Y_n \overset{q.s.}{\to} X + Y$.

b) $X_n Y_n \overset{q.s.}{\to} XY$.

c) $\forall C \in \mathbb{R}$, $CX_n \overset{q.s.}{\to} CX$.

d) $\forall g$ funció contínua, $g\,(X_n) \overset{q.s.}{\to} g\,(X)$.

Demostració.

a) Siguin A i B els successos de la demostració anterior. Sigui

$$C = \left\{\omega\in\Omega \mid X_n(\omega) + Y_n(\omega) \overset{n}{\to} X(\omega) + Y(\omega)\right\}.$$

Com $A\cap B \subseteq C$ tenim que $1 = p\,(A\cap B) \leq p\,(C) \leq 1$ i per tant $p\,(C)=1$, el que ens demostra que $X_n + Y_n \overset{q.s.}{\to} X+Y$.

Els altres apartats es deixen com a exercici al lector. Es demostren similarment al primer. □

I ara aquest lema ens permet enunair el primer teorema important en convergència quasi-segura.

Teorema 6.2.3. *Llei forta dels grans nombres.*
Sigui $\{X_n\}_{n\geq 1}$ una successió de variables aleatòries independents i idènticament distribuides (v.a.i.i.d.), $X_i \sim X$ amb $\mathbb{E}[|X|] < +\infty$. Aleshores, si $\mathbb{E}[X] = \mu$ i $S_n = \sum\limits_{i=1}^{n} X_i$, $\frac{S_n}{n} \overset{q.s.}{\to} \mu$.

Demostració. Per alleugerir la demostració del teorema i perquè formi part dels objectius d'aquest curs afegirem la hipòtesi extra que $\mathbb{V}\mathrm{ar}[X] = \sigma^2 < +\infty$. De totes maneres, aquesta hipòtesi no és necessària per la demostració del teorema.

Primer observem que $\frac{S_n}{n} \overset{q.s.}{\to} \mu \iff \frac{\sum\limits_{i=1}^{n}(X_i-\mu)}{n} \overset{q.s.}{\to} 0$, i per tant podem suposar que $\mathbb{E}[X]=0$. Això ens diu que el resultat que volem demostrar és equivalent a demostrar que $\frac{S_n}{n} \overset{q.s.}{\to} 0$.

Considerem $\forall\varepsilon > 0$, $A_n(\varepsilon) = \left\{\omega \in \Omega \mid \left|\frac{S_n}{n} - 0\right| > \varepsilon\right\} = \{\omega \in \Omega \mid |S_n| > n\varepsilon\}$. És fàcil veure que no és pot fitar $\sum\limits_{n\geq 1} p(A_n(\varepsilon))$, $\forall\epsilon > 0$ per demostrar el resultat fent servir la proposició 6.1.14.

Observem que si prenem $\frac{S_{n^2}}{n^2}$, aleshores $p(A_{n^2}(\varepsilon)) = p\left(|S_{n^2}| > n^2\varepsilon\right) \leq \frac{\sigma^2_{S_{n^2}}}{n^4\varepsilon^2}$, fent servir la desigualtat de Txebixov (2.2.12). Com S_n és una suma de variables aleatòries independents amb esperança finita, $\sigma^2_{S_{n^2}} = n\sigma^2$. Això ens diu que

$$\forall\varepsilon > 0, \sum_{n\geq 1} p(A_{n^2}(\varepsilon)) \leq \frac{\sigma^2}{\varepsilon^2}\sum_{n\geq 1}\frac{1}{n^2} < +\infty,$$

i per tant $\frac{S_{n^2}}{n^2} \xrightarrow{q.s.} 0$.

Per als valors que no són quadrats perfectes definim $D_n = \max\limits_{n^2<k<(n+1)^2} |S_k - S_{n^2}|$. És fàcil veure que aleshores $D_n^2 = \max\limits_{n^2<k<(n+1)^2} (S_k - S_{n^2})^2$, el que ens diu que

$$\begin{aligned}
\mathbb{E}\left[D_n^2\right] &= \mathbb{E}\left[\max_{n^2<k<(n+1)^2} (S_k - S_{n^2})^2\right] \leq \\
&\leq \sum_{k=n^2+1}^{(n+1)^2-1} \mathbb{E}\left[\max_{n^2<k<(n+1)^2} (S_k - S_{n^2})^2\right] = \sum_{k=n^2+1}^{n^2+2n}\left(k-n^2\right)\sigma^2 < \\
&< \sigma^2(2n)(2n) = 4n^2\sigma^2.
\end{aligned}$$

Sigui, $\forall\varepsilon > 0$, $C_n(\varepsilon) = \left\{\omega \in \Omega \mid \left|\frac{D_n}{n^2}\right| > \varepsilon\right\} = \left\{\omega \in \Omega \mid D_n > \varepsilon n^2\right\}$. Aleshores, fent servir la desigualtat de Markov (2.2.11),

$$p(C_n(\varepsilon)) = p\left(D_n > \varepsilon n^2\right) \leq p\left(D_n^2 > \varepsilon^2 n^4\right) \leq \frac{\mathbb{E}\left[D_n^2\right]}{\varepsilon^2 n^4} \leq \frac{4n^2\sigma^2}{\varepsilon^2 n^4} = \frac{4\sigma^2}{\varepsilon^2 n^2}.$$

Això ens diu que

$$\forall\varepsilon > 0, \sum_{n\geq 1} p(C_n(\varepsilon)) \leq \frac{4\sigma^2}{\varepsilon^2}\sum_{n\geq 1}\frac{1}{n^2} < +\infty,$$

i per tant $\frac{D_n}{n^2} \xrightarrow{q.s.} 0$. Això ens permet dir que, sigui $k \in \mathbb{N}$ un nombre no quadrat i sigui $n \in \mathbb{N}$ tal que $n^2 < k < (n+1)^2$, aleshores

$$0 \leq \left|\frac{S_k}{k}\right| \leq \frac{|S_{n^2} + D_n|}{k} \leq \frac{|S_{n^2}|}{n^2} + \frac{D_n}{n^2},$$

que acabem de demostrar que tendeixen les dues quasi-segurament cap a zero. Finalment, fent servir el Lema 6.2.2, tenim que la suma tendeix quasi-segurament cap a zero, el que ens diu que $\left|\frac{S_k}{k}\right| \xrightarrow{q.s.} 0$. □

Corol·lari 6.2.4. *Llei feble dels grans nombres.* Sigui $\{X_n\}_{n\geq 1}$ una successió de variables aleatòries independents i idènticament distribuides (v.a.i.i.d.), $X_i \sim X$ amb $\mathbb{E}[|X|] < +\infty$. Aleshores, si $\mathbb{E}[X] = \mu$ i $S_n = \sum\limits_{i=1}^{n} X_i$, $\frac{S_n}{n} \xrightarrow{p} \mu$.

Demostració. Com $\frac{S_n}{n} \xrightarrow{q.s.} \mu \implies \frac{S_n}{n} \xrightarrow{p} \mu$ la demostració és directa a partir del teorema anterior. □

Observació 6.2.5. A més, a partir del teorema també és cert que $\frac{S_n}{n} \xrightarrow{d} \mu$. El teorema del límit central, que veurem més endavant, és un refinament d'aquest resultat que descriu com fluctúa $\frac{S_n}{n}$ al voltant de μ.

Observació 6.2.6. Si les variables tenen variància finita $\mathbb{V}\mathrm{ar}[X] = \sigma^2 < +\infty$, aleshores,

$$\mathbb{E}\left[\left(\frac{S_n}{n} - \mu\right)^2\right] = \mathbb{E}\left[\frac{1}{n^2}(S_n - n\mu)^2\right] = \frac{1}{n^2}\mathbb{E}\left[\left(\sum_{i=1}^n X_i - n\mu\right)^2\right] = \frac{1}{n^2}\mathbb{V}\mathrm{ar}\left[\sum_{i=1}^n X_i\right] =$$

$$= \frac{1}{n^2}\mathbb{V}\mathrm{ar}[X]n = \frac{\sigma^2}{n} \xrightarrow{n\to\infty} 0.$$

Per tant, $\frac{S_n}{n} \xrightarrow{quad.} \mu \iff \mathbb{V}\mathrm{ar}[X] < +\infty$

6.3 Convergència en distribució. Teorema del límit central

Aquest mode de convergència (també anomenat feble o en llei) té propietats menys rígides que la convergència quasi-segura.

Observació 6.3.1. El límit, si existeix, no té per què ser únic. Si prenem $X \sim \mathrm{Ber}\left(\frac{1}{2}\right)$ i $Y = 1 - X$ i prenem X_n tal que $X_n = X\ \forall n \geq 1$. És clar que $X_n \xrightarrow{d} X$, però també és cert que $X_n \xrightarrow{d} 1 - X = Y$ (mateixa funció de distribució i $X \neq Y$.

Observació 6.3.2. $X_n \xrightarrow{d} X,\ Y_n \xrightarrow{d} Y \ \not\Longrightarrow\ X_n + Y_n \xrightarrow{d} X + Y$. Per exemple, si prenem $X, Y, X_n, Y_n \sim \mathrm{Ber}(\frac{1}{2})$ amb X_n, Y_n independents però $X = Y$, aleshores $X_n \xrightarrow{d} X$ i $Y_n \xrightarrow{d} Y$, i $X_n + Y_n$ satisfà

$$\left.\begin{array}{l} p(X_n + Y_n = 0) = \frac{1}{4} \\ p(X_n + Y_n = 1) = \frac{1}{2} \\ p(X_n + Y_n = 2) = \frac{1}{4} \end{array}\right\} \implies X_n + Y_n \sim \mathrm{Bin}\left(2, \frac{1}{2}\right).$$

Però $X + Y = 2X$ té distribució $\begin{cases} p(X + Y = 0) = \frac{1}{2} \\ p(X + Y = 2) = \frac{1}{2} \end{cases}$ i per tant $X_n + Y_n \stackrel{d}{\not\to} X + Y$.

Si afegim condicions addicionals, molts cops sí que podrem assegurar convergència.

Proposició 6.3.3. Siguin $\{X_n\}_{n\geq 1}, \{Y_n\}_{n\geq 1}$ successions de variables aleatòries amb $X_n \xrightarrow{d} X$, $Y_n \xrightarrow{p} \alpha$, $\alpha \in \mathbb{R}$, aleshores $X_n + Y_n \xrightarrow{d} X + \alpha$.

Demostració. Sigui x un punt de continuïtat de $F_{x+\alpha}$. Observem que

$$F_{x+\alpha}(x) = p(X + \alpha \leq x) = p(X \leq x - \alpha) = F_x(x - \alpha).$$

Per tant, $x - \alpha$ és un punt de continuïtat de F_X. Fem el cas $\alpha > 0$. Sigui δ tal que $0 < \delta < \alpha$. Com $Y_n \xrightarrow{p} \alpha$ existeix $n_0 = n_0(\delta)$ tal que $p(|Y_n - \alpha| > \delta) < \delta \; \forall n \geq n_0$. Per tant,

$$\begin{aligned} p(X_n + Y_n \leq x) &= p(X_n + Y_n \leq x, \, |Y_n - \alpha| \leq \delta) + p(X_n + Y_n \leq x, \, |Y_n - \alpha| > \delta) \\ &\leq p(X_n + Y_n \leq x, \, |Y_n - \alpha| \leq \delta) + p(|Y_n - \alpha| > \delta) \\ &\leq p(X_n \leq x - Y_n, \, |Y_n - \alpha| \leq \delta) + \delta \\ &\leq p(X_n \leq x + \delta - \alpha) + \delta. \end{aligned}$$

On a l'última desigualtat s'ha fet servir que $-\delta \leq Y_n - \alpha$. Per altra banda,

$$\begin{gathered} p(X_n + Y_n > x) \leq p(X_n + Y_n > x, \, |Y_n - \alpha| \leq \delta) + p(|Y_n - \alpha| > \delta) \leq \\ \leq p(X_n + Y_n > x, \, |Y_n - \alpha| \leq \delta) + \delta = p(X_n > x - Y_n, |Y_n - \alpha| \leq \delta) + \delta \leq \\ \overset{\delta \geq Y_n - \alpha}{\leq} p(X_n > x - \alpha - \delta) + \delta. \end{gathered}$$

Per tant, $1 - p(X_n + Y_n > x) \geq 1 - p(X_n > x - \alpha - \delta) - \delta \implies p(X_n + Y_n \leq x) \geq p(X_n \leq x - \alpha - \delta) - \delta$. D'on per a tota $0 < \delta < \alpha$ i tota $n \geq n_0(\delta)$ es té

$$p(X_n \leq x - \alpha - \delta) - \delta \leq p(X_n + Y_n \leq x) \leq p(X_n \leq x - \alpha + \delta) + \delta.$$

Sabent que $F_{X_n}(x - \alpha) \xrightarrow{n} F_X(x - \alpha)$, tenim que fent $\delta \to 0$,

$$\begin{gathered} \lim_{\delta \to 0, \, n \to \infty} p(X_n \leq x - \alpha - \delta) = \lim_{\delta \to 0, \, n \to \infty} p(X_n \leq x - \alpha + \delta) = \\ = p(X \leq x - \alpha) = F_X(x - \alpha) = F_{X+\alpha}(x). \end{gathered}$$

□

Malgrat tot l'anterior, tenim el següent resultat que relaciona la convergència feble amb la convergència quasi-segura.

Teorema 6.3.4. *Teorema de representació de Skorokhod.*
Sigui $\{X_n\}_{n \geq 1}$ successió de variables aleatòries amb $X_n \xrightarrow{d} X$ aleshores existeix un nou espai de probabilitat $(\Omega', \mathcal{A}', p')$ on podem definir les variables aleatòries $\{Y_n\}_{n \geq 1}$ i Y tals que

i) $F_{X_n}(x) = F_{Y_n}(x) \; \forall n \in \mathbb{N}$ i $F_X(x) = F_Y(x)$ (λ-gairebé arreu).

ii) $Y_n \xrightarrow{\text{q-s}} Y$.

Demostració. $\Omega' = [0,1]$, $\mathcal{A}' = \mathcal{B} \cap [0,1]$, $p' = \lambda_{[0,1]}$. Definim ara ($\omega \in \Omega'$):

$$\begin{aligned} Y_n(\omega) &= \inf \{x \in \mathbb{R} \text{ t.q. } \omega \leq F_{X_n}(x)\} \\ Y(\omega) &= \inf \{x \in \mathbb{R} \text{ t.q. } \omega \leq F_X(x)\}. \end{aligned}$$

Observem que $Y_n(\omega) \leq x \iff \omega \leq F_{X_n}(x)$ i $Y(w) \leq x \iff w \leq F_X(x)$. Vegem primer i):

$$\begin{gathered} F_{Y_n}(y) = p'\Big\{(\omega \in (0,1) \text{ t.q. } Y_n(\omega) \leq y\}\Big) = p'(\{\omega \in (0,1) \text{ t.q. } \omega \leq F_{X_n}(y)\} = \\ = p'\Big((0, F_{X_n}(y)]\Big) = F_{X_n}(y). \end{gathered}$$

Anàlogament, $F_Y(y) = F_X(y)\ \forall y \in \mathbb{R}$. Vegem ara ii). Sigui $\omega \in \Omega'$, $\varepsilon > 0$. Prenem ara x tal que $Y(\omega) - \varepsilon < x < Y(w)$ i tal que x és punt de continuïtat de F_X (existeix perquè el conjunt de punts de discontinuïtat de F_x té mesura zero en $\mathbb{R}$). Per ser x punt de continuïtat, $x < Y(\omega) \implies F_X(x) < \omega$ i com $F_{X_n}(x) \to F_X(x)$, tenim que per n prou gran, $F_{X_n}(x) < \omega \implies x \leq Y_n(\omega)$. Per tant, estem veient que si n és prou gran, $Y(w) - \varepsilon < x \leq Y_n(\omega)$. Fixat ara ω, si fem $\varepsilon \to 0$ i fem $n \to \infty$, tenim que

$$Y(\omega) \leq \liminf_n Y_n(\omega) \quad \forall \omega \in \Omega'.$$

Anem a buscar una altra fita. Prenem ara $w < w' < 1$ i sigui x punt de continuïtat de F_X que compleixi

$$Y(\omega') < x < Y(\omega') + \varepsilon \quad (\varepsilon > 0).$$

Com $\omega < \omega'$ i $Y(\omega') < x \implies \omega < \omega' < F_X(x)$. Per tant, per n prou gran, $w < F_{X_n}(x) \implies Y_n(w) \leq x$. Això implica que $Y_n(\omega) \leq x < Y(\omega') + \varepsilon$. Fent $\varepsilon \to 0$, $n \to \infty$, tenim

$$\limsup_n Y_n(\omega) \leq Y(\omega'), \quad \text{on } \omega < \omega'.$$

Per tant,

$$Y(\omega) \leq \liminf_n Y_n(\omega) \leq \limsup_n Y_n(\omega) \leq Y(\omega'), \quad \text{amb } \omega < \omega'. \tag{6.1}$$

Com Y és una funció creixent de $(0,1)$ en $\mathbb{R}$, el seu conjunt de punts de discontinuïtat té mesura zero. Per tant, per a tot punt ω de continuïtat de Y, $\lim\limits_{\omega' \to \omega} = Y(\omega)$ i, per tant, a 6.1 obtenim que si ω és punt de continuïtat de Y, $Y(\omega) = \lim\limits_n Y_n(\omega)$. Per tant, $p'\Big(\{\omega \in \Omega' \text{ t.q. } Y_n(\omega) \to Y(w)\}\Big) = p'\Big(\{\omega \in \Omega' \text{ t.q. } Y_n(\omega) \to Y(w), Y \text{ contínua en } \omega\}\Big)$ $+p'\Big(\{\omega \in \Omega' \text{ t.q. } Y_n(\omega) \to Y(w), Y \text{ no contínua en } \omega\}\Big) = 1 + 0 = 1$. Hi ha convergència quasi-segura. $\square$

Corol·lari 6.3.5. Si $X_n \xrightarrow{d} X$ i g és una funció contínua, aleshores $g(X_n) \xrightarrow{d} g(X)$.

Demostració. Aplicant el teorema anterior existeix $\{Y_n\}_{n\geq 1}$, Y sobre $(\Omega', \mathcal{A}', p')$ tal que $Y_n \xrightarrow{q-s} Y$. Com que g és contínua, $g(Y_n) \xrightarrow{q-s} g(Y) \implies g(Y_n) \xrightarrow{d} g(Y)$. Observem que $g(X_n)$ i $g(Y_n)$ tenen la mateixa distribució perquè tenen la mateixa funció de distribució. Dit d'una altra manera:

$$F_{X_n}(x) = p(X_n \leq x) = p\left(X_n \in (-\infty, x]\right) = p'\left(Y_n \in (-\infty, x]\right) = F_{Y_n}(x).$$

Que implica que $\forall \mathcal{B}$ borelià

$$p(X_n \in \mathcal{B}) = p'(Y_n \in \mathcal{B}).$$

Per tant,

$$p(g(X_n) \leq x) = p\left(g(X_n) \in (-\infty, x]\right) = p\Big(X_n \in g^{-1}(-\infty, x]\Big) = p'\Big(Y_n \in g^{-1}(-\infty, x]\Big) = p'\left(g(Y_n) \in (-\infty, x]\right) = p(g(Y_n) \leq x).$$

Per tant, $F_{g(X_n)}(x) = F_{g(Y_n)}(x)$ λ-g.a. i $F_{g(X)}(x) = F_{g(Y)}(x)$ λ-g.a.. Com $g(Y_n) \xrightarrow{d} g(Y) \implies F_{g(Y_n)}(x) \xrightarrow{n} F_{g(Y)}(x)$ en tot punt de continuïtat. Per tant, llevat d'un conjunt de mesura 0, $F_{g(Y_n)}(x) \xrightarrow{n} F_{g(Y)}(x) \implies F_{g(X_n)}(x) \xrightarrow{n} F_{g(X)}(x)$ llevat d'un conjunt de mesura 0. Com que els punts de discontinuïtat de $F_{g(X)}$ tenen mesura 0 i podem canviar F en un conjunt de mesura 0 queda demostrat el que volíem.

$\square$

Ara veurem com usar les funcions característiques en l'estudi de la convergència en distribució. El mètode que veurem és especialment útil quan les variables aleatòries que estudiem s'obtinguin com a suma de variables aleatòries independents.

Teorema 6.3.6. *Teorema de Lévy.*
Sigui $\{X_n\}_{n\geq 1}$ una successió de variables aleatòries amb les corresponents funcions característiques $\{\phi_n(t)\}_{n\geq 1}$. Aleshores

a) Si $X_n \xrightarrow{d} X$ amb $\phi(t)$ funció característica de X, aleshores $\phi_n(t) \xrightarrow{n} \phi(t)$ puntualment.

b) Si $\phi_n(t) \xrightarrow{n} \phi(t)$ puntualment i $\phi(t)$ és contínua en $t = 0$, aleshores $\exists X$ variable aleatòria amb funció característica $\phi(t)$ t. q. $X_n \xrightarrow{d} X$.

Demostració. La demostració d'aquest teorema es deixa com a exercici pel lector. □

Teorema 6.3.7. *Llei forta dels grans nombres (en distribució).*
Sigui $\{X_n\}_{n\geq 1}$ una successió de variables aleatòries independents idènticament distribuides amb $\mathbb{E}[X_n] = \mu < +\infty$. Sigui $S_n = \sum_{i=1}^{n} X_i$. Aleshores $\frac{S_n}{n} \xrightarrow{d} \mu$.

Demostració. Sigui $\{\phi_n(t)\}_{n\geq 1}$ la successió de funcions característiques de $\{X_n\}_{n\geq 1}$. Com les X_i són independents i idènticament distribuides tenim que

$$\phi_{\frac{S_n}{n}}(t) = \prod_{i=1}^{n} \phi_i\left(\frac{t}{n}\right) = \left(\phi_1\left(\frac{t}{n}\right)\right)^n.$$

Com l'esperança de les X_i és finita tenim que en un entorn de $t = 0$,

$$\phi_1(t) = \phi_1(0) + \phi_1'(0)\,t + o(t) = 1 + i\mu t + o(t).$$

Evaluant en $\frac{t}{n}$ tenim ara que el desenvolupament de Taylor val $\forall t$ fixat si n és prou gran, el que ens deixa que

$$\left(\phi_1\left(\frac{t}{n}\right)\right)^n = (1 + i\mu t + o(t))^n = \left(\left(1 + i\mu\frac{t}{n}(1 + o(1))\right)^{\frac{n}{i\mu t}}\right)^{i\mu t} \xrightarrow{n\to\infty} e^{i\mu t},$$

que és contínua en $t = 0$. Recordem que si X és una variable aleatòria tal que $X = \mu$ aleshores $\phi_X(t) = e^{i\mu t}$, i per tant, pel teorema de Lévy (6.3.6) $\frac{S_n}{n} \xrightarrow{d} \mu$. □

Teorema 6.3.8. *Teorema del límit central.*
Sigui $\{X_n\}_{n\geq 1}$ una successió de variables aleatòries independents idènticament distribuides amb $\mathbb{E}[X_1] = \mu$ i $\mathbb{V}\mathrm{ar}[X_1] = \sigma^2 < \infty$. Aleshores, si definim $S_n = \sum_{i=1}^{n} X_i$, tenim que $\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \xrightarrow{d} N(0,1)$.

Demostració. Siguin $Y_i = \frac{X_i - \mu}{\sigma}$. És fàcil veure que $\mathbb{E}[Y_i] = 0$, $\mathbb{V}\mathrm{ar}[Y_i] = 1$ i que és una famila de variables aleatòries independents idènticament distribuides. Fent el desenvolupament de Taylor en $t = 0$ tenim que per a un entorn d'aquest punt,

$$\phi_{Y_i}(t) = \phi_{Y_i}(0) + \phi_{Y_i}'(0)\,t + \frac{1}{2}\phi_{Y_i}''(0)\,t^2 + o\left(t^2\right) = 1 - \frac{1}{2}t^2 + o(t).$$

Com que $\frac{S_n - n\mu}{\sqrt{n\sigma^2}} = \frac{\sum_{i=1}^n (X_i-\mu)}{\sigma}\frac{1}{\sqrt{n}} = \sum_{i=1}^n \frac{Y_i}{\sqrt{n}}$, tenim que

$$\phi_{\frac{S_n-n\mu}{\sqrt{n\sigma^2}}}(t) = \prod_{i=1}^n \phi_{\frac{Y_i}{\sqrt{n}}}(t) = \prod_{i=1}^n \phi_{Y_i}\left(\frac{t}{\sqrt{n}}\right) = \left(1 - \frac{t^2}{2n} + o\left(\frac{t^2}{n}\right)\right)^n,$$

que ara si que és vàlid $\forall t$ fixat a partir d'un n prou gran. Per tant,

$$\left(1 - \frac{t^2}{2n} + o\left(\frac{t^2}{n}\right)\right)^n = \left(\left(1 - \frac{t^2}{2n}(1+o(1))\right)^{-\frac{2n}{t^2}}\right)^{-\frac{t^2}{2}} \overset{n\to\infty}{\to} e^{-\frac{t^2}{2}},$$

que és contínua en $t=0$. Però a més, és la funció característica d'una normal $N(0,1)$, i per tant, aplicant el teorema de Lévy (6.3.6), $\frac{S_n-n\mu}{\sqrt{n\sigma^2}} \xrightarrow{d} N(0,1)$. □

Observació 6.3.9. Prenem $X_n \sim \text{Bin}(n,p)$ amb p constant. Aleshores podem expressar $X_n = \sum_{i=1}^n Y_i$ on $Y_i \sim \text{Be}(p)$ independents amb $\mathbb{E}[Y_i] = p$ i $\mathbb{V}\text{ar}[Y_i] = p(1-p)$. Fent servir el teorema del límit central podem aproximar la binomial per una normal ja que $\frac{X_n - np}{\sqrt{np(1-p)}} \xrightarrow{d} N(0,1)$. Això no és pot fer si $p := p(n)$.

Observació 6.3.10. El teorema del límit central també pot servir per calcular límits d'expresions complicades. Sigui $\{P_i\}_{i\geq 1}$ una successió de variables aleatòries poisson amb $\lambda = 1$. Sigui $X_n = \sum_{i=1}^n P_i \sim \text{Po}(n)$. Aleshores

$$e^{-n}\sum_{i=0}^n \frac{n^i}{i!} = \sum_{k=0}^n p(X_n = k) = p(X_n \leq n) = p(X_n - n \leq 0) = p\left(\frac{x_n - n}{\sqrt{n}} \leq 0\right) \overset{n\to\infty}{\to} \frac{1}{2},$$

i per tant $\lim_{n\to\infty} e^{-n}\sum_{i=1}^n \frac{n^i}{i!} = \frac{1}{2}$.

6.4 Aplicacions

6.4.1 Conjunts lliures de sumes

Definició 6.4.1. *G grup abelià, $A \subseteq G$ és lliure de sumes si $\nexists a,b,c \in A$ tals que $a+b=c$.*

Problema 6.4.2. Sigui $G = \mathbb{Z}$, $B \subseteq \mathbb{N} \subseteq \mathbb{Z}$, $B = \{b_!, \dots, b_n\}$. Quina és la mida màxima d'un $A \subseteq B$ que sigui lliure de sumes?

Teorema 6.4.3. *(Erdős).*
Donat $B \subseteq \mathbb{N}$, $|B| = n$, existeix $A \subseteq B$ lliure de sumes amb $|A| > \frac{n}{3}$.

Demostració. $B = \{b_!, \dots, b_n\}$, prenem p nombre primer tal que $p > 2\max\{b_i\}$ i $p \equiv 2 \pmod 3$, i fem $p = 2+3k$.

$C = \{k+1, \dots, 2k+1\} \subseteq {\mathbb{Z}}/{p\mathbb{Z}}$ és lliure de sumes: $(k+1)+(k+1) = 2k+2 > 2k+1 \implies (k+1)+(k+1) \notin C$, i $(2k+1)+(2k+1) = 4k+2 \equiv k \pmod p \implies 4k+2 \notin C$

Mirem ara $B \subseteq {\mathbb{Z}}/{p\mathbb{Z}}$ via la projecció natural. Recordem que si $1 \leq x \leq p-1$, $a \equiv b \pmod p \Leftrightarrow ax \equiv bx \pmod p$. Definim $B_x = \{xb_1, \dots, xb_n\}$ (on $1 \leq x \leq p-1$). Ara triem x uniformement a l'atzar en $(1,\ldots,p-1)$. Sigui X la variable que compta quants elements de $B_x \in C$:

$$X = \sum_{i=1}^{n} X_i, \quad X_i = \begin{cases} 1 & \text{si } xb \in C \\ 0 & \text{si } xb \in C \end{cases}$$

$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X_i]$ i per simetria ($\mathbb{E}[X_i] = \mathbb{E}[X_j]$), $\mathbb{E}[X] = n\mathbb{E}[X_1]$. Observem que si $x \neq y$, $xb_1 \not\equiv xb_2 \pmod p$.

$$\mathbb{E}[X_1] = \frac{\text{tries de } x \text{ t. q. } xb_1 \in C}{\text{tries totals}} = \frac{|C|}{p-1} = \frac{k+1}{3k+1} > \frac{1}{3} \implies$$
$$\implies \mathbb{E}[X] > \frac{n}{3} \implies \exists x \in \{1,\ldots,p-1\} \text{ t. q. } B_x \cap C > \frac{n}{3}$$

$B_x \cap C$ és lliure de sumes. Si ara definim $A_x = B_x \cap C$, observem que si $xb_i, xb_j, xb_k \in A_x$, llavors $xb_i + xb_j \not\equiv xb_k \pmod p \iff b_i + b_j \not\equiv b_k \pmod p$, on hem multiplicat per x^{-1}.

Finalment triem $A \subseteq \mathbb{Z}/p\mathbb{Z} \hookrightarrow \mathbb{N}$, com $A = \left(x^{-1}bx, bx_i = B_x \cap C\right)$. $\square$

Observació 6.4.4. Millora (Bourgain): $\exists A$ amb $|A| > \frac{n+2}{3}$

Observació 6.4.5. Una conjectura afirma que $|A| > \frac{n}{3} + f(n)$ on $f(n)$ és una funció $f(n) = o(n)$.

6.4.2 Funcions de representació en els naturals

Definició 6.4.6. Donat $A \subseteq \mathbb{N}, |A| = +\infty$, definim

$$R_A(n) = \left|\left\{(a,b) \in A^2 \text{ t. q. } n = a+b, a \leq b\right\}\right|$$

Problema 6.4.7. Podem construir $A \subseteq \mathbb{N}, |A| = +\infty$ tal que $R_A(n) > 0 \quad \forall n \geq n_0$ de manera que $R_A(n)$ sigui fitada?

No es coneix la resposta a aquesta qüestió. La conjectura d'Erdős-Turán afirma que és impossible.

Teorema 6.4.8. *(Erdős).*
Existeix $A \subseteq \mathbb{N}, |A| = +\infty$ tal que $R_A(n) > 0 \quad \forall n \geq n_0$ de manera que $c_1 \log n \leq R_A(n) \leq c_2 \log n$ per certs valors c_1, c_2.

Demostració. Triem A aleatòriament, triant cada element independentment amb probabilitat $p(x \in A) = k\sqrt{\frac{\log x}{x}}$ amb una k suficientment gran (cada tria és un Bernoulli de paràmetre p_x.

Sigui $R_A(n)$ la variable aleatòria que compta les maneres d'escriure n com a suma de dos elements d'A

$$\begin{aligned}\mathbb{E}[R_A(n)] &= p(1 \in A)p(n-1 \in A) + p(2 \in A)p(n-2 \in A) + \cdots \\ &= \sum_{i=1}^{\lfloor \frac{n}{2} \rfloor} p(i \in A)p(n-i \in A) = \sum_{i=1}^{\lfloor \frac{n}{2} \rfloor} k^2 \left(\frac{\log(n)\log(n-i)}{n(n-i)}\right)^{\frac{1}{2}} \\ &= [...] \sim k^2 I \log n, \text{ on } I = \int_0^1 \frac{1}{\sqrt{s(1-s)}}\,\mathrm{d}s\end{aligned}$$

Definim $A_n = \left\{A \subseteq \mathbb{N} \text{ t. q. } |R_A(n) - \mathbb{E}[R_A(n)]| \geq \frac{1}{2}\mathbb{E}[R_A(n)]\right\}$. Aplicant el teorema de Chernoff (5.1.8) amb $\delta = \frac{1}{2}$,

$$p(A_n) \leq 2\exp\left(-\frac{\delta^2}{3}\mathbb{E}[R_A(n)]\right) \sim 2\exp\left(-\frac{1}{12}k^2 I \log n\right)$$
$$= 2\exp\left(\log\left(n^{\frac{-1}{12}}k^2 I\right)\right) = 2n^{-\frac{1}{12}k^2 I}$$

Aquí prenem k prou gran tal que

$$\frac{1}{12}k^2 I > 1, \quad \sum_{n\in\mathbb{N}} p(A_n) \leq 2\sum_{n\geq 1} n^{-\frac{1}{12}k^2 I} < +\infty$$

Per Borel-Cantelli (1.4.5), $p(\limsup A_n) = 0$. Si $A \in \limsup A_n$, aleshores

$$\exists n_0 \in \mathbb{N} \text{ t. q. } \forall n \geq n_0 \quad |R_A(n) - \mathbb{E}[R_A(n)]| \geq \mathbb{E}\left[\frac{1}{2}R_A(n)\right]$$

Això vol dir que només per un nombre finit dels conjunts és certa aquesta relació, perquè $p(\limsup A_n) = 0$. Per tant, amb probabilitat 1, quan triem A,

$$\exists n_0 \in \mathbb{N} \text{ t. q. } \forall n \geq n_0 \quad |R_A(n) - \mathbb{E}[R_A(n)]| \geq \frac{1}{2}\mathbb{E}\left[R_A(n)\right]$$

que equival a

$$-\frac{1}{2}\mathbb{E}[R_A(n)] \leq R_A(n) - \mathbb{E}[R_A(n)] \leq \mathbb{E}\left[\frac{1}{2}R_A(n)\right]$$

amb probabilitat 1, això és,

$$\frac{1}{2}C\log n \sim \frac{1}{2}\mathbb{E}[R_A(n)] \leq R_A(n) \leq \frac{3}{2}\mathbb{E}[R_A(n)] \sim \frac{3}{2}C\log n$$

$\square$

Índex alfabètic

conjunt
- d'esdeveniments, 1
- de successos, 1

convergència
- en distribució, 58
- en mitjana d'ordre r, 57
- en probabilitat, 57
- quasi-segura, 57

covariància, 17

desviació típica, 14

distribució
- χ quadrat, 41
- beta, 38
- binomial, 25
- binomial negativa, 28
- de Bernoulli, 25
- de Cauchy, 38
- de Fisher-Snedecor, 42
- de Poisson, 27
- de Weibull, 38
- exponencial, 37
- F, 42
- gamma, 38
- geomètrica, 28
- marginal, 16
- normal, 37, 41
- normal multivariant, 42
- t de Student, 42
- uniforme
  - contínua, 37
  - discreta, 26

esdeveniments independents, 5

espai
- de probabilitat, 1
  - producte, 5
- mostral, 1

esperança
- condicionada
  - contínua, 40
  - discreta, 30
  - discreta (2), 30
- mostral, 43

esperança d'una variable aleatòria, 13

família
- exponencial, 54
  - natural, 55

funció
- característica, 50
- de densitat de probabilitat, 35
  - conjunta, 38
  - d'un vector, 38
  - marginal, 39
- de distribució
  - condicionada, 29
  - de probabilitat, 11
- de probabilitat, 1
  - condicionada, 29
- generadora
  - de moments, 47
  - de probabilitat, 23

independència de variables aleatòries, 17

límit
- inferior d'esdeveniments, 6
- superior d'esdeveniments, 6

mesura absolutament contínua, 35

moment
- d'ordre r, 14
- factorial d'ordre r, 14

probabilitat condicionada, 4

subgrup
- lliure de sumes, 68

variable aleatòria, 9

absolutament contínua, 35
  condicionada, 40
contínua, 35
  condicionada, 40
discreta, 21
multidimensional, 16
variància, 14
  mostral, 44
vector de variables aleatòries, 16
  absolutament continu, 38
  continu, 38
  discret, 22