UNIVERSITAT DE BARCELONA
Facultat de Matemàtiques i Informàtica

GRAU DE MATEMÀTIQUES

Treball final de grau

PROBABILITATS AVANÇADES

Autor: Júlia Joana Carbó i Mitjans

Director: Dr. David Marquez Carreras
Realitzat a: Departament de Matemàtiques i Informàtica

Barcelona, 13 de juny de 2022

Abstract

The aim of this work is to extend the concepts studied in the subject of Probabilities of the degree in mathematics. In order to delve into familiar notions and deal with new concepts, we link probability theory to measurement theory. This construction provides us with a mathematical framework and a coherent language that unifies different ideas and eliminates the need to distinguish between fundamentally similar objects. As we have introduced, we will cover familiar concepts such as distributions, expectation, integration of random variables, or independence, and some new ones, such as the extent of measures, the product of probability spaces, or the characteristic function. We will also present some fundamental results such as the convergence theorems, the Radon-Nikodym theorem, the strong law of large numbers or the central limit theorem, among others.

Resum

L'objectiu d'aquest treball és ampliar els conceptes treballats a l'assignatura de Probabilitats del grau de matemàtiques. Per tal d'aprofundir en nocions conegudes i tractar conceptes nous, enllacem la teoria de la probabilitat amb la teoria de la mesura. Aquesta construcció ens proporciona un marc matemàtic i un llenguatge coherent que unifica diferents idees i elimina la necessitat de distingir entre objectes fonamentalment similars. Tal i com hem introduït, tractarem conceptes coneguts com les distribucions, l'esperança, l'integració de variables aleatòries o la independència, i alguns de nous, com poden ser l'extensió de mesures, el producte d'espais de probabilitat o la funció característica. Presentarem també diferents resultats fonamentals com els teoremes de convergència, el teorema de Radon-Nikodym, la llei forta des grans nombres o el teorema central del límit, entre altres.

2020 Mathematics Subject Classification. 60A10, 60F05

Agraïments

Vull agrair especialment al Dr. David Marquez Carreras per la guia i direcció, així com la constància i atenció que m'ha donat al llarg de la realització del treball.

M'agradaria mencionar també la meva amiga i companya de carrera Laia Sosa Soler, per l'ajuda prestada en la matèria i pel recolzament que hem compartit durant aquest últim semestre del Grau.

Agraïment especial també als meus pares, per les facilitats que m'han donat i la paciència que han tingut durant els darrers quatre anys.

"suam habet fortuna rationem"

— *Satiricó*, Gaius Petronius Arbiter.

Índex

0.1 Introducció 1

0.2 Estructura de la memòria 2

1 Mesures **3**

1.1 Estructures de classes de subconjunts 3

1.2 Extensió de mesures i caracterització de la σ-additivitat 5

1.3 Mesures exteriors i Teoremes de Carathéodory 7

1.4 Compleció d'un espai de mesures 10

1.5 Probabilitats en la recta real i funcions de distribució 12

2 Funcions mesurables i esperança matemàtica **18**

2.1 Funcions mesurables i variables aleatòries 18

2.2 Integrabilitat 20

2.2.1 Teoremes de convergència 24

2.2.2 Càlcul d'integrals 26

2.3 Esperança matemàtica 28

3 Producte d'espais mesurables, independència i distribucions condicionades **30**

3.1 Producte d'espais mesurables 30

3.1.1 Producte finit d'espais de mesura 31

3.2 Independència 33

3.3 Distribucions condicionades 37

4 Lleis dels grans nombres **39**

5 Funció característica i Teorema central del límit **44**

5.1 Convergència feble de probabilitats i convergència en llei 44

5.2 Funcions característiques 46

5.3 El teorema de continuïtat 50

5.4 Teorema central del límit i llei normal multidimensional 52

0.3 Conclusió 56

A **57**

A.1 Convergència de successions de variables aleatòries 57

A.1.1 Desigualtat de Txebixev i Desigualtat de Schwarz 57

A.1.2 Lemes de Borel-Cantelli 58

A.1.3 Convergència de variables aleatòries 59

A.2 Esperança Condicionada 59

A.3 Propietats fonamentals de les funcions característiques 62

A.4 Propietats de les lleis normals multidimensionals 64

Bibliografia **65**

0.1 Introducció

L'interès en l'estudi de les probabilitats és directe i natural de la convivència amb els fenòmens aleatoris que envolten la natura i les activitats humanes. L'objecte d'estudi d'aquesta branca de les matemàtiques són essencialment els *esdeveniments*, possibles resultats associats a una experiència aleatòria, i el que interessa és donar una estimació quantitativa de la possibilitat que es produeixin. Això ens porta a buscar la *probabilitat d'un esdeveniment* i el problema de definir i calcular les nocions intrínsecament relacionades amb aquest procés són la base d'aquest domini de les matemàtiques.

La creació d'una teoria matemàtica de la probabilitat va sorgir de dos matemàtics francesos, Blaise Pascal i Pierre de Fermat. L'any 1654, Chevalier de Méré, un noble francès amb interès en les qüestions de jocs i apostes, va cridar l'atenció de Pascal per una aparent contradicció respecte un popular joc de daus. Aquest problema i d'altres plantejats per de Méré, van provocar un intercanvi de cartes entre Pascal i Fermat en les quals van formular, per primera vegada, una base teòrica de la probabilitat [2].

A partir d'aquell moment, i com és habitual en l'evolució de la ciència, van sorgir diferents definicions matemàtiques de probabilitat, però van acabar trobant limitacions a mesura que avaçaven. Una de les dificultats més importants va ser arribar a una construcció prou precisa per utilitzar-la en matemàtiques, però suficientment genèrica per tal de ser aplicable a una àmplia gamma de fenòmens. La recerca d'una definició acceptable va durar gairebé tres segles i va estar marcada per molta controvèrsia, fins que la fonamentació teòrica va arribar a una construcció axiomàtica. *L'axiomàtica de Kolmogorov* [3] va ésser publicada el 1933 i relaciona la teoria de la probabilitat amb la teoria de la mesura, desenvolupada a finals del segle XIX per E. Borel, C. Carathéodory, H. Lebesgue entre altres.

Emmarcar la teoria de la probabilitat en la teoria de la mesura va permetre abordar conjuntament el cas discret i el cas continu així com el tractament més operatiu de diferents conceptes, com poden ser el condicionament, la independència [4] o la convergència en llei. Aquest treball pretén exposar alguns resultats fonamentals de la teoria de la probabilitat, començant des de la base més tècnica de la teoria de la mesura.

[2] l'article [11] detalla aquesta correspondència
[3] [13]
[4] Els conceptes de independència estocàstica i condicionament fan que la teoria de la probabilitat tingui consistència pròpia i diferenciada de la teoria de la mesura

0.2 Estructura de la memòria

El primer capítol d'aquest treball ha fet la funció de preliminars, ja que presenta diferents estructures i conceptes que es necessiten per provar resultats més avançats dels següents temes. En particular, s'introdueixen els conceptes nous de semiàlgebra, classe monòtona, mesura exterior i completació. A partir d'aquest primer tema podem parlar doncs, d'espais mesurables i espai de mesura. També s'estableixen les bases, gràcies als teoremes de Carathéodory i de la unicitat de l'extensió, per tal d'estendre una mesura σ-additiva. L'últim apartat d'aquest primer capítol el dediquem a les probabilitats sobre la recta real, on donada la semiàlgebra de Borel, intorduïm la funció de distribució i veiem diverses propietats i caracteritzacions.

Al segon capítol ja parlem de funcions mesurables i variables aleatòries i veiem una construcció per passos de l'integral respecte d'una mesura. Aquest capítol també presenta diferents resultats pel càlcul efectiu d'integrals, com el teorema de la mesura imatge, i propietats destacades de les integrals de funcions mesurables arbitràries, com els teoremes de convergència de Lebesgue. L'últim apartat d'aquest tema estudia l'esperança matemàtica com a cas particular de la integració de funcions mesurables reals.

Al tercer capítol introduïm el producte d'espais mesurables, ens centrem en el producte finit i demostrem el teorema de la mesura producte. Introduïm també el concepte d'independència de variables aleatòries, central en la teoria de la probabilitat, del qual en veiem diferents aplicacions. Aquest tema inclou també un petit incís de les distribucions condicionades i, a diferència de l'assignatura del grau, busquem una definició que ens permeti condicionar respecte un esdeveniment de probabilitat nul·la.

El quart capítol presenta els diferents lemes necesaris per tal de demostrar un resultat de la llei forta dels grans nombres.

El darrer capítol presenta el concepte de funció característica, que ens permet calcular la funció de distribució. Demostrarem també el teorema de continuïtat i, finalment, el teorema central del límit, així com la seva versió n-dimensional.

Capítol 1

Mesures

En aquest primer tema introduirem diferents estructures de classe de subconjunts, les extensions de mesures, les mesures exteriors i el concepte de completació. A l'últim apartat veurem, com a cas destacat, les probabilitats a la recta real i les funcions de distribució. Molts resultats seran coneguts pel que a moltes demostracions només donarem una indicació o deixarem indicada una referència.

1.1 Estructures de classes de subconjunts

En aquesta secció introduïrem les classes de subconjunts que ens seran d'utilitat en els següents capítols, així com diferents resultats rellevants que ens les relacionen.
Sigui $\mathcal{C} \subseteq \mathcal{P}$ [1] una col·lecció de subconjunts de Ω qualsevol. Considerem quatre possibles estructures per a $\mathcal{C}$: l'àlgebra, la σ-àlgebra, la semiàlgebra i la classe monòtona.

Definició 1.1.1. *Una classe de subconjunts $\mathcal{A} \subseteq \mathcal{P}$ és una àlgebra si compleix les condicions següents:*

(i) $\Omega \in \mathcal{A}$.

(ii) *$\mathcal{A}$ és estable per pas al complementari, és a dir que si $A \in \mathcal{A}$, és té també que $A^c \in \mathcal{A}$.*

(iii) *$\mathcal{A}$ és estable per reunions finites, és a dir que donats $\{A_1, A_2, \ldots, A_k\}$ de $\mathcal{A}$, és té també que $\cup_{i=1}^{k} A_i \in \mathcal{A}$.*

Definició 1.1.2. *Una classe de subconjunts $\mathcal{F} \subseteq \mathcal{P}$ és una σ-àlgebra si compleix les condicions (i) i (ii) de la defincició anterior i és estable per unions numerables, és a dir, donats $\{A_i, i \geq 1\}$ de $\mathcal{F}$, és té que $\cup_{i=1}^{\infty} A_i \in \mathcal{F}$.*

Notem que es compleix que tota σ-àlgebra és àlgebra.

Proposició 1.1.3. *Tota àlgebra $\mathcal{A}$ és tancada per interseccions finites i tota σ-àlgebra $\mathcal{F}$ és tancada per interseccions finites o numerables.*

Demostració. Fem la prova per el cas numerable, el cas finit és anàleg. Donada una successió d'esdeveniments $\{A_i, i \geq 1\}$ de $\mathcal{F}$ tenim $\cap_i^{\infty} A_i = (\cup_i^{\infty} A_i^c)^c$. El conjunt de la dreta pertany a $\mathcal{F}$ ja que una σ-àlgebra és estable per pas al complementari i per reunions numerables. $\square$

Ens interessa considerar una altra estructura: l'estructura de semiàlgebra.

[1] on $\mathcal{P} = \mathcal{P}(\Omega)$ denota el conjunt de les parts

Definició 1.1.4. *Una classe de subconjunts de Ω, $\mathcal{S}$, és una semiàlgebra si compleix*

(i) $\Omega \in \mathcal{S}$.

(ii) Si A, $B \in \mathcal{S}$ *es té que* $A \cap B \in \mathcal{S}$.

(iii) Si $A \in \mathcal{S}, A^c$ *és una reunió finita de subconjunts de* $\mathcal{S}$ *disjunts dos a dos.*

Tenim que tota àlgebra és semiàlgebra, a l'inrevés no és cert. L'última estructura que ens interessa és la de classe monòtona, que ens permetrà tractar l'existència i unicitat de la σ-àlgebra generada per l'àlgebra (Teorema 1.3.1). Aquests resultats els necessitarem per determinar més endevant la llei d'una variable aleatòria a partir de la seva funció de distribució.

Definició 1.1.5. *Una familia de subconjunts $\mathcal{M}$ de Ω és una classe monòtona si és no buida i tota successió monòtona, creixent o decreixent, d'esdeveniments de $\mathcal{M}$, $\{A_i, i \geq 1\}$, compleix* $\lim\limits_n A_n \in \mathcal{M}$ [2].

Notem que tenim que tota σ-àlgebra és una classe monòtona, ja que és tancada per reunions i interseccions numerables. El resultat següent ens permet relacionar els conceptes de σ-àlgebra, àlgebra i classe monòtona:

Proposició 1.1.6. *Una classe de subconjunts $\mathcal{F}$ és una σ-àlgebra si i només si és una àlgebra i una classe monòtona.*

Demostració. La implicació directa és trivial per definició. Per a la implicació inversa considerem una familia numerable de conjunts $\{A_i, i \geq 1\}$ de $\mathcal{F}$. Podem escriure per tot $n \in \mathbb{N}$,

$$\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} \bigcup_{k=1}^{n} A_k = \lim_{n\to\infty} \bigcup_{k=1}^{n} A_k \in \mathcal{F},$$

doncs $\{\cup_{k=1}^n A_k\}_{n\in\mathbb{N}}$ és una successió creixent de conjunts de $\mathcal{F}$, i per tant el seu límit també és de $\mathcal{F}$. □

Definició 1.1.7. *Sigui $\mathcal{C}$ una família qualsevol de subconjunts de Ω no buida. Es defineix $\sigma(\mathcal{C})$, la σ-àlgebra generada per $\mathcal{C}$, com la intersecció de totes les σ-àlgebres que contenen $\mathcal{C}$. Tenim que $\sigma(\mathcal{C})$ és la mínima σ-àlgebra que conté $\mathcal{C}$ en el sentit que està continguda es tota σ-àlgebra que contingui $\mathcal{C}$.*
Anàlogament podem definir $a(\mathcal{C})$, $s(\mathcal{C})$, $m(\mathcal{C})$, respectivament com l'algebra, la semiàlgebra i la classe monòtona generades per la classe de subconjunts de $\mathcal{C}$.

Proposició 1.1.8. *L'àlgebra generada per una semiàlgebra $\mathcal{S}$ coincideix amb la familia de les unions finites de conjunts de $\mathcal{S}$ disjunts dos a dos.*

Demostració. Podem consultar la prova d'aquest resultat a [10, Proposició 4.13]. □

Lema 1.1.9. *Sigui $\mathcal{A}$ una àlgebra i $m(\mathcal{A})$ la menor classe monòtona que conté $\mathcal{A}$. Aleshores, $m(\mathcal{A})$ també és una àlgebra.*

Demostració. Una prova detallada d'aquest lema la podem trobar [1, Lema 1.2.4]. Indiquem un esquema del que hauriem de veure:

1. Cal veure que $m(\mathcal{A})$ conté Ω. És directe del fet que $\Omega \in \mathcal{A}$.

[2] És a dir una classe monòtona és una classe de subconjunts tancada per pas al límit de successions monòtones.

2. Cal veure que $m(\mathcal{A})$ és una classe tancada per pas al complementari. Per comprovar-ho definim la classe dels conjunts $\mathcal{D} \subseteq m(\mathcal{A})$,

$$\mathcal{D} := \{A \in m(\mathcal{A}) : A^c \in m(\mathcal{A})\},$$

és a dir la classe dels conjunts que compleixen la propietat, i demostrem que és classe monòtona. Ho provaríem veient que $\mathcal{D}$ és no buida i tancada per successions monòtones.

3. Cal veure que $m(\mathcal{A})$ és tancada per reunions finites és a dir que donats $A, B \in m(\mathcal{A})$ es compliex que la seva reunió $A \cup B \in m(\mathcal{A})$. Per provar-ho fixarem un $A \in m(\mathcal{A})$ qualsevol i definirem $\mathcal{D}_A$ la classe dels conjunts següents,

$$\mathcal{D}_A := \{B \in m(\mathcal{A}) : A \cup B \in m(\mathcal{A})\} \subseteq m(\mathcal{A}),$$

anàlogament veuriem que $\mathcal{D}_A$ és classe monòtona tindrem la igualtat $\mathcal{D}_A = m(\mathcal{A})$ per a tot $A \in m(\mathcal{A})$ i hauriem acabat la demostració.

□

Teorema 1.1.10 (Teorema de les classes monòtones de Halmo's). *Sigui $\mathcal{A}$ una àlgebra. Sigui $\mathcal{M}$ una classe monòtona tal que $\mathcal{A} \subseteq \mathcal{M}$. Aleshores $\sigma(\mathcal{A}) \subseteq \mathcal{M}$.*

Demostració. Si veiem que $m(\mathcal{A}) = \sigma(\mathcal{A})$ haurem demostrat el teorema ja que $m(\mathcal{A}) \subseteq \mathcal{M}$. En efecte d'una banda $m(\mathcal{A}) \subseteq \sigma(\mathcal{A})$ ja que tota σ-àlgebra és classe monòtona. Però d'altra banda, pel lema anterior tenim que $m(\mathcal{A})$ és àlgebra, i també és òbviament classe monòtona, aleshores per la Proposició 1.1.6 és una σ-àlgebra i necessàriament tenim que $\sigma(\mathcal{A}) \subseteq m(\mathcal{A})$. □

Aquest teorema ens diu també que la classe monòtona i la σ-àlgebra generades per una àlgebra coincideixen.

1.2 Extensió de mesures i caracterització de la σ-additivitat

En aquest apartat exposarem alguns resultats sobre la construcció i les propietats que necesitarem en l'estudi de les mesures.

Definició 1.2.1. *Sigui Ω un conjunt i $\mathcal{C}$ una col·lecció de parts de Ω tal que $\emptyset \in \mathcal{C}$. Considerem una funció $\mu : \mathcal{C} \longrightarrow [0, \infty]$.*

(i) Direm que μ és additiva si $\mu(A \cup B) = \mu(A) + \mu(B)$, sempre que A,B, $A \cup B \in \mathcal{C}$ i $A \cap B = \emptyset$.

(ii) Direm que μ és σ-additiva si $\mu(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$, sempre que $A_n \in \mathcal{C}$ per a tot $n \geq 1$, $\cup_{n=1}^{\infty} A_n \in \mathcal{C}$ i $A_n \cap A_m = \emptyset$, si $n \neq m$.

Una funció additiva que no sigui idènticament infinita compleix $\mu(\emptyset) = 0$.

Definició 1.2.2. *Sigui Ω un conjunt i $\mathcal{F}$ una σ-àlgebra de $\mathcal{P}(\Omega)$. El parell $(\Omega, \mathcal{F})$ és un espai mesurable i els elements de $\mathcal{F}$ s'anomenen conjunts mesurables de $(\Omega, \mathcal{F})$. La terna $(\Omega, \mathcal{F}, \mu)$, on μ és una funció σ-additiva sobre $\mathcal{F}$, s'anomena espai de mesura.*

Entenem com a *mesura* una funció $\mu : \mathcal{C} \longrightarrow [0, \infty]$ σ-additiva que compleix $\mu(\emptyset) = 0$. Passem a veure diferents propietats de les funcions σ-additives.

Proposició 1.2.3. *Sigui $\mathcal{A}$ una àlgebra associada a un conjunt Ω. Sigui μ una funció additiva que compleix $\mu(\emptyset) = 0$ definida sobre $\mathcal{A}$. Aleshores es compleix:*

(i) *Per a tot $A, B \in \mathcal{A}$ es té $\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$.*

(ii) *Per a tot $A, B \in \mathcal{A}$ tals que $A \subseteq B$, es té $\mu(A) \leq \mu(B)$.*

(iii) *Si $A_1, ..., A_n \in \mathcal{A}$ aleshores es té que μ és una funció sub-additiva, és a dir que compleix $\mu(\cup_{i=1}^{n} A_i) \leq \sum_{i=1}^{n} \mu(A_i)$.*

(iv) *Si μ és σ-additiva, és compleix la subadditivitat numerable $\mu(\cup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} \mu(A_i)$.*

Demostració. Una prova de *(i)* es pot trobar a [1, Proposició 1.1.17] i una prova de *(ii)* a [4, Volum I, Proposició 1.3.9]. Veiem els resultats *(iii)* i *(iv)*:

(iii) Podem convertir una reunió d'elements finita $\{A_1, ..., A_n\}$ qualsevol en una reunió disjunta definint $B_1 = A_1$ i $B_n = A_n \cap (\cap_{i=1}^{n-1} A_i^c)$ i es compleix que,

$$\bigcup_{i=1}^{n} A_i = \bigcup_{i=1}^{n} B_i,$$

aleshores,

$$\mu\Big(\bigcup_{i=1}^{n} A_i\Big) = \mu\Big(\bigcup_{i=1}^{n} B_i\Big) = \sum_{i=1}^{n} \mu(B_i) \leq \sum_{i=1}^{n} \mu(A_i).$$

(iv) El raonament anterior es pot estendre al cas numerable si μ és σ-additiva.

□

Definició 1.2.4. *Sigui Ω un conjunt, $\mathcal{A}$ un subconjunt de parts de Ω i μ una funció σ-additiva sobre $\mathcal{A}$. Direm que μ es σ-finita si existeix una descomposició $\Omega = \cup_{n=1}^{\infty} A_n$ del conjunt Ω en conjunts $A_n \in \mathcal{A}$ tals que $\mu(A_n) < \infty$ per a tot n. La successió $\{A_i, i \geq 1\}$ es pot agafar creixent* [3] *o formada per conjunts disjunts dos a dos* [4].

Definició 1.2.5. *Sigui $(\Omega, \mathcal{A}, \mu)$ un espai de mesura. Aleshores μ és una mesura de probabilitat (o una probabilitat) si compleix que $\mu(\Omega) = 1$; en tal cas tenim que $(\Omega, \mathcal{A}, \mu)$ s'anomena espai de probabilitat, i dels conjunts mesurables se'n diu també esdeveniments.*

Els següents resultats ens serviran per comprovar si una funció μ es σ-additiva. Veurem en la Proposició 1.2.6 que si μ és una funció additiva en una àlgebra, aleshores la σ-additivitat equival a la continuïtat de μ sobre successions creixents (i), o sobre successions decreixents (ii) i (iii), en les quals caldrà afegir la condició de $\mu(\Omega) < \infty$.

Proposició 1.2.6. *Sigui $\mu : \mathcal{A} \longrightarrow [0, \infty]$ una funció additiva, $\mu(\emptyset) = 0$ i $\mathcal{A}$ és una àlgebra de parts de Ω. Aleshores,*

(i) *La funció μ és σ-additiva si i només si per a tota successió creixent $\{A_n, n \geq 1\} \subseteq \mathcal{A}$ tal que $A = \cup_{n=1}^{\infty} A_n \in \mathcal{A}$ es compleix $\mu(A_n) \uparrow \mu(A)$.*

(ii) *Si la funció μ és σ-additiva, aleshores per a tota successió decreixent $\{A_n, n \geq 1\} \subseteq \mathcal{A}$ tal que $A = \cap_{n=1}^{\infty} A_n \in \mathcal{A}$ i $\mu(A_1) < \infty$ es compleix $\mu(A_n) \downarrow \mu(A)$.*

(iii) *Si per tota successió decreixent $\{A_n, n \geq 1\} \subseteq \mathcal{A}$ tal que $\cap_{n=1}^{\infty} A_n = \emptyset$ es compleix $\mu(A_n) \downarrow 0$, aleshores μ és σ-additiva. .*

Demostració. Una prova detallada d'aquesta proposició la podem trobar a [10, Proposició 4.2]. □

[3] només cal prendre $A'_n = A_1 \cup A_2 \cup ... \cup A_n$

[4] només cal prendre $A''_n = A'_n - (A_1 \cup A_2 \cup .. \cup A_{n-1}), n > 1$

1.3 Mesures exteriors i Teoremes de Carathéodory

Sigui $\mathcal{A}$ una àlgebra. L'objectiu d'aquesta secció és veure que podem estendre μ de manera única a $\sigma(\mathcal{A})$ i per tant definir una mesura sobre $(\Omega, \sigma(\mathcal{A}))$. Primerament veurem el següent teorema que mostra la unicitat de l'extensió d'una mesura d'una àlgebra a una σ-àlgebra.

Teorema 1.3.1 (Teorema de la unicitat de l'extensió). *Sigui $\mathcal{A}$ una àlgebra de parts d'un conjunt Ω. Si μ_1, μ_2, són dues mesures sobre la σ-àlgebra $\sigma(\mathcal{A})$ tals que coincideixen sobre $\mathcal{A}$ i són σ-finites sobre $\mathcal{A}$, aleshores $\mu_1 = \mu_2$.*

Demostració. Si μ_1 i μ_2 són finites, definim la classe de conjunts de la σ-àlgebra on coincideixen: $\mathcal{C} = \{A \in \sigma(\mathcal{A}) : \mu_1(A) = \mu_2(A)\}$. És fàcil veure que $\mathcal{C}$ és una classe monòtona i inclou per definició $\mathcal{A}$. Per tant, pel Teorema de les classes monòtones (Teorema 1.1.10) inclou també $\sigma(A)$. En el cas σ-finit, podem considerar $\Omega = \cup_n \Omega_n$ si $\mu_i(\Omega_n) < \infty$ per $i = 1, 2$ i considerar les mesures finites $\mu_{1,n}$ i $\mu_{2,n}$, definides sobre cada Ω_n com

$$\mu_{i,n}(B) := \mu_i(B \cap \Omega_n) \quad \text{per} \quad i = 1, 2.$$

Aleshores tenim $\mu_{1,n} = \mu_{2,n}$ per a tot n i per tant coincideixen en general. Notem que la suma numerable de mesures és una mesura. □

Per tal de demostrar els teoremes d'extensió de mesures cal introduïr el concepte de mesura exterior.

Definició 1.3.2. *Sigui Ω un conjunt. Una aplicació μ^*: $\mathcal{P}(\Omega) \longrightarrow [0, +\infty]$ s'anomena una mesura exterior si compleix les següents propietats:*

(i) $\mu^*(\emptyset) = 0$.

(ii) *μ^* és creixent, és a dir, si $A \subseteq B$ aleshores $\mu^*(A) \leq \mu^*(B)$.*

(iii) *μ^* és subadditiva, és a dir que per a qualsevol successió numerable de subconjunts de Ω, $\{A_n, n \geq 1\}$, es compleix*
$$\mu^*\Big(\bigcup_{n=1}^{\infty} A_n\Big) \leq \sum_{n=1}^{\infty} \mu^*(A_n).$$

Notem que una mesura σ-additiva definida sobre $\mathcal{P}$ és una mesura exterior però no a l'inrevés, és a dir una mesura exterior no és necessàriament σ-additiva.

Definició 1.3.3. *Donada una mesura exterior μ^* direm que un conjunt $A \subseteq \Omega$ és μ^*-mesurable si per a tot conjunt $B \subseteq \Omega$ es compleix $\mu^*(B) = \mu^*(A \cap B) + \mu^*(A^c \cap B)$.*

Passem a veure i demostrar els Teoremes de Carathéodory.

Teorema 1.3.4 (Primer teorema de Carathéodory). *Sigui $\mathcal{A}_{\mu^*}$ la familia dels conjunts μ^*-mesurables. Aleshores $\mathcal{A}_{\mu^*}$ és una σ-àlgebra i μ^* és σ-additiva sobre $\mathcal{A}_{\mu^*}$.*

Demostració. Veurem primer que $\mathcal{A}_{\mu^*}$ és una àlgebra. El conjunt $\emptyset$ és de $\mathcal{A}_{\mu^*}$, ja que per a tot $B \subseteq \Omega$ es compleix $\mu^*(B) = \mu^*(\emptyset \cap B) + \mu^*(\Omega \cap B)$.
Clarament $\mathcal{A}_{\mu^*}$ és tancada per pas al complementari perquè la condició de μ^*-mesurabilitat és simètrica respecte el pas al complementari. Ens queda provar doncs, l'estabililtat per reunions finites: Siguin $A_1, A_2 \in \mathcal{A}_{\mu^*}$ i fixem $B \subseteq \Omega$, tenim

$$\begin{aligned}\mu^*(B) &= \mu^*(A_1 \cap B) + \mu^*(A_1^c \cap B)\\ &= \mu^*(A_1 \cap A_2 \cap B) + \mu^*(A_1^c \cap A_2^c \cap B) + \mu^*(A_1^c \cap A_2 \cap B) + \mu^*(A_1 \cap A_2^c \cap B).\end{aligned}$$

Si apliquem aquesta fórmula general al conjunt $B \cap (A_1 \cup A_2)$ obtenim:

$$\mu^*(B \cap (A_1 \cap A_2)) = \mu^*(B \cap A_1 \cap A_2) + \mu^*(B \cap A_1 \cap A_2^c) + \mu^*(B \cap A_1^c \cap A_2),$$

ja que $(A_1 \cap A_2^c) \subseteq (A_1 \cup A_2)$ i $(A_1^c \cap A_2) \subseteq (A_1 \cup A_2)$.

En conclusió tenim que per a tot $B \subseteq \Omega$,

$$\mu^*(B) = \mu^*(B \cap (A_1 \cup A_2)) + \mu^*(B \cap (A_1 \cup A_2)^c),$$

la qual cosa implica que $A_1 \cup A_2 \in \mathcal{A}_\mu *$.

Demostrem per últim que $\mathcal{A}_\mu *$ és una σ-àlgebra i que μ^* és σ-additiva sobre $\mathcal{A}_\mu *$. Considerem una familia de conjunts de $\mathcal{A}_\mu *$, $\{A_n,\ \text{n} \geq 1\}$ i sense pèrdua de generalitat podem suposar que són disjunts dos a dos. Posem $B_n = \cup_{k=1}^n A_k$ per $n \geq 1$ amb $B_0 = \emptyset$ i $A = \cup_{n=1}^\infty A_n$. Sabem que $B_n \in \mathcal{A}_\mu *$ per a tot n. Sigui $C \subseteq \Omega$ i $p \geq 1$. Es compleix:

$$\begin{aligned}\mu^*(C \cap B_p) &= \mu^*(C \cap B_p \cap B_{p-1}) + \mu^*(C \cap B_p \cap B_{p-1}^c) \\ &= \mu^*(C \cap B_{p-1}) + \mu^*(C \cap A_p).\end{aligned}$$

Sumant per p=1 fins a n obtenim:

$$\mu^*(C \cap B_n) = \sum_{p=1}^n \mu^*(C \cap A_p).$$

per tant, com que $B_n \subseteq A$, ens queda:

$$\mu^*(C) = \mu^*(C \cap B_n) + \mu^*(C \cap B_n^c) = \sum_{p=1}^n \mu^*(C \cap A_p) + \mu^*(C \cap B_n^c)$$

$$\geq \sum_{p=1}^n \mu^*(C \cap A_p) + \mu^*(C \cap A^c).$$

Fent la n tendir a infinit $n \to \infty$ i degut a la subadditivitat de μ^*, tindrem

$$\mu^*(C) \geq \sum_{p=1}^\infty \mu^*(C \cap A_p) + \mu^*(C \cap A^c) \geq \mu^*(C \cap A) + \mu^*(C \cap A^c),$$

la qual cosa implica que $A \in \mathcal{A}_\mu *$ i, per tant, $\mathcal{A}_\mu *$ és una σ-àlgebra.
Finalment, posant $C = A$ en l'expressió anterior obtenim:

$$\mu^*(A) \geq \sum_{p=1}^\infty \mu^*(A_p).$$

□

Teorema 1.3.5 (Segon teorema de Carathéodory). *Sigui $\mathcal{A}$ una àlgebra de parts d'un conjunt de Ω i μ una funció σ-additiva sobre $\mathcal{A}$. Per a tot $\mathcal{B} \subseteq \Omega$ definim,*

$$\mu^*(B) = \inf \sum_{n=1}^\infty \mu(A_n),$$

on l'ínfim es pren sobre tots els recobriments numerables $\{A_n, n \geq 1\}$ del conjunt B per conjunts $A_n \in \mathcal{A}$. Aleshores es compleix que μ^* és una mesura exterior, $\mathcal{A} \subseteq \mathcal{A}_\mu *$ i $\mu^*|_{\mathcal{A}} = \mu$.

Demostració. Veurem primer que $\mu^*|_{\mathcal{A}} = \mu$. La desigualtat $\mu(A) \geq \mu^*(A)$ és immediata, és a dir que només hem de provar la desigualtat contrària. Fixem $A \in \mathcal{A}$ i un recobriment $\{A_n, n \geq 1\} \subseteq \mathcal{A}$ de A. Tindrem que $\{A_n \cap A, n \geq 1\}$ és una família de conjunts tal que $A = \cup_{n=1}^{\infty}(A_n \cap A)$. Com que μ és subadditiva podem escriure,

$$\mu(A) = \mu\Big(\bigcup_{n=1}^{\infty}(A_n \cap A)\Big) \leq \sum_{n=1}^{\infty}\mu(A_n \cap A) \leq \sum_{n=1}^{\infty}\mu(A_n).$$

la qual cosa implica que $\mu(A) \leq \mu^*(A)$ i ja tenim provada la igualtat.
Veurem ara que μ^* és una mesura exterior, notem que compleix:

(i) $\mu^*(\emptyset) = 0$.

(ii) Si $A \subseteq B$, tot recobriment numerable B és també un recobriment d'A, pel que $\mu^*(A) \leq \mu^*(B)$ i μ^* és creixent.

(iii) Cal veure que μ^* és subadditiva. Si $B = \cup_{n=1}^{\infty}B_n$, tenim que per a tot $\varepsilon > 0$ i per a tot $n \geq 1$ existirà un recobriment $\{B_n^k, k \geq 1\} \subseteq \mathcal{A}$ de B_n tal que

$$\sum_{k=1}^{\infty}\mu(B_n^k) \leq \mu^*(B_n) + \frac{\varepsilon}{2^n}.$$

La col·lecció de conjunts $\{B_n^k, k, n \geq 1\}$ és un recobriment numerable de B i per a tot $\varepsilon > 0$,

$$\mu^*(B) \leq \sum_{\substack{n=1\\k=1}}^{\infty}\mu(B_n^k) \leq \sum_{n=1}^{\infty}\mu^*(B_n) + \varepsilon.$$

en conseqüencia, com que ε és arbitrari,

$$\mu^*(B) \leq \sum_{n=1}^{\infty}\mu^*(B_n).$$

Tenim doncs que μ^* és subadditiva i μ^* és mesura exterior. Per últim ens queda veure que $\mathcal{A} \subseteq \mathcal{A}_{\mu^*}$. Fixem $A \in \mathcal{A}$ i $B \subseteq \Omega$. Per a tot $\varepsilon > 0$ existeix un recobriment de B, $\{A_n, n \geq 1\} \subseteq \mathcal{A}$, tal que

$$\sum_{n=1}^{\infty}\mu(A_n) \leq \mu^*(B) + \varepsilon.$$

Ara bé,

$$\sum_{n=1}^{\infty}\mu(A_n) = \sum_{n=1}^{\infty}[\mu(A \cap A_n) + \mu(A^c \cap A_n)] \geq \mu^*(A \cap B) + \mu^*(A^c \cap B).$$

Per tant, $\mu^*(B) \geq \mu^*(A \cap B) + \mu^*(A^c \cap B)$, la qual cosa implica que $A \in \mathcal{A}_{\mu^*}$. ☐

Observem que els dos teoremes combinats ens diuen que si tenim definida una funció σ-additiva sobre una àlgebra associada a un conjunt Ω, la podem estendre a una σ-àlgebra $\mathcal{A}_{\mu^*}$ que òbviament inclou $\sigma(\mathcal{A})$. Pel teorema d'unicitat tenim, a més, que si la funció es σ-finita, l'extensió a $\sigma(\mathcal{A})$ és única.

1.4 Compleció d'un espai de mesures

Acabem de veure que donada μ una funció σ-additiva en una àlgebra $\mathcal{A}$ de parts de Ω, μ es pot estendre a la σ-àlgebra $\sigma(\mathcal{A})$ i aquesta extensió és única si μ és σ-finita sobre $\mathcal{A}$. Això ho hem vist construint una mesura exterior μ^* mitjançant recobriments numerables amb conjunts de l'àlgebra $\mathcal{A}$ i prenent després la σ-àlgebra $\mathcal{A}_{\mu^*}$, tal que $\sigma(\mathcal{A}) \subseteq \mathcal{A}_{\mu^*}$, dels conjunts μ^*-mesurables. Això ho hem pogut fer perquè aquest mètode ens permet estendre μ a una σ-àlgebra $\mathcal{A}_{\mu^*}$ més gran que $\sigma(\mathcal{A})$. Veurem que no hi ha molta diferència entre elles però, per fer aquestes consideracions, hem d'introduir els conceptes de conjunt negligible i de compleció d'un espai de mesura.

Definició 1.4.1. *Considerem un espai de mesura $(\Omega, \mathcal{A}, \mu)$. Diem que un conjunt $N \subseteq \Omega$* [5] *és neglibible si existeix un conjunt $A \in \mathcal{A}$ tal que $N \subseteq A$ i $\mu(A) = 0$.*

Definició 1.4.2. *Direm que l'espai de mesura és complet si la σ-àlgebra $\mathcal{A}$ conté tots els conjunts negligibles.*

Definició 1.4.3. *Sigui $(\Omega, \mathcal{A}, \mu)$ un espai de mesura. Definim el conjunt $\bar{\mathcal{A}} := \{A \cup N$ tals que $A \in \mathcal{A}$ i N és negligible$\}$ i la següent aplicació $\bar{\mu} : \bar{\mathcal{A}} :\longrightarrow [0, +\infty]$ de manera que $\bar{\mu}(A \cup N) = \mu(A)$. Tenim que $(\Omega, \bar{\mathcal{A}}, \bar{\mu})$ és un espai complet i s'anomena la compleció de $(\Omega, \mathcal{A}, \mu)$.*

Utilitzant la definició anterior podem comprovar que $\bar{\mathcal{A}}$ és una σ-àlgebra que conté $\mathcal{A}$. També es compleix que $\bar{\mu}$ és σ-additiva i $\bar{\mu}|_{\mathcal{A}} = \mu$.

Teorema 1.4.4. *Sigui μ una funció σ-additiva i σ-finita en una àlgebra $\mathcal{A}$ de parts d'un conjunt Ω. Aleshores l'espai de mesura $(\Omega, \mathcal{A}_{\mu^*}, \mu^*)$* [6] *és la compleció de $(\Omega, \sigma(\mathcal{A}), \mu^*)$.*

Demostració. En primer lloc veurem que $(\Omega, \mathcal{A}_{\mu^*}, \mu^*)$ és complet. En efecte, si N és negligible, existeix $A \in \mathcal{A}_{\mu^*}$ tal que $N \subseteq A$ i $\mu^*(A) = 0$. Això implica que per a tot $B \subseteq \Omega$,

$$\mu^*(B) \geq \mu^*(B \cap N) + \mu^*(B \cap N^c),$$

ja que $\mu^*(B \cap N) = 0$. Per tant, $N \in \mathcal{A}_{\mu^*}$. És clar que $\overline{\sigma(\mathcal{A})} \subseteq \mathcal{A}_{\mu^*}$, ja que $(\Omega, \mathcal{A}_{\mu^*}, \mu^*)$ és complet.

Fixem un $A \in \mathcal{A}_{\mu^*}$, volem veure que $A \in \overline{\sigma(\mathcal{A})}$: per a cada $n \geq 1$ existeix un recobriment $\{A_n^k, k \geq 1\}$ de A, format per conjunts de $\mathcal{A}$ tal que

$$\frac{1}{n} + \mu^*(A) \geq \sum_{k=1}^{\infty} \mu(A_n^k) \geq \mu\Big(\bigcup_{k=1}^{\infty} A_n^k\Big) = \mu^*(G_n) \geq \mu^*(A),$$

on $G_n = \cup_{k=1}^{\infty} A_n^k \in \sigma(\mathcal{A})$. Definim $G = \cap_{n=1}^{\infty} G_n \in \sigma(\mathcal{A})$. Es compleix $A \subseteq G$ i per a tot $n \geq 1$ tenim

$$\frac{1}{n} + \mu^*(A) \geq \mu^*(G_n) \geq \mu^*(G) \geq \mu^*(A).$$

En conseqüència $\mu^*(G) = \mu^*(A)$.

Acabem de veure que per tot conjunt $B \in \mathcal{A}_{\mu^*}$, existeix un $H \in \sigma(A)$ tal que $B \subseteq H$ i $\mu^*(B) = \mu^*(H)$. Si apliquem el resultat a $G - A$ tindrem $G - A \subseteq F$, $F \in \sigma(\mathcal{A})$, $\mu^*(F) = 0$. Finalment, $A = (G - F) \cup (A \cap F)$ on $G - F \in \sigma(\mathcal{A})$ i $\mu(A \cap F) = 0$, és a dir $A \in \overline{\sigma(\mathcal{A})}$. □

[5] N pot no pertànyer a $\mathcal{A}$

[6] usant la notació dels teoremes de Carathéodory

Reduirem encara més la col·lecció de conjunts sobre els que cal especificar una mesura per tal de tenir-la determinada. Concretament veurem al següent teorema que és suficient fer-ho sobre una semiàlgebra.

Teorema 1.4.5. *Sigui $\mathcal{S}$ una semiàlgebra de Ω i μ una funció additiva definida sobre $\mathcal{S}$, és a dir $\mu : \mathcal{S} \longrightarrow [0, +\infty]$. Aleshores,*

(i) Existeix una única extensió additiva de μ a sobre $a(\mathcal{S})$.

(ii) Si a més μ és σ-additiva en $\mathcal{S}$, també ho serà l'extensió a $a(\mathcal{S})$.

(iii) Si a més μ és σ-finita en $\mathcal{S}$, també ho serà l'extensió a $a(\mathcal{S})$.

Demostració.

(i) Sigui $A \in a(\mathcal{S})$, per la Proposició 1.1.8 sabem que $a(\mathcal{S})$ és la família de les unions finites de conjunts de $\mathcal{S}$ disjunts dos a dos, aleshores tenim $A = \cup_{i=1}^n A_i$, $A_i \in \mathcal{S}$ amb $A_i \neq A_j$ si $i \neq j$. Podem escriure,

$$\mu(A) = \sum_{i=1}^n \mu(A_i).$$

Aquesta definició es consistent ja que si $\cup_{i=1}^n A_i = \cup_{j=1}^m B_j$ i ambdues descomposicions són de conjunts disjunts de $\mathcal{S}$ tindrem,

$$\begin{aligned}\sum_{i=1}^n \mu(A_i) &= \sum_{i=1}^n \mu\Big(\bigcup_{j=1}^m (A_i \cap B_j)\Big) = \sum_{i=1}^n \sum_{j=1}^m \mu(A_i \cap B_j) \\ &= \sum_{j=1}^m \mu\Big(\bigcup_{i=1}^n (A_i \cap B_j)\Big) = \sum_{j=1}^m \mu(B_j),\end{aligned}$$

ja que tots els conjunts $A_i \cap B_j$ són de $\mathcal{S}$.

D'altra banda, veiem que μ és additiva sobre $a(\mathcal{S})$: si $A = \cup_{i=1}^n A_i$ i $B = \cup_{j=1}^m B_j$ són dos conjunts disjunts de $a(\mathcal{S})$ amb els A_i i B_j de $\mathcal{S}$, tenim

$$\mu(A \cup B) = \mu\Big(\Big(\bigcup_{i=1}^n A_i\Big) \bigcup \Big(\bigcup_{j=1}^m B_j\Big)\Big) = \sum_{i=1}^n \mu(A_i) + \sum_{j=1}^m \mu(B_j) = \mu(A) + \mu(B).$$

Observem finalment que l'extensió és única: Si μ_1 i μ_2 coincideixen sobre $\mathcal{S}$ i són additives sobre $a(\mathcal{S})$, per a tot $A = \cup_{i=1}^n A_i$ amb els A_i disjunts i de $\mathcal{S}$, tenim

$$\mu_1(A) = \mu_1\Big(\bigcup_{i=1}^n A_i\Big) = \sum_{i=1}^n \mu_1(A_i) = \sum_{i=1}^n \mu_2(A_i) = \mu_2(A).$$

(ii) Si ara tenim que μ és σ-additiva sobre $\mathcal{S}$, també ho serà sobre $a(\mathcal{S})$. Sigui $\{A_n, n \geq 1\}$ una família d'elements disjunts de $a(\mathcal{S})$ que compleixen $A = \cup_{n=1}^\infty A_n \in a(\mathcal{S})$. Tenim que si $A = \cup_{j=1}^m C_j$ i $A_n = \cup_{i=1}^{m_n} B_i^{(n)}$, amb C_j i $B_i^{(n)}$ respectivament disjunts i de $\mathcal{S}$. Aleshores, usant la σ-additivitat sobre $\mathcal{S}$ i l'additivitat sobre $a(\mathcal{S})$ tenim:

$$\begin{aligned}\mu(A) &= \sum_{j=1}^m \mu(C_j) = \sum_{j=1}^m \mu(C_j \cap A) = \sum_{j=1}^m \mu\Big(\bigcup_{n=1}^\infty (C_j \cap A_n)\Big) \\ &= \sum_{j=1}^m \mu\Big(\bigcup_{n=1}^\infty \bigcup_{i=1}^{m_n} (C_j \cap B_i^{(n)})\Big) = \sum_{j=1}^m \sum_{n=1}^\infty \sum_{i=1}^{m_n} \mu(C_j \cap B_i^{(n)}) \\ &= \sum_{n=1}^\infty \mu\Big(\bigcup_{j=1}^m \bigcup_{i=1}^{m_n} (C_j \cap B_i^{(n)})\Big) = \sum_{n=1}^\infty \mu(A_n \cap A) = \sum_{n=1}^\infty \mu(A_n),\end{aligned}$$

i això demostra la σ-additivitat de μ sobre $a(\mathcal{S})$.

(iii) Si μ és σ-finita sobre $\mathcal{S}$ això vol dir que existeix una partició de Ω en conjunts Ω_n de $\mathcal{S}$ tal que $\mu(\Omega_n) < \infty$. La mateixa partició és de $a(\mathcal{S})$ i, per tant, μ és σ-finita sobre $a(\mathcal{S})$.

□

El que hem vist en aquest apartat juntament amb els resultats dels teoremes de Caratheodory ens diu que si μ és una mesura σ-additiva generada per $\mathcal{S}$, μ es pot estendre a una mesura sobre $\sigma(\mathcal{S})$: la σ-àlgebra generada per $\mathcal{S}$. Si tenim a més que μ és σ-finita, només hi ha una extensió possible.

1.5 Probabilitats en la recta real i funcions de distribució

El conjunt de resultats de moltes experiències aleatòries és un conjunt de nombres reals, això ens porta a estudiar les probabilitats sobre la recta real mitjançant les funcions de distribució. Estudiarem en aquest apartat el cas particular important de l'espai mesurable $(\mathbb{R}, \mathcal{B})$. Donada $\mathcal{S}$ una semiàlgebra de parts de $\mathbb{R}$ formada pels intervals $(a, b]$ amb $a < b$, les semirectes $(a, +\infty), (-\infty, a], a \in \mathbb{R}, \emptyset$ i $\mathbb{R}$, la σ-àlgebra generada per $\mathcal{S}$ s'anomena la σ-àlgebra de Borel, $\mathcal{B}$, i coincideix amb la σ-àlgebra generada pels conjunts oberts o tancats de $\mathbb{R}$.

Utilitzant els Teoremes de Carathéodory podem definir una mesura de probabilitat sobre $(\mathbb{R}, \mathcal{B})$, una possibilitat és definir una mesura de probabilitat σ-additiva sobre $\mathcal{S}$. El concepte de funció de distribució i la seva caracterització ens permetran fer-ho.

Proposició 1.5.1. *Si P és una probabilitat en $(\mathbb{R}, \mathcal{B})$ es defineix la seva funció de distribució F de la següent manera,*

$$\begin{aligned} F : \mathbb{R} &\longrightarrow [0, 1] \\ x &\longmapsto P((-\infty, x]), \end{aligned}$$

que té les següents propietats:

(i) F és creixent.

(ii) F és contínua per la dreta.

(iii) Es compleix que $\lim_{x \to +\infty} F(x) = 1$.

(iv) Es compleix que $\lim_{x \to -\infty} F(x) = 0$.

Demostració.

(i) En efecte si $x \leq y$ aleshores es té $(-\infty, x] \subseteq (-\infty, y]$ i es dedueix:

$$F(x) = P((-\infty, x]) \leq P((-\infty, y]) = F(y).$$

(ii) Demostrem la continuïtat per successions, és dir, donada una successió decreixent x_n de $\mathbb{R}$ que convergeix a x [7], es compleix que $(-\infty, x] = \cap_{n=1}^{\infty}(-\infty, x_n]$. Per les propietats de continuïtat de les mesures de la Proposició 1.2.6 obtenim,

$$F(x) = P((-\infty, x]) = \lim_n P((-\infty, x_n]) = \lim_n F(x_n).$$

[7] a partir d'ara usarem la notació $x_n \downarrow x$

(iii) En efecte, per tota successió $x_n \uparrow \infty$, es compleix que $\lim_n F(x_n) = 1$ ja que $\cup_{n=1}^{\infty}(-\infty, x_n] = \Omega$.

(iv) En efecte, per tota successió $x_n \downarrow -\infty$, es compleix que $\lim_n F(x_n) = 0$ ja que $\cap_{n=1}^{\infty}(-\infty, x_n] = \emptyset$.

□

Proposició 1.5.2. *Si P_1 i P_2 són dues mesures de probabilitat a $\mathbb{R}$ i les seves funcions de distribució F i G coincideixen, les probabilitats coincideixen sobre $\mathcal{S}$.*

Demostració. Si tenim $F(x) = G(x)$ per tot $x \in \mathbb{R}$, tenim $P_1((-\infty, x]) = P_2((-\infty, x])$. Donats $a, b \in \mathbb{R}$ amb $a < b$, tenim:

- $P_1((a, b]) = F(b) - F(a) = G(b) - G(a) = P_2((a, b])$.
- $P_1((a, \infty]) = 1 - F(a) = 1 - G(a) = P_2((a, \infty))$.

□

Aleshores tenim que si dues probabilitats P_1 i P_2 tenen la mateixa funció de distribució F, coincideixen sobre la semiàlgebra $\mathcal{S}$ i, en conseqüència $P_1 = P_2$. És a dir, la funció de distribució d'una probabilitat caracteritza completament aquesta probabilitat. El següent Teorema ens diu que, com a conseqüència dels resultats d'extensió de mesures, hi ha equivalència entre funcions de distribució i probabilitats sobre $\mathbb{R}$.

Teorema 1.5.3. *Donada una funció de distribució F existeix una única probabilitat P sobre $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ tal que $F(x) = P((-\infty, x])$ per a tot $x \in \mathbb{R}$.*

Demostració. Sigui F una funció de distribució complint les quatre propietats, podem definir P sobre la semiàlgebra $\mathcal{S}$ de la següent manera: per $a, b \in \mathbb{R}$ amb $a < b$,

$$P((a, b]) = F(b) - F(a), \quad P((-\infty, b]) = F(b), \quad P((a, \infty) = 1 - F(a),$$
$$P(\mathbb{R}) = 1 \quad \text{i} \quad P(\emptyset) = 0.$$

Si veiem que aquesta mesura P és σ-additiva sobre la semiàlgebra $\mathcal{S}$, com que ja sabem que és finita i per tant σ-finita, usant els teoremes de Carathéodory tindrem que admet una extensió única a $\mathcal{B}$ la σ-àlgebra de Borel [8].
Per les propietats de F, suposant que $a < c < b$ es compleix que $(a, b] = (a, c] \cup (c, b]$ i tenim $P((a, b]) = P((a, c]) + P((c, b])$, com que podem fer el mateix per les particions finites de les semirectes, tenim que P és additiva sobre $\mathcal{B}$. Per veure la σ-additivitat prenem $(a, b] = \cup_{n \geq 1}(a_n, b_n]$ amb $b_n = a_{n+1}$ per tot n, $(a, b]$,$(a_n, b_n] \in \mathcal{S}$ i els $(a_n, b_n]$ disjunts dos a dos. Recordem que P admet una extensió additiva a l'algebra generada per $\mathcal{S}$, aleshores per $m \geq 1$,

$$\sum_{n=1}^{m} P((a_n, b_n]) = P\Big(\bigcup_{n=1}^{m}(a_n, b_n]\Big) \leq P((a, b]),$$

i passant al límit tenim,

$$\sum_{n=1}^{\infty} P((a_n, b_n]) \leq P((a, b]).$$

Per la desigualtat contrària, denotem $I := (a, b]$ i $I_n = (a_n, b_n]$, pel que $I = \cup_{n=1}^{\infty} I_n$. Necessitem utilitzar la següent propietat:

[8]la σ-àlgebra generada per la semiàlgebra $\mathcal{S}$ de $\mathbb{R}$

Donat $\varepsilon > 0$ i $J \in \mathcal{S}$, existeixen J_1, $J_2 \in \mathcal{S}$ tals que $\overline{J_1}$ és compacte, $\overline{J_1} \subseteq J \subseteq \mathring{J}_2$, $P(J) - P(J_1) < \varepsilon$ i $P(J_2) - P(J) < \varepsilon$.
Considerem les diferents possiblitats:

(i) Si $J = (a, b]$ prenem $J_1 = (a', b]$ i $J_2 = (a, b']$ amb $a < a'$, $b < b'$, $F(a') - F(a) < \varepsilon$ i $F(b') - F(b) < \varepsilon$.

(ii) Si $J = (a, +\infty)$ prenem $J_1 = (a', c]$ i $J_2 = J$ amb $a < a' < c$, $F(a') - F(a) < \frac{\varepsilon}{2}$ i $1 - F(c) < \frac{\varepsilon}{2}$.

(iii) Si $J = (-\infty, b]$ prenem $J_1 = (a', b]$ i $J_2 = (-\infty, b']$ amb $a' < b < b'$, $F(a') < \varepsilon$ i $F(b') - F(b) < \varepsilon$.

Fixem $\varepsilon > 0$ i sigui $I' \in \mathcal{S}$, amb $\overline{I'}$ compacte tal que $\overline{I'} \subseteq I$ i $P(I) - P(I') < \frac{\varepsilon}{2}$. També per cada $n \geq 1$ prenem $I'_n \in \mathcal{S}$ tals que $I_n \subseteq \mathring{I}'_n$ i $P(I'_n) - P(I_n) < \frac{\varepsilon}{2^{n+1}}$. Com que $\{\mathring{I}'_n, n \geq 1\}$ és un recobriment obert del compacte $\overline{I'}$ existirà un subrecobriment finit $I' \subseteq \cup_{n=1}^m \mathring{I}'_n$. Aleshores,

$$P(I) < P(I') + \frac{\varepsilon}{2} \leq \sum_{n=1}^{m} P(I'_n) + \frac{\varepsilon}{2} \leq \sum_{n=1}^{m} P(I_n) + \sum_{n=1}^{m} \frac{\varepsilon}{2^{n+1}} + \frac{\varepsilon}{2} \leq \sum_{n=1}^{\infty} P(I_n) + \varepsilon,$$

això demostra l'altra desigualtat i, per tant, la σ-additivitat de la mesura P sobre la semiàlgebra $\mathcal{S}$. □

Observació 1.5.4. Donada una probabilitat P amb funció de distribució F les següents fórmules es compleixen i ens permeten calcular la probabilitat de qualsevol interval, siguin $a, b \in \mathbb{R}$ amb $a < b$, tenim

$$P(a) = F(a) - F(a^-), \quad P([a, b]) = F(b) - F(a^-), \quad P([a, b)) = F(b^-) - F(a^-).$$

Notem que no tota mesura μ sobre $(\mathbb{R}, \mathcal{B})$ es pot obtenir a partir d'una funció de distribució mitjançant la fòrmula $\mu((a, b]) = F(b) - F(a)$, ja que totes les mesures construïdes d'aquesta manera donen una mesura finita per a tot interval acotat. Veurem que són precisament aquestes les que es poden determinar amb funcions de distribució. Una àmplia classe de mesures, en particular totes les mesures finites i per tant les probabilitats, compleixen aquesta propietat, els hi donem un nom en la definció següent.

Definició 1.5.5. *Una mesura de Lebesgue-Stieljes és una mesura μ en $(\mathbb{R}, \mathcal{B})$ tal que per tot $A \in \mathcal{B}(\mathbb{R})$ acotat es compleix que $\mu(A) < \infty$.*

La probabilitat anomenada *Delta de Dirac* en $a \in \mathbb{R}$, que designarem δ_a, ve donada per la funció de distribució següent: es té, per qualsevol $A \in \mathcal{B}(\mathbb{R})$,

$$\begin{aligned} \delta_a : \mathcal{B}(\mathbb{R}) &\longrightarrow [0, 1] \\ A &\longmapsto \delta_a(A) = \begin{cases} 1, & si \quad a \in A \\ 0, & si \quad a \notin A. \end{cases} \end{aligned}$$

La necessitem per definir les probabilitats discretes.

Definició 1.5.6. *Es diu que una mesura de probabilitat P a $\mathbb{R}$ és discreta si i només si existeix un conjunt finit o numerable I tal que,*

$$P = \sum_{i \in I} p_i \cdot \delta_{a_i},$$

amb $p_i \in (0, 1]$, $\sum_{i \in I} p_i = 1$ i δ_{a_i} són Deltes de Dirac en els punts $a_i \in \mathbb{R}$ per a tot $i \in I$.

Observació 1.5.7. Observem que la funció de distribució d'una mesura de probabilitat discreta és la funció esglaonada:

$$F(x) = P((-\infty, x]) = \sum_{i:a_i \leq x} p_i,$$

on P modela una experiència aleatòria en la qual s'obtenen els resultats a_i amb probabilitat p_i.

Definició 1.5.8. *És diu que una mesura de probabilitat és contínua si i només si $P(x) = 0$ per a tot $x \in \mathbb{R}$.*

Observem que P és contínua si i només si la F ho és, i això implica que $P(x) = 0$.

Si tenim P una probabilitat i F la seva funció de distribució, la continuïtat per la dreta, juntament amb l'existència del $\lim_{x \to x^-} F(x)$ per tota $x \in \mathbb{R}$, ens asseguren que F només té discontinuïtats de salt. A més, com que $F(x) - F(x^-) = P(\{X = x\})$, tenim que la longitud del salt de F en cada punt és la probabilitat de que la variable tingui tal valor.

Passem a veure que el conjunt dels punts de discontinuïtat d'una funció de distribució és nul, finit o numerable.

Teorema 1.5.9. *El conjunt de punts de discontinuïtat de la funció de distribució d'una variable aleatòria és numerable.*

Demostració. Sigui S el conjunt de punts de discontinuïtat d'una funció de distribució F. Per tot $n \in \mathbb{N}$, definim el conjunt $E_n \subseteq S$, com el conjunt de punts en que la longitud de salt és major o igual a $\frac{1}{n}$, és a dir:

$$S = \{x \in \mathbb{R};\, F(x) - F(x^-) > 0\} \quad \text{i} \quad E_n = \{x \in \mathbb{R};\, F(x) - F(x^-) \geq \tfrac{1}{n}\}.$$

Aleshores tenim que $x \in S$ si i només si es compleix que existeix $n \in \mathbb{N}$ de manera que $F(x) - F(x) \geq \frac{1}{n}$, és a dir, si i només si, per tot $n \in \mathbb{N}$ tenim $x \in E_n$ o, anàlogament, $x \in \cup_{n=1}^{\infty} E_n$. Per tant, $S = \cup_{n=1}^{\infty} E_n$. Notem també que per tot $n \in \mathbb{N}$, el conjunt E_n té com a màxim n punts perquè, en cas contrari, la suma dels salts de F seria més gran que 1, però això no pot passar ja que $F(-\infty) = 0$, $F(+\infty) = 1$ i F no és decreixent. Per tant, $S = \cup_{n=1}^{\infty} E_n$ és una unió numerable de conjunts finits i és, per tant, numerable. ☐

Proposició 1.5.10. *Tota mesura de probabilitat a $\mathbb{R}$ es descompon de manera única en una combinació lineal d'una mesura de probabilitat contínua i una mesura de probabilitat discreta.*

Demostració. Sigui P una mesura de probabilitat sobre $\mathbb{R}$ i F la seva funció de distribució. Sabem que el conjunt de discontinuïtats de F és nul, finit o numerable, el denotem S. Aleshores podem definir,

$$\mu_d := \sum_{y \in S} (F(y) - F(y-))\delta_y,$$

que és una mesura finita de massa $\mu_d(\mathbb{R}) = \mu_d(S) = P(S) \leq 1$ i $\mu_c := P - \mu_d$, que és una altra mesura finita de massa $1 - \mu_d(S) = 1 - P(S) = P(S^c) \leq 1$. Aleshores, per a qualsevol borelià $B \in \mathcal{B}$, podem definir les probabilitats següents:

$$P_d(B) = \frac{\mu_d(B)}{P(S)} \quad \text{i} \quad P_c(B) = \frac{\mu_c(B)}{P(S^c)}.$$

Observem que, $P(B) = P_d(B)P(S) + P_c(B)P(S^c) = \mu_d(B) + \mu_c(B)$ i, en particular, pel Teorema de les probabilitats totals tenim $P_d(B) = P(B|S)$ i $P_c(B) = P(B|S^c)$ i això demostra la unicitat de la descomposició. Falta veure que P_d és una probabilitat discreta i que P_c és una probabilitat contínua. En el primer cas tenim,

$$P_d := \sum_{y\in S} \frac{F(y) - F(y-)}{P(S)} \delta_y \quad \text{amb} \quad \sum_{y\in S} \frac{F(y) - F(y-)}{P(S)} = 1,$$

pel que és una probabilitat discreta. En el segon cas tenim $\mu_c(x) = P(x) - \mu_d(x)$, i aquesta expressió val sempre 0. $\square$

Considerem F una funció de distribució d'una probabilitat μ i S el conjunt nul, finit o numerable dels seus punts de discontinuïtat. Com ja hem vist, es defineix la funció de distribució de la mesura discreta com $\mu_d := \sum_{y\in S}(F(y) - F(y-))\delta_y$. Si definim la part purament discontínua de F com $F_d := \sum_{y\in S}[F(y) - F(y-)]$, és la funció de distribució de la mesura discreta μ_d.
Notem que $S = \{y : \mu(y) > 0\}$ i $\mu_d = \sum_{y\in S} \mu(y)\delta_y$, d'aquí es dedueix que la probabilitat μ és discreta si $\mu = \mu_d$ i que la diferència $\mu_c = \mu - \mu_d$ és una mesura contínua que té per funció de distribució $F_c = F - F_d$.

Els resultats anteriors es poden estendre al cas de probabilitats de $\mathbb{R}^n$. Descriurem breument aquesta extensió sense entrar en detalls: Fixats dos punts $a = (a_1, ..., a_n)$, $b = (b_1, ..., b_n)$ de $\mathbb{R}^n$ tals que $a_i < b_i$ per a tot $i \in \{1, ..., n\}$ es defineix el rectangle n-dimensional $(a, b]$ per $(a, b] = \prod_{i=1}^n (a_i, b_i]$ [9]. La família d'aquests rectangles afegint el conjunt buit constitueixen una σ-àlgebra de Borel de $\mathbb{R}^n$.

Definició 1.5.11. *Sigui μ una probabilitat en l'espai mesurable $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$. Si tenim $X_1, ..., X_n$ variables aleatòries, definim la seva funció de distribució conjunta com la funció de distribució n-dimensional de la llei del vector aleatori $X = (X_1, ..., X_n)$, és a dir, per $x = (x_1, ..., x_n) \in \mathbb{R}^n$,*

$$\begin{aligned} F\colon\ \mathbb{R}^n &\longrightarrow [0, 1] \\ x &\longmapsto F(x) = \mu((-\infty, x]) = \mu(\{X_1 \le x_1, ..., X_n \le x_n\}), \end{aligned}$$

Com en el cas real, la funció de distribució determina la llei de X.

Pasem a veure, fent servir la mateixa notació, algunes de les seves propietats:

(i) $\triangle_{ab} F \le 0$, per tota parella de punts a, b de $\mathbb{R}^n$ tals que $a_i < b_i$ amb $i \in \{1, ..., n\}$ on, per definició:

$$\triangle_{ab} F = \sum_{\substack{\varepsilon\in\{0,1\}^n \\ \varepsilon=(\varepsilon_1,...,\varepsilon_n)}} (-1)^{\varepsilon_1+...+\varepsilon_n} F(b_1 + \varepsilon_1(a_1 - b_1), ..., b_n + \varepsilon_n(a_n - b_n)).$$

$\triangle_{ab} F$ representa l'increment n-dimensional de la funció F en el rectangle $(a, b]$. La positivitat d'aquest increment es redueix del fet que coincideix amb $\mu((a, b])$.

(ii) F és contínua per la dreta en totes les variables simultàniament: $\lim_{x_i\downarrow y_i, i\in\{1,...,n\}} F(x) = F(y)$.

(iii) Es compleixen les següents propietats asimptòtiques següents:

$$\lim_{\substack{x_i\to\infty \\ i\in\{1,...,n\}}} F(x) = 1 \quad \text{i} \quad \lim_{\substack{x_i\to-\infty \\ i\in\{1,...,n\}}} F(x) = 0.$$

[9] on les coordenades a_i poden ser $-\infty$ i les coordenades b_i poden ser $+\infty$

Tota funció F: $\mathbb{R}^n \longrightarrow [0,1]$ que compleix aquestes propietats és funció de distribució; notem que existeix una única probabilitat μ en $\mathbb{R}^n$ tal que $F(x) = \mu((-\infty, x])$ [10].

També veurem que es poden construïr probabilitats sobre la recta real mitjançant funcions de densitat i les integrals de Lebesgue. Més endevant, quan veiem el Teorema de Radon-Nikodym (Teorema 2.2.23), introduïrem els conceptes de densitat de probabilitat i funcions de distribució absolutament contínues.

[10] la unicitat prové del fet que dues probabilitats μ_1, μ_2 amb la mateixa funció de distribució F han de coincidir sobre la semiàlgebra $\mathcal{S}$ ja que $\mu_1((a,b]) = \mu_2((a,b]) = \triangle_{ab} F$

Capítol 2

Funcions mesurables i esperança matemàtica

L'objectiu d'aquesta secció és caracteritzar l'esperança matemàtica i per fer-ho introduirem previament la noció de variable aleatòria així com els resultats necessaris d'integració de funcions mesurables.

2.1 Funcions mesurables i variables aleatòries

En aquest apartat definirem les funcions mesurables, les variables i vectors aleatoris i les funcions elementals.

Definició 2.1.1. *Siguin* $(\Omega_1, \mathcal{A}_1), (\Omega_2, \mathcal{A}_2)$ *espais mesurables i* $X : \Omega_1 \longrightarrow \Omega_2$ *una aplicació. Direm que* X *és una funció mesurable respecte* $\mathcal{A}_1$ *i* $\mathcal{A}_2$ *si i només si per tot* $A \in \mathcal{A}_2$ *es compleix que* $X^{-1}(A) \in \mathcal{A}_1$. *Abreviant diem que* X *és* $(\mathcal{A}_1, \mathcal{A}_2)$*-mesurable.*

Proposició 2.1.2. *La composició d'aplicacions mesurables és una aplicació mesurable.*

Demostració. Sigui X una aplicació mesurable $X : \Omega \longrightarrow E$ i Y una aplicacio mesurable $Y : E \longrightarrow G$. Denotem per $\mathcal{G}$ la σ-àlgebra associada a G, $\mathcal{A}$ la σ-àlgebra associada a Ω i $\mathcal{E}$ la σ-àlgebra associada a E. Aleshores la composició $Y \circ X$ és una aplicació mesurable de Ω en G ja que per a tot subconjunt $A \in \mathcal{G}$ la seva antiimatge per Y pertany a $\mathcal{E}$ i l'antiimatge d'aquest conjunt per X pertant a $\mathcal{A}$, per ser Y i X mesurables. Per tant, $(Y \circ X)^{-1}(A) = X^{-1}(Y^{-1}(A)) \in \mathcal{A}$, i $Y \circ X$ és una aplicació mesurable· □

Observació 2.1.3. Notem que per comprovar que una aplicació $X : (\Omega_1, \mathcal{A}_1) \longrightarrow (\Omega_2, \mathcal{A}_2)$ és mesurable cal comprovar únicament que $X^{-1}(B) \in \mathcal{A}_1$ per a tot $B \in \mathcal{G}$ on $\mathcal{G}$ és una família de conjunts que genera la σ-àlgebra $\mathcal{A}_2$.

Definició 2.1.4. *Sigui* $\{(E_i, \mathcal{E}_i), i \in I\}$ *una familia arbitrària d'espais mesurables i* $\{X_i : \Omega \longrightarrow E_i, i \in I\}$ *una familia d'aplicacions. S'anomena* σ*-àlgebra generada per* $\{X_i, i \in I\}$[1] *a la* σ*-àlgebra de* $\mathcal{P}(\Omega)$ *més petita que fa mesurables totes les aplicacions de la familia* $\{X_i, i \in I\}$.

Sigui $(\Omega, \mathcal{A}, P)$ un espai de probabilitat general associat a una certa experiència aleatòria.

Definició 2.1.5. *Definim una variable aleatòria associada a una certa experiència aleatòria en un espai de probabilitat* $(\Omega, \mathcal{A}, P)$ *com una aplicació mesurable* $X : \Omega \longrightarrow \mathbb{R}$*, a la*

[1] també se l'anomena σ-àlgebra inicial de la familia $\{X_i, i \in I\}$

qual cada resultat possible ω de l'experiència li correspon un nombre real $X(\omega)$ [2]. *Anàlogament una aplicació mesurable $(X_1, ..., X_n) : \Omega \longrightarrow \mathbb{R}^n$ s'anomena un vector aleatori* [3].

Proposició 2.1.6. *Si $\{X_n\}_{n\in\mathbb{N}}$ és una successió de variables aleatòries, també seràn mesurables les següents funcions:* $\inf_n X_n$, $\sup_n X_n$, $\liminf_n X_n$ i $\limsup_n X_n$.

Demostració. Aquesta propietat es dedueix de les següents relacions:

$$\{\sup_n X_n \le a\} = \bigcap_n \{X_n \le a\} \in \mathcal{A} \text{ per a tot } a \in \mathbb{R},$$

$$\sup_n X_n = -\inf_n(-X_n), \quad \liminf_n X_n = \sup_n \inf_{m\ge n} X_m \quad \text{i} \quad \limsup_n X_n = \inf_n \sup_{m\ge n} X_m. \quad \square$$

Notem que les idees de la demostració anterior permeten demostrar que el màxim i el mínim de dues variables aleatòries són també variables aleatòries, i en particular:

$$X^+ := X \vee 0\ , \quad X^- := -(X \wedge 0) \quad \text{i} \quad |X| = X^+ + X^-, \quad \text{són variables aleatòries.}$$

Definició 2.1.7. *Si $A \in \mathcal{P}(\Omega)$ la funció $\mathbb{1}_A : \Omega \longrightarrow \mathbb{R}$ s'anomena indicador del conjunt A i està definida de la manera següent:*

$$\mathbb{1}_A(\omega) := \begin{cases} 1, & si \quad \omega \in A \\ 0, & si \quad \omega \notin A \end{cases}$$

Proposició 2.1.8. *Sigui $(\Omega, \mathcal{A}, P)$ un espai de probabilitats. Aleshores tenim que $\mathbb{1}_A$ és una variable aleatòria si, i només si, $A \in \mathcal{A}$.*

Demostració. Observem que si $B \in \mathcal{B}(\mathbb{R})$, $\mathbb{1}_A^{-1}(B)$ només pot ser un dels conjunts següents: $\Omega, A, A^c, \emptyset$. Per tant, si A és mesurable tots aquests conjunts ho són. En cas contrari, no. $\square$

Definició 2.1.9. *Una funció mesurable $X : \Omega \longrightarrow \mathbb{R}$ diem que és elemental si pren un nombre fint de de valors diferents. Això és equivalent a dir que X és una combinació lineal finita de funcions indicatrius mesurables:*

$$X = \sum_{i=1}^{n} a_i \cdot \mathbb{1}_{A_i},$$

on $a_i \neq a_j$ i $A_i \in \mathcal{A}$, amb $A_i \cap A_j = \emptyset$ si $i \neq j$ [4].

Poden haver altres definicions de les funcions mesurables elementals depenent de com definim els a_i i els A_i. En el cas escollit, la tria dels a_i diferents entre ells i dels A_i disjunts dos a dos estalvia justificacions. Ens interessarem per les funcions elementals positives, és a dir, les variables aleatòries que prenen un nombre finit de valors a $[0, +\infty)$.

Observació 2.1.10. El conjunt $\mathcal{E}$ que denota les funcions elementals sobre un espai mesurable $(\Omega, \mathcal{A})$ és un espai vectorial real respecte les operacions habituals de suma i producte per escalars, ja que al fer combinacions lineals d'elements de $\mathcal{E}$ s'obtenen de nou funcions mesurables que prenen només un conjunt finit de valors diferents.

[2] una aplicació mesurable de $(\Omega, \mathcal{A})$ a $(\mathbb{R}, \mathcal{B})$

[3] les components X_j amb $j \in \{1, ..., n\}$ són variables aleatòries

[4] Notem que la respresentació és única ja que els $\{a_1, ..., a_n\}$ són els valors que pren la variable X i d'altra banda $A_i = \{X = a_i\}$

Proposició 2.1.11. *Tota funció mesurable $X : \Omega \longrightarrow [0, +\infty]$ és límit d'una successió creixent de funcions elmentals.*

Demostració. Definim per a cada $n \in \mathbb{N}$,

$$X_n := \sum_{k=1}^{n2^n-1} \frac{k}{2^n} \cdot \mathbb{1}_{\{\frac{k}{2^n} \leq X < \frac{k+1}{2^n}\}} + n \cdot \mathbb{1}_{\{n \leq X\}}.$$

És rutinari comprovar que les X_n són funcions elementals, que formen una successió creixent i que $\lim_{n\to\infty} X_n(\omega) = X(\omega)$, per tot $\omega \in \Omega$. □

Proposició 2.1.12. *Tota funció mesurable $X : \Omega \longrightarrow \mathbb{R}$ és límit puntual de funcions elmentals.*

Demostració. Les funcions X^+ i X^- són funcions mesurables positives. Sigui $\{Y_n\}_{n\in\mathbb{N}}$ una successió creixent de funcions elementals tal que $\lim_{n\to\infty} Y_n = X^+$. Sigui$\{Z_n\}_{n\in\mathbb{N}}$ una successió creixent de funcions elementals tal que $\lim_{n\to\infty} Z_n = X^-$. Aleshores,

$$\lim_{n\to\infty} (Y_n - Z_n) = X^+ - X^- = X.$$

□

2.2 Integrabilitat

Sempre que tenim definida una mesura μ en un espai mesurable $(\Omega, \mathcal{A})$, podem construir una integral relativa a ella. El procés que porta a la construcció de la integral pot reproduir-se de manera semblant per a espais de mesura qualsevol: Consisteix en una construcció per passos, on inicialment considerem la integral d'una funció elemental positiva, en segon lloc sobre una funció mesurable positiva i, per últim, definim la integral sobre una funció mesurable arbitrària. La teoria resultant també s'anomena Teoria de la Integral de Lebesgue.

Suposarem a partir d'ara que totes les mesures de les que parlem són no trivials (és a dir $\mu \neq 0, \mu \neq +\infty$). Considerem l'espai de mesura $(\Omega, \mathcal{A}, \mu)$ i

$$X = \sum_{i=1}^{n} a_i \cdot \mathbb{1}_{A_i},$$

una funció elemental positiva. Recordem que els a_i són tots diferents i els A_i són disjunts dos a dos [5]. Es defineix la integral de X respecte μ com:

$$\int_\Omega X d\mu = \sum_{i=1}^{n} a_i \mu(A_i).$$

La integral [6] pren valors a $[0, +\infty)$, recordem que per la tria *canònica* que hem fet de la funció elemental la representació és única. Passem a veure les propietats de la integral de funcions elementals positives.

[5]Farem servir el conveni $0 \cdot \infty = 0$ quan algun dels conjunts A_i tingui mesura infinita i a_i valgui zero

[6]Utilitzarem també les següents notacions per la integral

$$\int_\Omega X(\omega) d\mu(\omega) \quad \text{o bé} \quad \int_\Omega X(\omega)\mu(d\omega)$$

Proposició 2.2.1. *Siguin X, Y funcions elementals positives. La integral respecte μ de funcions elementals compleix les propietats següents:*

(i) Si a és un nombre real no negatiu, aleshores es compleix $\int_\Omega aXd\mu = a\int_\Omega Xd\mu$.

(ii) $\int_\Omega (X+Y)d\mu = \int_\Omega Xd\mu + \int_\Omega Yd\mu$.

(iii) Si $X \leq Y$, aleshores es compleix $\int_\Omega Xd\mu \leq \int_\Omega Yd\mu$.

(iv) Si $\{X_n\}_{n\in\mathbb{N}}$ és una successió creixent de variables aleatòries simples i positives i X és una variable aleatòria elemental positiva tal que $X = \lim\limits_{n\to\infty} X_n$, aleshores es compleix $\int_\Omega X_n d\mu \uparrow \int_\Omega Xd\mu$.

Demostració.

(i) Sigui $X = \sum\limits_{i=1}^{n} a_i \cdot \mathbb{1}_{A_i}$. Tindrem que $aX = \sum\limits_{i=1}^{n}(aa_i)\cdot \mathbb{1}_{A_i}$. Aleshores,

$$\int_\Omega aXd\mu = \sum_{i=1}^{n}(aa_i)\mu(A_i) = a\sum_{i=1}^{n} a_i\mu(A_i) = a\int_\Omega Xd\mu.$$

(ii) Siguin $X = \sum\limits_{i=1}^{n} a_i \cdot \mathbb{1}_{A_i}$, i $Y = \sum\limits_{j=1}^{m} b_j \cdot \mathbb{1}_{B_j}$. Llavors,

$$\begin{aligned}\int_\Omega (X+Y)d\mu &= \int_\Omega \sum_{i=1,j=1}^{n,m}(a_i+b_j)\cdot \mathbb{1}_{A_i\cap B_j} = \sum_{i=1}^{n}\sum_{j=1}^{m}(a_i+b_j)\mu(A_i\cap B_j)\\ &= \sum_{i=1}^{n} a_i \sum_{j=1}^{m}\mu(A_i\cap B_j) + \sum_{j=1}^{m} b_j \sum_{i=1}^{n}\mu(A_i\cap B_j)\\ &= \sum_{i=1}^{n} a_i\mu(A_i) + \sum_{j=1}^{m} b_j\mu(b_j) = \int_\Omega Xd\mu + \int_\Omega Yd\mu.\end{aligned}$$

(iii) Siguin $X = \sum\limits_{i=1}^{n} a_i \cdot \mathbb{1}_{A_i}$, i $Y = \sum\limits_{j=1}^{m} b_j \cdot \mathbb{1}_{B_j}$. Llavors,

$$\begin{aligned}\int_\Omega Xd\mu &= \sum_{i=1}^{n} a_i\mu(A_i) = \sum_{i=1}^{n}\sum_{j=1}^{m} a_i\mu(A_i\cap B_j)\\ &\leq \sum_{i=1}^{n}\sum_{j=1}^{m} b_j\mu(A_i\cap B_j) = \sum_{j=1}^{m} b_i\mu(B_j) = \int_\Omega Yd\mu.\end{aligned}$$

La demostració *(iv)* la podem trobar al llibre [1, Proposició 2.2.8]. □

La segona etapa en la construcció de la integral consisteix, tal i com hem avançat a l'inici de la secció, en definir la integral d'una funció mesurable positiva $X : \Omega \longrightarrow [0, +\infty]$.

Definició 2.2.2. *Sigui $(\Omega, \mathcal{A}, \mu)$ un espai de mesura i X una funció mesurable positiva $X : \Omega \longrightarrow [0, +\infty)$. Definim la integral de X respecte μ de la següent manera:*

$$\int_\Omega Xd\mu = \lim_{n\to\infty}\int_\Omega X_n d\mu,$$

on $\{X_n\}_{n\in\mathbb{N}}$ és una successió creixent de funcions elementals positves que convergeix a X.

Notem que per la Proposició 2.1.11 podem justificar l'existencia d'aquesta successió. Per veure que està ben definida, és a dir que el valor de $\int_\Omega X d\mu$ no depèn de la successió $\{X_n\}_{n\in\mathbb{N}}$ escollida, veurem el següent lema.

Lema 2.2.3. *Siguin $\{X_n\}_{n\in\mathbb{N}}$ i $\{Y_n\}_{n\in\mathbb{N}}$ dues successions creixents de funcions elementals no negatives. Aleshores si $\lim\limits_{n\to\infty} X_n \leq \lim\limits_{n\to\infty} Y_n$ es compleix,*

$$\lim_{n\to\infty} \int_\Omega X_n d\mu \leq \lim_{n\to\infty} \int_\Omega Y_n d\mu.$$

Demostració. Fixat un $m \in \mathbb{N}$ qualsevol, tenim que $X_m = \lim\limits_{n\to\infty} \inf(X_m, Y_n)$. Utilitzant la propietat *(iv)* de la Proposició 2.2.1 obtenim,

$$\int_\Omega X_m d\mu = \lim_{n\to\infty} \int_\Omega \inf(X_m, Y_n) d\mu \leq \lim_{n\to\infty} \int_\Omega Y_n d\mu,$$

i fent tendir $m \to \infty$ obtenim $\lim\limits_{m\to\infty} \int_\Omega X_m d\mu \leq \lim\limits_{n\to\infty} \int_\Omega Y_n d\mu$. □

Aquest lema implica doncs, que si els límits de les funcions coincideixen, els límits de les integrals han de coincidir. Passem a veure algunes propietats de la nova construcció de la integral, notem que les propietats *(i),(ii),(iii)* de monotonia i linealitat de la Proposició 2.2.1 són anàlogues per les funcions mesurables positives [7].

Proposició 2.2.4. *Siguin X, Y funcions mesurables positives. La integral respecte μ de funcions mesurables positives compleix les propietats següents:*

(i) Si tenim $\int_\Omega X d\mu < +\infty$, aleshores es compleix $\mu(\{X = +\infty\}) = 0$. És a dir, si una funció té integral finita, la funció és finita a menys d'un conjunt de mesura zero.

(ii) Es compleix $\int_\Omega X d\mu = 0$ si i només si $\mu(\{X > 0\}) = 0$. És a dir, la integral d'una funció és zero si i només si la funció positiva és zero a menys d'un conjunt de mesura nul·la.

(iii) Si $\{X_n\}_{n\in\mathbb{N}}$ és una successió creixent de funcions mesurables positives que convergeixen puntualment cap a X, aleshores $\int_\Omega X_n d\mu \uparrow \int_\Omega X d\mu$.

Demostració.

(i) Posem $m = \int_\Omega X d\mu$. Tindrem $m \geq \int_\Omega X \cdot \mathbb{1}_{\{X\geq n\}} d\mu \geq n\mu(\{X \geq n\})$, i en conseqüència, $\mu(\{X = \infty\}) = \lim_n \mu(\{X \geq n\}) = 0$.

(ii) Suposem primer que $\int_\Omega X d\mu = 0$. Tindrem $\mu(\{X > 0\}) = \lim_n \mu(\{X \geq \frac{1}{n}\}) \leq \lim_n \int_\Omega nX d\mu = 0$, ja que $\mathbb{1}_{\{X\geq\frac{1}{n}\}} \leq nX$. Recíprocament,

$$\int_\Omega X d\mu = \int_\Omega X \cdot \mathbb{1}_{\{X>0\}} d\mu \leq \lim_n \int_\Omega n \cdot \mathbb{1}_{\{X>0\}} d\mu = 0.$$

La demostració *(iii)* la podem trobar al llibre [10, Proposició 6.2.2]. □

Finalment, definim la integral en el cas més general. La idea és descompondre la funció en la part positiva i en la part negativa, i fer que la part negativa contribueixi a la integral amb signe negatiu.

[7] ho podem veure per pas al límit

Definició 2.2.5. *Sigui $X : \Omega \longrightarrow \mathbb{R}$ una funció mesurable. Diem que X és integrable si compleix $\int_\Omega |X| d\mu < \infty$ i definim la integral de X respecte μ com*

$$\int_\Omega X d\mu = \int_\Omega X^+ d\mu - \int_\Omega X^- d\mu. \tag{2.2.1}$$

Observació 2.2.6. Equivalentment per veure que X és integrable podem comprovar si $\int_\Omega X d\mu$ existeix [8] i és finita.

Veiem en el resultat següent que les propietats enunciades a la Proposicó 2.2.1 es poden extendre de manera gairebé directe a la nova definció d'integral.

Proposició 2.2.7. *La integració respecte μ de funcions mesurables arbitràries té les propietats següents:*

(i) Si $a \in \mathbb{R}$, i la funció X és integrable, aleshores la funció aX és integrable i es compleix $\int_\Omega aX d\mu = a \int_\Omega X d\mu$.

(ii) Si les funcions X i Y són integrables, aleshores $X + Y$ és integrable i es compleix $\int_\Omega (X + Y) d\mu = \int_\Omega X d\mu + \int_\Omega Y d\mu$.

(iii) Si X i Y són integrables i $X \leq Y$, aleshores es compleix $\int_\Omega X d\mu \leq \int_\Omega Y d\mu$.

(iv) Si X és integrable aleshores $|\int_\Omega X d\mu| \leq \int_\Omega |X| d\mu$.

Demostració. Una prova detalla d'aquests resultats la podem trobar al llibre [6, Proposició 4.1]. □

Seguidament introduïm un concepte més feble per parlar de les propietats de la integral de funciones mesurables arbitràries. Notem que quan diem que una propietat és certa en un conjunt Ω, estem dient que *tots els elements $\omega \in \Omega$ la compleixen*, parlarem ara de la certesa de les propietats de manera *quasi segura*.

Definició 2.2.8. *Diem que una propietat és certa quasi per tot respecte μ (usarem la notació μ-q.p.t) si i només si es certa per a tot $\omega \in \Omega$ excepte d'un conjunt $N \in \mathcal{A}$ de mesura zero, és a dir $\mu(N) = 0$.*

Proposició 2.2.9. *Sigui X una funció integrable en un espai de mesura $(\Omega, \mathcal{A}, \mu)$ tal que $\int_A X d\mu = 0$ per a tot $A \in \mathcal{A}$. Aleshores $X = 0$ μ-q.p.t.*

Demostració. En particular tenim $\int_{\{X>0\}} X d\mu = 0$ i$\int_{\{X<0\}} (-X) d\mu = 0$. Mitjançant les propietats de la integral de les funcions no negatives s'obté $\mu(\{X > 0\}) = \mu(\{X < 0\}) = 0$. És a dir, $\mu(\{X \neq 0\}) = 0$. □

Proposició 2.2.10. *Considerem un espai de mesura $(\Omega, \mathcal{A}, \mu)$ i $X, Y :\longrightarrow \mathbb{R}$ funcions mesurables tals que $X = Y$ μ-q.p.t. Aleshores si la integral de X existeix, la integral de Y també existeix i coincideixen $\int_\Omega X d\mu = \int_\Omega Y d\mu$. En particular, si X és integrable Y també ho és.*

Demostració. Una prova detallada d'aquest resultat la podem trobar al llibre [1, Proposició 2.2.16]. □

[8]sempre que l'expressió de la dreta 2.2.1 no sigui $\infty - \infty$; si ho és diem que la integral no existeix

Observació 2.2.11. Recordem que $X = X^- - X^+$ i $|X| = X^- + X^+$, la condició $\int_\Omega |X| d\mu < \infty$ implica que la funció X és finita a menys d'un conjunt de mesura zero. Pel que, de la Proposició 2.2.10 es desprèn que per parlar de la integral d'una funció mesurable arbitrària X només cal que estigui definida μ-q.p.t [9]. Per tant, tenim que en la teoria de la probabilitat si dues variables coincideixen quasi per tot respecte μ, es poden identificar.

Observació 2.2.12. Les funcions $X : \Omega \longrightarrow \mathbb{C}$, les podem considerar definint les seves integrals separant la part real de la part imaginària. Recordem que amb a notació binomial $a + ib$, amb $a, b \in \mathbb{R}$, totes les operacions lineals sobre $\mathbb{R}$ s'estenen de manera natural al cas complex fent-les actuar formalment sobre l'expressió $a + ib$. La integral de funcions integrables és una operació lineal, per tant, la definició natural d'integral per funcions complexes considerant un espai de mesura $(\Omega, \mathcal{A}, \mu)$ és la següent: Definim la integral de X respecte μ com la integral,

$$\int_\Omega X d\mu = \int_\Omega ReX d\mu + i \int_\Omega ImX d\mu.$$

Diem aleshores que X és integrable [10].

2.2.1 Teoremes de convergència

Agrupem en aquest apartat tres propietats destacades de les integrals de funcions mesurables arbitràries. Els següents resultats són coneguts com els Teoremes de Convergència de Lebesgue.

Teorema 2.2.13 (Teorema de convergència monòtona). *Sigui $\{X_n, n \geq 1\}$ una successió creixent q.s de variables aleatòries integrables amb límit X. Suposem que $\sup_n \int_\Omega X_n d\mu < \infty$. Aleshores X és integrable i,*

$$\int_\Omega X_n d\mu \uparrow \int_\Omega X d\mu.$$

Anàlogament, si $\{X_n, n \geq 1\}$ una successió decreixent q.s de variables aleatòries integrables amb límit X que compleix $\inf_n \int_\Omega X_n d\mu > -\infty$, es té aleshores que X és integrable i,

$$\int_\Omega X_n d\mu \downarrow \int_\Omega X d\mu.$$

Demostració. Per una prova detallada veure el llibre [1, Teorema 2.3.1]. Donem una indicació de com provar el cas creixent (el cas decreixant es redueix al mateix mitjançant un canvi de signe). En el cas creixent nomes cal substituir X_n i X per $X_n - X_1$ i $X - X_1$ respectivament i aplicar les propietats de monotonia de la integral de funcions mesurables positives. □

Teorema 2.2.14 (Lema de Fatou). *Per a tota successió $\{X_n, n \geq 1\}$ de variables aleatòries i X funció mesurable arbitrària tals que compleixen*

(i) $X_n \geq X$ per a tot $n \geq 1$, on X és una variable aleatòria.

(ii) X, X_n, $\liminf X_n$ són integrables

---

[9] és a dir, pot haver-hi un conjunt de mesura zero N tal que X no estigui definida sobre N

[10] sempre i quan ReX i ImX siguin funcions reals integrables. En cas contrari diem que X no existeix o que X no és integrable

Aleshores,

$$\int_\Omega (\liminf X_n) d\mu \le \liminf \int_\Omega X_n d\mu.$$

Demostració. La successió $\inf_{m\ge n} X_m$ creix cap a $\liminf X_n$. Utilitzant la propietat *(iv)* de convergència monòtona de la Proposició 2.2.1 obtenim,

$$\begin{aligned}\liminf \int_\Omega X_n d\mu &= \lim_n \inf_{m\ge n} \int_\Omega X_m d\mu \ge \lim_n \int_\Omega (\inf_{m\ge n} X_m) d\mu \\ &= \int_\Omega \lim_n (\inf_{m\ge n} X_m) d\mu = \int_\Omega (\liminf X_n) d\mu.\end{aligned}$$

□

Proposició 2.2.15 (Lema de Fatou invertit). *Per a tota successió $\{X_n, n \ge 1\}$ de variables aleatòries i X funció mesurable arbitrària tals que compleixen*

(i) $X_n \ge X$ *per a tot* $n \ge 1$, *on* X *és una variable aleatòria.*

(ii) X, X_n, $\limsup X_n$ *són integrables*

Aleshores,

$$\limsup \int_\Omega X_n d\mu \le \int_\Omega (\limsup X_n) d\mu.$$

Teorema 2.2.16 (Teorema de convergència dominada). *Sigui $\{X_n, n \ge 1\}$ una successió de variables aleatòries que compleixen les següents propietats,*

(i) $\lim_n X_n = X$.

(ii) $|X_n| \le Y$, *per a tot* $n \ge 1$, *on* Y *és integrable.*

Aleshores X *és integrable i es compleix,*

$$\int_\Omega X_n d\mu = \lim \int_\Omega X_n d\mu.$$

Demostració. La relació $|X_n| \le Y$ implica que X és integrable i aplicant el Lema de Fatou (Lema 2.2.14) i el Lema de Fatou invertit (Lema 2.2.15) tindrem

$$\begin{aligned}\int_\Omega X_n d\mu = \int_\Omega (\liminf X_n) d\mu &\le \liminf \int_\Omega X_n d\mu \\ \le \limsup \int_\Omega X_n d\mu &\le \int_\Omega (\limsup X_n) d\mu = \lim \int_\Omega X_n d\mu.\end{aligned}$$

□

Observació 2.2.17. El Teorema 2.2.16 és una altra propietat que es pot enunciar directament per a funcions complexes, només interpretant el valor absolut com a mòdul. Si una funció Y domina el mòdul de X, també domina les seves parts real i imaginària ($max\{|ReX|, |ImX| \le |X|\}$) i per tant el Teorema 2.2.16 s'aplica a cada una d'elles. La fòrmula dels límits s'obté sumant parts reals i parts imaginàries.

2.2.2 Càlcul d'integrals

En aquest apartat veurem els teoremes de la mesura imatge i de Radon-Nikodýn que ens faciliten el càlcul efectiu d'integrals, també introduïrem el concepte de densitat i de variables absolutament contínues.

Considerem un espai de mesura $(\Omega, \mathcal{A}, \mu)$, un espai mesurable $(E, \mathcal{E})$ i una aplicació mesurable $X : \Omega \longrightarrow E$. L'aplicació X indueix una mesura en l'espai $(E, \mathcal{E})$ definida de la manera següent:

$$(\mu \circ X^{-1})(B) = \mu(X^{-1}(B)), \quad B \in \mathcal{E}.$$

Definició 2.2.18. *Anomenem mesura imatge de μ per X a l'aplicació $(\mu \circ X^{-1})$ induïda per X en l'espai mesurable $(E, \mathcal{E})$ i la denotarem μ_X.*

Abans d'introduïr el teorema de la mesura imatge fem incís en la següent notació, tenint en compte l'espai de mesura i l'aplicació mesurable anteriors, introduïm el símbol següent:

$$\int_A X d\mu := \int_\Omega X \cdot \mathbb{1}_A d\mu, \quad \text{per qualsevol } A \in \mathcal{A}.$$

Notem que si la integral X existeix, aleshores la integral de $X \cdot \mathbb{1}_A$ existeix, atès que $(X \cdot \mathbb{1}_A)^+ \leq X^+$ i $(X \cdot \mathbb{1}_A)^- \leq X^-$, i el símbol $\int_A X d\mu$ està ben definit. En particular si X és integrable tenim que $X \cdot \mathbb{1}_A$ també ho és.

El teorema de la mesura imatge és, de fet, un teorema de canvi de variable molt general.

Teorema 2.2.19 (Teorema de la mesura imatge). *Sigui $f : E \longrightarrow \mathbb{R}$ una funció mesurable. Llavors f és integrable en l'espai de mesura $(E, \mathcal{E}, \mu_X)$ si i només si $f \circ X$ és integrable en $(\Omega, \mathcal{A}, \mu)$. En aquest cas tindrem*

$$\int_\Omega (f \circ X) d\mu = \int_E f d\mu_X.$$

Demostració. Per demostrar aquest teorema seguirem el següent esquema: suposarem que la funció f és un indicador, una funció elemental positiva, una funció mesurable positiva i una funció mesurable arbitrària [11] .

(i) Indicadors: $f = \mathbb{1}_B$, $B \in \mathcal{E}$,

$$\int_\Omega f \circ X d\mu = \int_\Omega \mathbb{1}_B \circ X d\mu = \int_\Omega \mathbb{1}_{X^{-1}(B)} d\mu = \mu(X^{-1}(B)) = \mu_X(B) = \int_E \mathbb{1}_B d\mu_X.$$

(ii) Elementals positives: $f = \sum_{i=1}^n a_i \cdot \mathbb{1}_{B_i}$, $B_i \in \mathcal{E}$ per $i \in \{1, ..., n\}$,

$$\int_\Omega f \circ X d\mu = \sum_{i=1}^n a_i \int_\Omega \mathbb{1}_{B_i} \circ X d\mu = \sum_{i=1}^n a_i \int_E \mathbb{1}_{B_i} d\mu_X = \int_E f d\mu_X.$$

On a la segona igualtat hem usat el cas *(i)*.

(iii) Mesurables positives: Sigui $\{f_n\}_{n\in\mathbb{N}}$ una successió creixent de funcions elementals positives amb $\lim_{n\to\infty} f_n = f$. Observem que $\{f_n \circ X\}_{n\in\mathbb{N}}$ és una successió creixent de funcions elementals positives sobre $(\Omega, \mathcal{A})$ que convergeixen a $f \circ X$. Per tant,

$$\int_\Omega f \circ X d\mu = \lim_{n\to\infty} \int_\Omega f_n \circ X d\mu = \lim_{n\to\infty} \int_\Omega f_n d\mu_X = \int_E f d\mu_X.$$

[11] aquest és un procediment habitual per provar enunciats que involucren integrals

On al segon igual hem usat el cas *(ii)*, i a les altres igualtats els teoremes de convergència.

(iv) Mesurables arbitràries: Pel cas anterior tenim:

$$\int_\Omega (f\circ X)^+ d\mu = \int_\Omega f^+ \circ X d\mu = \int_E f^+ d\mu_X,$$
$$\int_\Omega (f\circ X)^- d\mu = \int_\Omega f^- \circ X d\mu = \int_E f^- d\mu_X.$$

Per tant, la integral de f respecte μ_X és integrable si i només si la integral de $f\circ X$ respecte μ és integrable. En cas afirmatiu,

$$\int_\Omega f\circ X d\mu = \int_\Omega (f\circ X)^+ d\mu - \int_\Omega (f\circ X)^- d\mu = \int_E f^+ d\mu_X - \int_E f^- d\mu_X = \int_E f d\mu_X.$$

□

Definició 2.2.20. *Sigui* $(\Omega, \mathcal{A}, \mu)$ *un espai de mesura i* $f : \Omega \longrightarrow [0, +\infty)$ *una funció mesurable. Definim sobre* $\mathcal{A}$ *la mesura*

$$\nu(A) := \int_A f d\mu,$$

amb $A \in \mathcal{A}$*, i l'anomenem densitat de* ν *respecte* μ *i l'escrivim com* $f = \frac{d\nu}{d\mu}$ [12].

Observacions 2.2.21.

(i) Notem que si μ és una mesura σ-finita i $f < +\infty$ μ-q.p.t, llavors tenim que ν és σ-finita.

(ii) Si f és integrable, llavors ν és finita.

(iii) Si $\int_\Omega f d\mu = 1$, llavors ν és una probabilitat.

Notem que les densitats són funcions mesurables que compleixen $\int_{-\infty}^{+\infty} f(x)dx = 1$, és per això que també s'anomenen densitats de probabilitat.

Definició 2.2.22. *Siguin* μ *i* ν *dues mesures sobre un espai mesurable* $(\Omega, \mathcal{A})$*. Diem que* ν *és absolutament contínua respecte* μ *si i només si es compleix que per tot* $A \in \mathcal{A}$*,* $\mu(A) = 0$ *implica* $\nu(A) = 0$*. Ho denotem per* $\nu \ll \mu$*.*

Tota funció de distribució absolutament contínua és contínua. El recíproc no és cert.

Teorema 2.2.23 (Teorema de Radon-Nikodým). *Siguin* μ *i* ν *dues mesures sobre un espai mesurable* $(E, \mathcal{E})$ *amb* μ σ*-finita. Aleshores són equivalents els següents enunciats:*

(i) ν *és absolutament contínua respecte* μ*:* $\nu \ll \mu$ *.*

(ii) Existeix una funció $f : E \longrightarrow [0, +\infty)$ *tal que* f *és una densitat de* ν *respecte* μ*:* $f = \frac{d\nu}{d\mu}$*.*

Demostració. La implicació *(ii)* $\Rightarrow$ *(i)* és directe i no cal utilitzar que μ és σ-finita:

$$\text{Si } \mu(A) = 0 \text{ aleshores } \nu(A) = \int_A f \, \mathrm{d}\mu = \int_\Omega 0 \, \mathrm{d}\mu = 0, \text{ perquè } f \cdot \mathbb{1}_A = 0 \ \sigma\text{-q-p.t.}$$

L'altre implicació és molt llarga i no l'escriurem aquí, es pot consultar el llibre [3, Section 32]. □

[12] també és diu que f és la derivada de Radon-Nikodým de ν respecte μ

2.3 Esperança matemàtica

L'objectiu d'aquest apartat és estudiar l'esperança matemàtica d'una variable aleatòria com a cas particular de la integració de funcions mesurables reals.

Considerem un espai de probabilitat $(\Omega, \mathcal{A}, P)$. Recordem que tota variable aleatòria $X : \Omega \longrightarrow \mathbb{R}$ és una funció mesurable que indueix una probabilitat sobre la recta, $P_X = (P \circ X^{-1})$, que s'anomena llei o distribució de probabilitat de la variable X [13], és a dir

$$P_X(B) := P(X^{-1}(B)), \quad B \in \mathcal{B}.$$

Anomenarem funció de distribució de la variable X a la funció de distribució de la seva llei, és a dir:

$$F_X(x) := P_X((-\infty, x]) = P(\{X \leq x\}), \quad x \in \mathbb{R}.$$

Definició 2.3.1. *Sigui $(\Omega, \mathcal{A}, P)$ un espai de probabilitat. Si X és una variable aleatòria integrable respecte P, o bé X és una variable positiva, definim l'esperança matemàtica de la variable X respecte P com la integral de X respecte de la mesura P. És a dir,*

$$E(X) := \int_\Omega X dP.$$

Gràcies al Teorema de la mesura imatge (Teorema 2.2.19), l'esperança de X es pot calcular mitjançant una integral sobre $\mathbb{R}$ si coneixem la seva llei:

$$E(X) := \int_{\mathbb{R}} x P_X(dx), \text{ on } P_X = (P \circ X^{-1}).$$

Més en general, si $f : \mathbb{R} \longrightarrow \mathbb{R}$ és una funció mesurable positiva o amb $f \circ X$ integrable respecte P, aleshores tenim

$$E[f(X)] = E[f \circ X] = \int_\Omega f(X) dP = \int_{\mathbb{R}} f(x) P_X(dx).$$

Veiem com queda l'expressió segons si X és una variable discreta o absolutament contínua:

(i) Si X és una variable discreta alehores tenim per un I finit o numerable: $P_X = \sum_{i \in I} \alpha_i \delta_i$ amb $x_i \in \mathbb{R}$, $\alpha_i \in [0,1]$ i $\sum_{i \in I} \alpha_i = 1$. L'esperança matemàtica en aquest cas ens queda:

$$E[f(X)] = \sum_{i \in I} f(x_i)\alpha_i.$$

(ii) Si X és una variable absolutament contínua amb densitat $\varphi(x)$ l'expressió de l'esperança matemàtica ens queda:

$$E[f(X)] = \int_{\mathbb{R}} f(x)\varphi(x) dx.$$

Passem a definir diferents conceptes associats a la noció d'esperança: moments, variància, desviació típica i covariància.

Definició 2.3.2. *Per cada natural $p \geq 1$ podem definir, si existeix, $E(X^p)$ i s'anomena moment d'ordre p de X. Generalment el denotarem m_p.*

[13] P_X és la mesura imatge de P per X

S'anomena variància d'una variable aleatòria X, i es denota $\mathrm{Var}(X)$, al moment d'ordre 2 de $X - E(X)$, si aquesta expressió té sentit, és a dir si tenim $E(X^2) < \infty$, aleshores:

$$\mathrm{Var}(X) := E[(X - E(X))^2] = E(X^2) - (E(X))^2.$$

Gràcies a la Desigualtat de Schwarz [14] tenim que $E(X^2) < \infty$ implica que $E(|X|) < \infty$, ja que es compleix $E(|X|) \leq \sqrt{E(X^2)}$. Notem que la variància sempre és positiva. Anomenem desviació típica de X, $\sigma(X)$, a l'arrel quadrada positiva de la variància de X:

$$\sigma(X) := \sqrt{\mathrm{Var}(X)}.$$

És també habitual escriure la variància com $\sigma^2(X)$. Considerem també com a cas particular el moment d'ordre 1 de X, m_1, que anomenem valor mig, ja que si interpretem μ com una distribució de masses sobre la recta, el valor mig representa el *centre*. Concretament, la variància mesura el grau de disperció de la mesura μ respecte el valor mig m_1.

Observació 2.3.3. Els conceptes d'esperança, variància i desviació típica, i, en general, el moment d'ordre p d'una variable i qualsevol quantitat que en sigui funció, només depèn de la llei de la variable. És a dir, podem parlar també de esperança, variància i moment d'ordre p d'una probabilitat en $(\mathbb{R}, \mathcal{B})$. Notem que no és el mateix variable aleatòria (funció) i llei de la variable aleatòria (probabilitat): dues variables definides de manera diferent, fins i tot en espais de probabilitat diversos, poden tenir la mateixa llei.

Siguin X, Y dues variables aleatòries de quadrat integrable. Es defineix la seva covariància com,

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

Si dues variables tenen covariància nul·la es diu que estan *incorrelacionades.*

Definim el coeficient de correlació per dues variables X, Y de quadrat integrable i de variància no nul·la de la següent manera,

$$\rho_{XY} = \frac{cov(X, Y)}{\sigma(X)\sigma(Y)}.$$

Com a conseqüència de la desigualtat de Schwarz tenim que $|\rho_{XY}| \leq 1$. Es defineix també, per un vector aleatori $X = (X_1, ..., X_n)$ amb components de quadrat integrable, la matriu de variàncies i covariàncies com $\Lambda = (cov(X_i, X_j))_{1 \leq i,j \leq n}$ [15]. La matriu Λ és simètrica i definida no negativa ja que per a tot $t \in \mathbb{R}^n$ es té,

$$t^*\Lambda t = \left[t^* E[(X - E(X))(X - E(X))^* t\right] = E\left[(t^*(X - E(X)))^2\right] \geq 0.$$

[14] es pot consultar a l'annex

[15] també es pot escriure com $\Lambda = E[(X - E(X))(X - E(X))^*]$ amb el conveni que els vectors s'escriuen com a matriu d'una columna

Capítol 3

Producte d'espais mesurables, independència i distribucions condicionades

3.1 Producte d'espais mesurables

En aquest apartat introduïm els productes d'espais mesurables. Sigui $(\Omega_i, \mathcal{A}_i)_{i\in I}$ una família arbitrària d'espais mesurables. Sigui $\Omega := \times_{i\in I}\Omega_i$ el producte cartesià de la familia $\{\Omega_i\}_{i\in I}$ i per tot $i \in I$ numerable, considerem les aplicacions projecció $\pi_i : \Omega \longrightarrow \Omega_i$ de Ω.

Definició 3.1.1. *S'anomena σ-àlgebra producte de $\{\mathcal{A}_i\}_{i\in I}$ la σ-àlgebra de $\mathcal{P}(\Omega)$ generada per $\{\pi_i\}_{i\in I}$. Es denota $\mathcal{A} = \otimes_{i\in I}\mathcal{A}_i$. Aleshores $(\Omega, \mathcal{A})$ s'anomena l'espai mesurable producte* [1] *de $\{(\Omega_i, \mathcal{A}_i)\}_{i\in I}$.*

Equivalentment, tenim que la σ-àlgebra producte es pot expressar com $\otimes_{i\in I}\mathcal{A}_i = \{\pi_i^{-1}(A_i);\ A_i \in \mathcal{A}_i,\ i \in I\}$. Un conjunt de la forma $\pi_i^{-1}(A_i)$ s'anomena un cilindre de base A_i.

Observació 3.1.2. En la situació anterior si $(E, \mathcal{E})$ és un altre espai mesurable, aleshores la funció $X : E \longrightarrow \Omega$ és mesurable si i només si per tot $i \in I$, es té $\pi_i \circ X :\longrightarrow \Omega_i$ és mesurable. Aquest fet és conseqüència directa de la Proposició 2.1.2.

La definició anterior és molt general, però només ens interessarà el cas en què la familia d'espais mesurables és finita. En aquesta situació la σ-àlgebra producte es pot definir d'una altra manera equivalent:

Definició 3.1.3. *Sigui $(\Omega, \mathcal{A})$ l'espai mesurable producte de la familia finita $(\Omega_i, \mathcal{A}_i)_{i=1}^n$. S'anomena rectangle mesurable tot subconjunt de Ω de la forma $A_1 \times \cdots \times A_n$, amb $A_i \in \mathcal{A}_i$.*

La σ-àlgebra generada pels rectangles coincideix amb la generada pels cilindres, ja que tot cilindre és un rectangle mesurable, $\pi_i^{-1}(A_i) = \Omega_1 \times \cdots \times \Omega_{i-1} \times A_i \times \Omega_{i+1} \times \cdots \times \Omega_n$, i tot rectangle és intersecció finita de cilindres, $A_1 \times \cdots \times A_n = \cap_{i=1}^n \pi_i^{-1}(A_i)$.

Proposició 3.1.4. *Siguin $(\Omega_1, \mathcal{A}_1)$ i $(\Omega_2, \mathcal{A}_2)$ dos espais mesurables i $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2)$ el seu espai mesurable producte. Sigui $(E, \mathcal{E})$ un espai mesurable i X una funció mesurable tal que $X : \Omega_1 \times \Omega_2 \longrightarrow E$. Aleshores per cada $\omega_1 \in \Omega_1$, la funció parcial següent és mesurable:*

[1] observem l'analogia amb la definició de topologia producte, que és la mínima topologia que fa contínues les projeccions

$$\begin{aligned} X(\omega_1, \cdot) \colon \Omega_2 &\longrightarrow E \\ \omega_2 &\longmapsto X(\omega_1, \omega_2). \end{aligned}$$

Demostració. Gràcies a la Observació 3.1.2, la funció $i_{\omega_1} \colon \Omega_2 \longrightarrow \Omega_1 \times \Omega_2$ amb $i_{\omega_1}(\omega_2) = (\omega_1, \omega_2)$, és mesurable perquè composada amb cadascuna de les projeccions dóna lloc a una funció mesurable, en particular obtenim o bé una funció constant, o bé la identitat. Pel que, $X(\omega_1, \cdot)$ és la composició de les funions mesurables $(X \circ i_{\omega_1})$, i per tant, és mesurable. □

3.1.1 Producte finit d'espais de mesura

Volem construir mesures en un espai mesurable producte (finit), ho farem per dos espais i aplicant iterativament el procediment es farà la construcció sobre un producte finit de n factors. L'eina bàsica és el concepte que definim tot seguit.

Definició 3.1.5. *Una mesura de transició d'un espai mesurable $(\Omega_1, \mathcal{A}_1)$ en un altre espai mesurable $(\Omega_2, \mathcal{A}_2)$ és una aplicació $\tau : \Omega_1 \times \mathcal{A}_2 \longrightarrow [0, \infty)$ tal que:*

(i) Per a tot $\omega_1 \in \Omega_1$ es compleix $\tau(\omega_1, \cdot) : \mathcal{A}_2 \longrightarrow [0, \infty)$ és una mesura en $(\Omega_2, \mathcal{A}_2)$.

(ii) Per a tot $A_2 \in \mathcal{A}_2$ es compleix $\tau(\cdot, A_2) : \Omega_1 \longrightarrow [0, \infty)$ és mesurable.

Quan la imatge de la funció τ és [0,1], diem que és una probabilitat de transició i aquestes modelen la següent situació: Si Ω_1, Ω_2 són els conjunts de resultats de dos experiments aleatoris, $\tau(\omega_1, A_2)$ respresenta la probabilitat que es produeixi A_2 en el segon experiment sabent que el resultat del primer ha sigut ω_1. Un cas particular important és aquell en què $\tau(\cdot, A_2)$ és una constant $\tau(A_2)$, per cada $A_2 \in \mathcal{A}_2$. Aleshores,

$$\begin{aligned} \tau \colon \mathcal{A}_2 &\longrightarrow [0, 1] \\ A_2 &\longmapsto \tau(A_2), \end{aligned}$$

és una probabilitat en $(\Omega_2, \mathcal{A}_2)$. Això vol dir que el segon experiment no està influït pel primer resultat. Ho precisarem més endevant amb el concepte d'independència.

Definició 3.1.6. *Una família de mesures $\{\mu_i\}_{i \in I}$ en un espai mesurable $(\Omega, \mathcal{A})$ és uniformement σ-finita si i només si existeix una descomposicio $\Omega = \cup_{n=1}^{\infty} B_n$ amb $B_n \in \mathcal{A}$ i constants $K_n \in \mathbb{R}$ tals que per tot $n \in \mathbb{N}$ i per tot $i \in I$ es compleix $\mu_i(B_n) \leq K_n$.*

Teorema 3.1.7 (Teorema de la mesura producte). *Sigui $(\Omega_1, \mathcal{A}_1, P_1)$ un espai de probabilitat amb P_1 σ-finita, $(\Omega_2, \mathcal{A}_2)$ un espai mesurable i τ una probabilitat de transició de $(\Omega_1, \mathcal{A}_1)$ a $(\Omega_2, \mathcal{A}_2)$, és a dir $\tau : (\Omega_1, \mathcal{A}_2) \longrightarrow [0, 1]$. Aleshores, usant la notació $(\Omega, \mathcal{A}) := (\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2)$ es compleix,*

(i) Existeix una única probabilitat Q en l'espai producte $(\Omega, \mathcal{A})$ tal que,

$$Q(A_1 \times A_2) = \int_{A_1} \tau(\omega_1, A_2) P_1(d\omega_1), \quad \text{per a tots} \quad A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2. \tag{3.1.1}$$

(ii) Si $X : \Omega \longrightarrow [0, \infty)$ és una funció mesurable, es compleix:

$$\int_{\Omega} X dQ = \int_{\Omega_1} \Big(\int_{\Omega_2} X(\omega_1, \omega_2) \tau(\omega_1, d\omega_2) \Big) P_1(d\omega_1). \tag{3.1.2}$$

A més, si $X : \Omega \longrightarrow \mathbb{R}$ és integrable respecte Q val també la fòrmula (3.1.2) en el sentit de que la integral $\int_{\Omega_2} X(\omega_1, \omega_2)\tau(\omega_1, d\omega_2)$ existeix per a tot ω_1 a menys d'un conjunt de probabilitat P_1 zero i defineix una funció de ω_1 integrable respecte a P_1.

Demostració.

(i) Tenint en compte els resultats sobre extensió de mesures, l'existència i la unicitat quedaran provades si veiem que Q (definida per la fòrmula (3.1.1)) és σ-additiva en la semiàlgebra $\mathcal{S}$ dels rectangles mesurables, és a dir en $\mathcal{A}_1 \otimes \mathcal{A}_2$. Suposem que $A_1 \times A_2 = \cup_{i=1}^{\infty}[A_1^i \times A_2^i]$ on els $A_1^i \times A_2^i$ són disjunts dos a dos, amb $A_1^i \in \mathcal{A}_1$ i $A_2^i \in \mathcal{A}_2$ per tot i. Llavors podem escriure, usant la definició de τ i el Teorema de la convergència monòtona (Teorema 2.2.13),

$$\begin{aligned}
\sum_{i=1}^{\infty} Q(A_1^i \times A_2^i) &= \sum_{i=1}^{\infty} \int_{A_1^i} \tau(\omega_1, A_2^i) P_1(d\omega_1) = \int_{\Omega_1} \Big(\sum_{i=1}^{\infty} \mathbb{1}_{A_1^i}(\omega_1) \cdot \tau(\omega_1, A_2^i) \Big) P_1(d\omega_1) \\
&= \int_{\Omega_1} \Big(\sum_{i=1}^{\infty} \int_{\Omega_2} \mathbb{1}_{A_1^i}(\omega_1) \cdot \mathbb{1}_{A_2^i}(\omega_2) \cdot \tau(\omega_1, d\omega_2) \Big) P_1(d\omega_1) \\
&= \int_{\Omega_1} \Big(\int_{\Omega_2} \Big(\sum_{i=1}^{\infty} \mathbb{1}_{A_1^i}(\omega_1) \cdot \mathbb{1}_{A_2^i}(\omega_2) \Big) \tau(\omega_1, d\omega_2) \Big) P_1(d\omega_1) \\
&= \int_{\Omega_1} \Big(\int_{\Omega_2} \mathbb{1}_{A_1 \times A_2}(\omega_1, \omega_2) \cdot \tau(\omega_1, d\omega_2) \Big) P_1(d\omega_1) \\
&= \int_{A_1} \tau(\omega_1, A_2) P_1(d\omega_1) = Q(A_1 \times A_2).
\end{aligned}$$

(ii) Per provar aquesta segona part, seguirem el camí habitual: primer ho veurem per indicadors, per funcions elementals positives, per funcions mesurables positives i, finalment, per funcions mesurables arbitràries:

(a) Indicadors: $X(\omega_1, \omega_2) = \mathbb{1}_A(\omega_1, \omega_2)$. Definim la familia $\mathcal{C}$ com els $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$, tals que compleixen:

- $\omega_1 \longmapsto \int_{\Omega_2} \mathbb{1}_A(\omega_1, \omega_2)\tau(\omega_1, d\omega_2)$ és mesurable.
- $Q(A) = \int_{\Omega_1} \left(\int_{\Omega_2} \mathbb{1}_A(\omega_1, \omega_2)\tau(\omega_1, d\omega_2) \right) P_1(d\omega_1)$.

Volem veure que $\mathcal{C} = \mathcal{A}_1 \otimes \mathcal{A}_2$. Aquesta igualtat surt de les consideracions següents:

I. $\mathcal{C}$ conté la semiàlgebra $\mathcal{S}$ dels rectangles mesurables.

II. $\mathcal{C}$ conté l'àlgebra $a(\mathcal{S})$.

III. $\mathcal{C}$ és classe monòtona.

Aleshores tenim que $\mathcal{A}_1 \otimes \mathcal{A}_2 = \sigma(\mathcal{S}) \subseteq \mathcal{C}$, conseqüència dels punts anteriors i del Teorema de les classes monòtones (Teorema 1.1.10).

(b) Funcions elementals positives. Caldria aplicar que la suma de funcions mesurables és mesurable i la integral de la suma és la suma d'integrals, tenint en compte que totes les funcions involucrades són positives.

(c) Ho veiem ara per funcions mesurables positives: $X = \lim_{n\to\infty} X_n$, on $\{X_n\}_{n\in\mathbb{N}}$ és una successió creixent de funcions elementals positives. Per definició,

$$\int_{\Omega_2} X(\omega_1, \omega_2)\tau(\omega_1, d\omega_2) = \lim_{n\to\infty} \int_{\Omega_2} X_n(\omega_1, \omega_2)\tau(\omega_1, d\omega_2). \tag{3.1.3}$$

La funció $\omega_1 \longmapsto \int_{\Omega_2} X(\omega_1,\omega_2)\tau(\omega_1,d\omega_2)$ és mesurable pel fet que ha de ser límit de mesurables.

$$\begin{aligned}\int_{\Omega_1\times\Omega_2} X dQ &= \lim_{n\to\infty}\int_{\Omega_1\times\Omega_2} X_n dQ\\ &= \lim_{n\to\infty}\int_{\Omega_1}\Big(\int_{\Omega_2} X_n(\omega_1,\omega_2)\tau(\omega_1,d\omega_2)\Big)P_1(d\omega_1)\\ &= \int_{\Omega_1}\Big(\int_{\Omega_2} X(\omega_1,\omega_2)\tau(\omega_1,d\omega_2)\Big)P_1(d\omega_1),\end{aligned}$$

on la segona igualtat és per (b) i la tercera igualtat és resultat d'aplicar el Teorema de la convergència monòtona (Teorema 2.2.13) i del fet que la successió (3.1.3) d'integrals és creixent i positiva.

(d) Finalment ho veiem per funcions mesurables arbitràries: Suposem que X^- és integrable, el resultat de (c) aplicat a X^- implica:

$$P_1\Big(\Big\{\omega_1\in\Omega_1 : \int_{\Omega_2} X^-(\omega_1,\omega_2)\tau(\omega_1,d\omega_2) = +\infty\Big\}\Big) = 0$$

i, per tant, la funció:

$$\begin{aligned}\omega_1 \longmapsto &\int_{\Omega_2} X(\omega_1,\omega_2)\tau(\omega_1,d\omega_2)\\ &= \int_{\Omega_2} X^+(\omega_1,\omega_2)\tau(\omega_1,d\omega_2) - \int_{\Omega_2} X^-(\omega_1,\omega_2)\tau(\omega_1,d\omega_2),\end{aligned}$$

està ben definida P_1-quasi segurament per tot $\omega_1\in\Omega_1$. Aleshores usant (c) obtenim,

$$\begin{aligned}\int_{\Omega_1\times\Omega_2} X dQ &= \int_{\Omega_1\times\Omega_2} X^+ dQ - \int_{\Omega_1\times\Omega_2} X^- dQ\\ &= \int_{\Omega_1}\Big(\int_{\Omega_2} X(\omega_1,\omega_2)\tau(\omega_1,d\omega_2)\Big)P_1(d\omega_1).\end{aligned}$$

□

Observació 3.1.8. En el cas particular en què la mesura de transició no depèn de ω_1 [2], la fòrmula (3.1.2) és coneguda com el Teorema de Fubini:

$$\int_{\Omega_1\times\Omega_2} X dQ = \int_{\Omega_1}\int_{\Omega_2} X(\omega_1,\omega_2)\tau(d\omega_2)P_1(d\omega_1). \tag{3.1.4}$$

El clàssic Teorema de Fubini en $\mathbb{R}^2$, on P_1 i τ coincideixen amb la mesura de Lebesgue és un cas particular.

3.2 Independència

El concepte d'independència ocupa un lloc central en la teoria de la probabilitat. És característic dels espais de probabilitat, és a dir, no té extensió a mesures generals. En aquest apartat estendrem la noció d'independència a col·leccions de σ-àlgebres i de variables aleatòries; establirem també diverses caracteritzacions de la independència.

[2] τ és simplement una mesura σ-finita en $(\Omega_2,\mathcal{A}_2)$

Considerem per tota la secció l'espai de probabilitat $(\Omega, \mathcal{A}, P)$. Recordem que dos esdeveniments $A, B \in \mathcal{A}$ són independents si i només si $P(A \cap B) = P(A)P(B)$. Una família arbitrària d'esdeveniments $\{A_i\}_{i\in I}$ és independent si i només si per tot subconjunt finit $J \subseteq I$, es té $P(\cap_{j\in J} A_j) = \prod_{j\in J} P(A_j)$ [3].

Definició 3.2.1. *Sigui $\{\mathcal{C}_i\}_{i\in I}$, amb $\mathcal{C}_i \in \mathcal{A}$ una família de col·leccions d'esdeveniments. Diem que $\{\mathcal{C}_i\}_{i\in I}$ és independent si i només si tota la família d'esdeveniments $\{A_i\}_{i\in I}$, amb $A_i \in \mathcal{C}_i$ per a tota $i \in I$, és independent.*

Definició 3.2.2. *Sigui $\{X_i : \Omega \longrightarrow \mathbb{R}\}_{i\in I}$ una família de variables aleatòries. Diem que $\{X_i\}_{i\in I}$ és independent si i només si la col·lecció de σ-àlgebres $\{X_i^{-1}(\mathcal{B}(\mathbb{R}))\}_{i\in I}$ és independent.*

Tenint en compte les definicions anteriors, podem considerar el concepte d'independència de variables aleatòries en termes d'independència d'esdeveniments: *$\{X_i\}_{i\in I}$ és independent si i només si per a tot subconjunt finit $\{i_1, ..., i_n\} \subseteq I$, i per a tots $B_1, ..., B_n \in \mathcal{B}(\mathbb{R})$, es compleix*

$$P(\{(X_{i_1} \in B_1, ..., X_{i_n} \in B_n\}) = P(\{X_{i_1} \in B_1\} \times \cdots \times P(\{X_{i_n} \in B_n\}).$$

En general direm que una familia $\{X_i\}_{i\in I}$, $X_i : (\Omega : \mathcal{A}) \longrightarrow (E_i, \mathcal{E}_i)$ d'aplicacions mesurables és independent si ho és la col·lecció de σ-àlgebres $\{X_i^{-1}(\mathcal{E}_i)\}_{i\in I}$.

Abans de veure diferents propietats de les variables aleatòries independents fem un apunt, en les següents observacions, d'algunes característiques de les distribucions marginals, les distribucions conjuntes i les densitats de les lleis absolutament contínues.

Observació 3.2.3. Les funcions de distribució marginals és poden obtenir a partir de la funció de distribució conjunta mitjançant un pas al límit [4]:

$$\begin{aligned} F_{X_i}(x_i) &= P(\{X_i \le x_i\}) = P(\{X_1 \in \mathbb{R}, ..., X_i \le x_i, ..., X_n \in \mathbb{R}\}) \\ &= \lim_{\substack{x_j \to \infty \\ j \neq i}} P(\{X_1 \le x_1, ..., X_i \le x_i, ..., X_n \le x_n\}) = \lim_{\substack{x_j \to \infty \\ j \neq i}} F_X(x_1, ..., x_n). \end{aligned}$$

Ara bé, les funcions de distribució marginals $F_{X_1}, ..., F_{X_n}$ no determinen la distribució conjunta excepte en situacions particulars, com és el cas de la independència de les variables $X_1, ..., X_n$.

Observació 3.2.4. Si $X = (X_1, ..., X_n)$ és un vector aleatori que té llei conjunta $P_X = (P \circ X^{-1})$ absolutament contínua [5] amb densitat $f_X(x_1, ..., x_n)$ aleshores, cada variable X_i té una llei marginal absolutament contínua amb densitat:

$$f_{X_i}(x_i) = \int_{\mathbb{R}^{n-1}} f_X(x_1, ..., x_n) dx_1 ....dx_{i-1} dx_{i+1} ...dx_n.$$

En efecte, per tot borelià B de $\mathbb{R}$, si posem $\overline{B} = \mathbb{R}^{i-1} \times B \times \mathbb{R}^{n-i}$, tindrem pel Teorema de Fubini,

$$P(\{X_i \in B\}) = P(\{X \in \overline{B}\}) = \int_{\overline{B}} f_X(x) dx = \int_B f_{X_i}(x_i) dx_i.$$

El recíproc no és cert, ja que si X_1 és una variable aleatòria amb llei absolutament contínua en $\mathbb{R}$, el vector aleatori (X_1, X_1) no pot tenir densitat perquè la seva distribució està concentrada en la recta $x_1 = x_2$ de $\mathbb{R}^2$.

[3] Notem que la independència d'una família d'esdeveniments no és equivalent al fet que els seus elements siguin independents dos a dos

[4] usant la propietat de continuïtat de les mesures per successions creixents

[5] respecte la mesura de Lebesgue en $\mathbb{R}^n$

Passem a veure algunes propietats de les variables aleatòries independents:

(i) Si $\{X_i\}_{i\in I}$ són variables aleatòries independents i $g_i : \mathbb{R} \longrightarrow \mathbb{R}$ són funcions mesurables, aleshores les variables $\{g_i(X_i)\}_{i\in I}$ són també independents.

Demostració. Per a cada $i \in I$ tenim,

$$(g_i(X_i))^{-1}(\mathcal{B}(\mathbb{R})) = X_i^{-1}\left(g_i^{-1}(\mathcal{B}(\mathbb{R}))\right) \subseteq (X_i)^{-1}(\mathcal{B}(\mathbb{R})). \qquad \square$$

(ii) Les variables aleatòries $X_1, ..., X_n$ són independents si i només si la llei del vector aleatori $X = (X_1, ..., X_n)$ és el producte de les lleis marginals, és a dir:

$$P_X = P \circ (X_1, ..., X_n)^{-1} = (P \circ X_1^{-1}) \times \cdots \times (P \circ X_n^{-1}) = P_{X_1} \times \cdots \times P_{X_n}. \quad (3.2.1)$$

Demostració. Suposem primer que les variable són independents: Per a tot $B_1, ..., B_n \in \mathcal{B}(\mathbb{R})$ tindrem:

$$\begin{aligned}(P \circ X^{-1})(B_1 \times \cdots \times B_n) &= P(\{X_1 \in B_1, ..., X_n \in B_n\}) \\ &= P(\{X_1 \in B_1\}) \times \cdots \times P(\{X_n \in B_n\}) \\ &= P_{X_1}(B_1) \times \cdots \times P_{X_n}(B_n),\end{aligned}$$

la qual cosa implica (3.2.1) ja que les dues probabilitats en $\mathbb{R}^n$ queden determinades pels seus valors en els rectangles mesurables. Recíprocament si (3.2.1) és cert, per a tot conjunt finit $\{i_1, ..., i_m\} \subseteq \{1, ..., n\}$ i per tots $B_1, ..., B_m \in \mathcal{B}(\mathbb{R})$ tindrem,

$$\begin{aligned}&P(\{X_{i_1} \in B_1, ..., X_{i_m} \in B_m\}) \\ &= P(\{X_{i_1} \in B_1, ..., X_{i_m} \in B_m; X_j \in \mathbb{R}, j \in \{1, 2, ..., n\} - \{i_1, ..., i_m\}\}) \\ &= P(\{X_{i_1} \in B_1\}) \times \cdots \times P(\{X_{i_m} \in B_m\}).\end{aligned}$$

$\square$

(iii) Les variables aleatòries $X_1, ..., X_n$ són independents si i només si la funció de distribució conjunta és el producte de les funcions de distribució marginals, és a dir:

$$F_X(x_1, ..., x_n) = F_X(x_1) \times \cdots \times F_X(x_n).$$

Demostració. La funció de distribució de la probabilitat producte de P_{X_1},...,P_{X_n} val:

$$\begin{aligned}(P_{X_1} \times \cdots \times P_{X_n})((-\infty, x]) &= P_{X_1}((\infty, x_1]) \times \cdots \times P_{X_n}((-\infty, x_n]) \\ &= F_{X_1}(x_1) \times \cdots \times F_{X_n}(x_n).\end{aligned}$$

Notem que aquesta funció de distribució és igual a la de P_X si i només si $X_1, ..., X_n$ són independents per (ii). $\square$

(iv) Siguin $X_1, ..., X_n$ variables aleatòries amb lleis marginals absolutament contínues. Aleshores $X_1, ..., X_n$ són independents si i només si la llei conjunta és absolutament contínua amb densitat,

$$f_X(x_1, ..., x_n) = f_{X_1}(x_1) \times \cdots \times f_{X_n}(x_n).$$

Demostració. Una de les implicacions és directe de la Observació 3.2.4. Degut a la propietat (ii) la independència de les variables $X_1, ..., X_n$ equival a que,

$$P(X_1 \in B_1, ..., X_n \in B_n) = \left(\int_{B_1} f_{X_1}(x_1)dx_1\right) \times \cdots \times \left(\int_{B_n} f_{X_n}(x_n)dx_n\right).$$

Només cal veure que pel Teorema de Fubini, el segon membre d'aquesta igualtat val

$$\int_{B_1\times\cdots\times B_n} f_{X_1}(x_1) \times \cdots \times f_{X_n}(x_n)dx_1...dx_n.$$ □

(v) Si les variables aleatòries $X_1, ..., X_n$ prenen valors en un mateix conjunt finit o numerable $S = \{a_i\}_{i\in I} \subseteq \mathbb{R}$, tenim que són independents únicament si per a tots els $a_{i_1}, ..., a_{i_n} \in S$ es compleix:

$$P(\{X_1 = a_{i_1}, ..., X_n = a_{i_n}\}) = P(\{X_1 = a_{i_1}) \times \cdots \times P(\{X_n = a_{i_n})$$

Demostració. Una prova detallada d'aquest resultat la podem trobar a [6, Capítol 8. Teorema 8.1]. □

(vi) Suposem que $X_1, ..., X_n$ són variables aleatòries independents i integrables. Aleshores, el producte $X_1...X_n$ és integrable i l'esperança del producte és igual al producte de les esperances dels factors, és a dir

$$E(X_1 \times \cdots \times X_n) = E(X_1) \times \cdots \times E(X_n).$$

Demostració. Suposem primer que les variables són positives. Utilitzant la caracterització (3.2.1) de la independència i el Teorema de Fubini, obtenim:

$$\begin{aligned} E(X_1 \times \cdots \times X_n) &= \int_{\mathbb{R}^n} x_1 \times \cdots \times x_n P_X(dx_1, ..., dx_n) \\ &= \int_{\mathbb{R}^n} x_1 \times \cdots \times x_n \left(P_{X_1} \times \cdots \times P_{X_n}\right)(dx_1, ..., dx_n) \\ &= \left(\int_{\mathbb{R}} x_1 P_{X_1}(dx_1)\right) \times \cdots \times \left(\int_{\mathbb{R}} x_1 P_{X_n}(dx_n)\right) \\ &= E(X_1) \times \cdots \times E(X_n). \end{aligned}$$

□

(vii) Siguin $X_1, ..., X_n$ variables aleatòries independents amb moment de segon ordre finit. Aleshores,

$$\mathrm{Var}[X_1 + ... + X_n] = \mathrm{Var}[X_1] + ... + \mathrm{Var}[X_n].$$

Demostració. Una prova detallada d'aquest resultat la podem trobar al llibre [12, Capítol 7. Secció 7.4. Proposició 4.2]. □

(viii) Sigui $X = (X_1, ...X_n)$ un vector aleatori tal que les seves components són independents i de quadrat integrable. Aleshores la matriu de variàncies i covariàncies és diagonal.

Demostració. Com que les components $X = (X_1, ...X_n)$ són de quadrat integrable, la matriu $\Lambda = (cov(X_i, X_j))_{1\le i,j\le n}$ està ben definida i usant que són independents tenim,

$$cov(X_i, X_j) = E[(X_i - E(X_i))(X_j - E(X_j))] = E(X_iX_j) - E(X_i)E(X_j) = 0,$$

per a tot $i \neq j$, pel que la matriu és diagonal [6]. □

3.3 Distribucions condicionades

Donat $(\Omega, \mathcal{A}, P)$ un espai de probabilitat. Es defineix la probabilitat condicionada d'un esdeveniment $A \in \mathcal{A}$ per un altre esdeveniment $B \in \mathcal{A}$ tal que $P(B) > 0$, de la següent manera:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \tag{3.3.1}$$

La probabilitat condicionada té la següent interpretació: quan tenim la informació addicional que un cert experiment B s'ha realitzat (és a dir que $\omega \in B$), cal modificar convenienment el model inicial que haviem construït. Per exemple, B ha de passar a tenir probabilitat 1. Amb aquesta definició, notem que tindrem que els esdeveniments A, B són independents si i només si $P(A|B) = P(A)$, és a dir, la informació que s'ha produït B no aporta res de nou sobre A. En aquest apartat volem estendre la definició de probabilitat condicionada anterior per tal de poder condicionar per esdeveniments de probabilitat zero. Això ens porta buscar una definició de la distribució de probabilitat d'una variable aleatòria Y condicionada per la realització d'una variable X.

Definició 3.3.1. *Siguin $X, Y : \Omega \longrightarrow \mathbb{R}$ variables aleatòries. S'anomena llei de Y condicionada per X a tota probabilitat de transició $p(X, B)$ de $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ en $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ que compleixi, per a tot $A, B \in \mathcal{B}(\mathbb{R})$,*

$$P(\{X \in A, Y \in B\}) := \int_A p(X, B)P_X(dx) = \int_{\{X\in A\}} p(X(\omega), B)P(dx), \tag{3.3.2}$$

Fem un incís de notació: a vegades s'escriu $p(x, B) = P(\{Y \in B|X = x\})$. Observem que la notació és consistent [7] si $P(\{X = x\}) > 0$. Si prenem $A = \{x\}$ tenim per (3.3.2),

$$P(\{X = x, Y \in B\}) = \int_{\{x\}} p(y, B)P_X(dy) = p(x, B)P(\{X = x\}),$$

i doncs,

$$p(x, B) = \frac{P(\{X = x, Y \in B\})}{P(\{X = x\})}.$$

Observació 3.3.2. Com que el primer membre de (3.3.2) val $[P \circ (X, Y)^{-1}](A \times B)$, la definició anterior equival a dir que la llei $P \circ (X, Y)^{-1} = P_{(X,Y)}$ del vector aleatori (X, Y) s'obté a partir de la llei de X i la probabilitat de transició $p(x, B)$. Aleshores tenint en compte el Teorema de la mesura producte (Teorema 3.1.7), tindrem que per tota funció mesurable $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ tal que $E(|f(X, Y)|) < \infty$ [8] es compleix,

$$E(f(X, Y)) = \int_{\mathbb{R}^2} fP_{(X,Y)}(dx) = \int_{\mathbb{R}} P_X(dx)\Big[\int_{\mathbb{R}} f(x, y)P(Y \in dy|X = x)\Big].$$

Aquest resultat s'esten al cas de dos vectors aleatoris $X = (X_1, ..., X_n)$, $Y = (Y_1, ..., Y_n)$.

[6] en particular hem vist que si dues variables aleatòries X, Y són de quadrat integrable i independents estan incorrelacionades

[7] coincideix amb la probabilitat condicionada de l'esdeveniment $\{Y \in B\}$ per l'esdeveniment $\{X = x\}$

[8] f integrable respecte de la llei del vector aleatori (X, Y)

En particular hem vist que si coneixem la llei d'una variable X i la distribució d'una altra variable Y condicionada per X, podrem calcular la llei del vector aleatori (X, Y) i, en particular, la llei de la variable Y. No és immediat que existeixin sempre lleis condicionades: és cert, però no ho demostrarem. Concretament es té la proposició següent.

Proposició 3.3.3. *Donades dues variables aleatòries $X, Y : \Omega \longrightarrow \mathbb{R}$, sempre existeix una llei de Y condicionada per X. Si p és una d'elles, aleshores p' també ho és si i només si per a cada $B \in \mathcal{B}(\mathbb{R})$ es compleix $p(x, B) = p'(x, B)$ quasi segur en x respecte P_X.*

Demostració. Una prova detallada de la segona part d'aquest resultat la podem trobar al llibre [10, Proposició 7.12]. $\square$

Tot seguit veurem la unicitat de les distribucions condicionades, llevat de conjunts de probabilitat zero.

Proposició 3.3.4. *Siguin $X, Y : \Omega \longrightarrow \mathbb{R}$ variables aleatòries i p, p' dues lleis de Y condicionades per X. Sigui $g : \mathbb{R} \longrightarrow \mathbb{R}$ tal que existeix la integral de g respecte $p(x, \cdot)$, $P_X - q.p.t.x$. Aleshores, existeix la integral de g respecte $p'(x, \cdot)$, $P_X - q.p.t.x$ i*

$$\int_{\mathbb{R}} g(y)p(x, dy) = \int_{\mathbb{R}} g(y)p'(x, dy), \quad P_X - q.s.$$

Demostració. Farem una indicació de l'esquema de la demostració:

(i) Indicadors: si $g = \mathbb{1}_B$, hem de veure que $p(x, \cdot) = p'(x, \cdot)$ q.s., i això ho sabem per la Proposició 3.3.3.

(ii) Elementals: Si $g = \sum_{i=1}^{n} a_i \cdot \mathbb{1}_{B_i}$, sabem pel cas anterior que $p(x, B_i) = p'(x, B_i)$ excepte en un conjunt N_i de probabilitat zero, i per tant,

$$\sum_{i=1}^{n} a_i p(x, B_i) = \sum_{i=1}^{n} a_i p'(x, B_i),$$

excepte en el conjunt de probabilitat zero $\cup_{i=1}^{n} N_i$.

(iii) Mesurables positives: Si g és mesurable positiva, aleshores $g = \lim_{n\to\infty} g_n$, per una certa successió creixent $\{g_n\}_{n\in\mathbb{N}}$ de funcions elementals positives. Pel cas anterior,

$$\int_{\mathbb{R}} g_n(y)p(x, dy) = \int_{\mathbb{R}} g_n(y)p'(x, dy),$$

excepte sobre un conjunt N_n de probabilitat zero. Per tant, excepte sobre $\cup_{n=1}^{\infty} N_n$, que també té probabilitat zero,

$$\int_{\mathbb{R}} g(y)p(x, dy) = \lim_{n\to\infty} \int_{\mathbb{R}} g_n(y)p(x, dy) = \lim_{n\to\infty} \int_{\mathbb{R}} g_n(y)p'(x, dy) = \int_{\mathbb{R}} g(y)p'(x, dy).$$

(iv) Mesurables arbitràries: Només cal descompondre g en la part positiva i en la part negativa. Es dedueix que si les integrals respecte les probabilitats $p(x, \cdot)$ existeixen, aleshores també existeixen respecte les probabilitats $p'(x, \cdot)$, i que coincideixen $P_X - q.s.$ $\square$

L'esperança condicionada respecte un esdeveniment, una variable i una σ-àlgebra es pot consultar a l'annex.

Capítol 4

Lleis dels grans nombres

L'expressió *Llei dels grans nombres* inventada per Simeón-Denis Poisson a finals del segle XVIII, es refereix a la situació següent: Sigui $\{X_n\}_{n\in\mathbb{N}}$ una successió de variables aleatòries independents i $\{S_n\}_{n\in\mathbb{N}}$ la successió de sumes parcials de $\{X_n\}_{n\in\mathbb{N}}$ és a dir, $S_n = X_1 + ... + X_n$ amb $n \in \mathbb{N}$. Ens demanem per la convergència en probabilitat o quasi segura de la successió $\{\frac{S_n}{n}\}_{n\in\mathbb{N}}$ [1].

En aquest apartat utilitzarem alguns resultats preliminars de l'estudi de convergència de variables que es poden consultar a l'annex.

Teorema 4.0.1 (Llei feble dels grans nombres). *Sigui $\{X_n\}_{n\in\mathbb{N}}$ una successió de variables aleatòries independents i identicament distribuïdes (i.i.d). Suposem $E[X_1] = m$, $E[X_1^2] < \infty$ i $S_n = X_1 + ... + X_n$ per $n \in \mathbb{N}$. Aleshores,*

$$\frac{S_n}{n} \xrightarrow[n\to\infty]{L^2} m.$$

Demostració.

$$E\Big[|\frac{S_n}{n} - m|^2\Big] = \frac{1}{n^2}E[|S_n - nm|^2] = \frac{1}{n^2}\text{Var}[S_n] = \frac{1}{n}\text{Var}[X_1] \xrightarrow[n\to\infty]{} 0. \qquad \square$$

L'objectiu d'aquest apartat és enunciar i demostrar una Llei forta dels grans nombres:

Teorema 4.0.2 (Llei forta dels grans nombres [2]). *Sigui $\{X_n\}_{n\in\mathbb{N}}$ una successió de variables aleatòries independents i identicament distribuïdes (i.i.d) i $\{S_n\}_{n\in\mathbb{N}}$ la successió de sumes parcials de $\{X_n\}_{n\in\mathbb{N}}$: $S_n = X_1 + ... + X_n$. Aleshores,*

(i) Si es compleix $E[|X_i|] < \infty$, aleshores es té $\frac{S_n}{n} \xrightarrow[n\to\infty]{q.s} m = E[|X_i|]$.

(ii) Si $E[|X_i|] = \infty$ aleshores es té $\limsup_{n\to\infty} \frac{|S_n|}{n} = +\infty$ quasi segurament.

La demostració és llarga i utilitza un seguit de resultats que tenen interès en si mateixos, els estudiarem prèviament anomentant-los lemes.

Lema 4.0.3 (Desigualtat de Kolmogorov). *Siguin $X_1, ..., X_n$ variables aleatòries independents, de quadrat integrable i centrades. Posem $S_k = X_1 + .. + X_k$ amb $1 \leq k \leq n$. Aleshores, per a tot $\varepsilon > 0$ tenim*

$$P(\{\max_{1\leq k\leq n} |S_k| > \varepsilon\}) \leq \frac{\text{Var}(S_n)}{\varepsilon^2}.$$

[1] els resultats de converegència en probabilitat s'anomenen Lleis febles dels grans nombres i els resultats de convergència quasi segura s'anomenen Lleis fortes dels grans nombres

[2] també s'anomena Llei forta de Kolmogorov

Demostració. Notem que per $n = 1$ el resultat es redueix a la desigualtat de Txebixev. Fixem $\varepsilon > 0$ i posem $A := \{\max_{1\leq k\leq n}|S_k| > \varepsilon\}$. Definim:

$$A_1 := \{|S_1| > \varepsilon\} \quad \text{i} \quad A_k := \{\max_{1\leq j\leq k-1}|S_j| \leq \varepsilon, |S_k| > \varepsilon\}, \text{ per } 2 \leq k \leq n.$$

Els conjunts A_k són mesurables, disjunts dos a dos i es compleix $A = A_1 \cup A_2 \cup ... \cup A_n$. Aleshores,

$$\begin{aligned}
\mathrm{Var}[S_n] &= \int_\Omega S_n^2 dP \geq \int_A S_n^2 dP = \sum_{k=1}^n \int_{A_k} S_n^2 dP = \sum_{k=1}^n \int_{A_k} [S_k + (S_n - S_k)]^2 dP \\
&\geq \sum_{k=1}^n \int_{A_k} [S_k^2 + 2S_k(S_n - S_k)] dP = \sum_{k=1}^n \int_{A_k} S_k^2 dP + \sum_{k=1}^n \int_{A_k} 2S_k(S_n - S_k) dP \\
&\geq \varepsilon^2 \sum_{k=1}^n P(A_k) + 2\sum_{k=1}^n E[S_k \cdot \mathbb{1}_{A_K} \cdot (S_n - S_k)] \\
&= \varepsilon^2 P(A) + 2\sum_{k=1}^n E[S_k \cdot \mathbb{1}_{A_k}] \cdot E[S_n - S_k] = \varepsilon^2 P(A).
\end{aligned}$$

Usant que l'esperança de S_n és zero, la variància coincideix amb el moment de segon ordre. També hem fet servir que $S_k \cdot \mathbb{1}_{A_K}$ i $S_n - S_k$ són funcions, respectivament, de $(X_1, ..., X_k)$ i $(X_{k+1}, ..., X_n)$ que són vectors independents pel que $S_k \cdot \mathbb{1}_{A_K}$ i $S_n - S_k$ són variables aleatòries independents. □

Lema 4.0.4 (Criteri de convergència quasi segura de sèries). *Sigui $\{X_n\}_{n\in\mathbb{N}}$ una successió de variables aleatòries de quadrat integrable i centrades, tals que $E[X_n] = 0$ per tot $n \in \mathbb{N}$ i $\sum_{n=1}^\infty \mathrm{Var}[X_n] < \infty$. Aleshores, la sèrie $\sum_{n=1}^\infty X_n$ convergeix quasi segurament.*

Demostració. Considerem $S_n = X_1 + ...X_n$, $n \in \mathbb{N}$. Per cada $m \in \mathbb{N}$ obtenim, usant la desigual de Kolmogorov (Lema 4.0.3),

$$\begin{aligned}
P(\{\sup_{k\in\mathbb{N}}|S_{m+k} - S_m| > \varepsilon\}) &= \lim_{n\to\infty} P(\{\sup_{1\leq k\leq n}|S_{m+k} - S_m| > \varepsilon\}) \leq \lim_{n\to\infty} \frac{\mathrm{Var}[S_{m+n} - S_m]}{\varepsilon^2} \\
&= \lim_{n\to\infty} \frac{1}{\varepsilon^2} \sum_{k=m+1}^{m+n} \mathrm{Var}[X_k] = \frac{1}{\varepsilon^2} \sum_{k=m+1}^{\infty} \mathrm{Var}[X_k].
\end{aligned}$$

Fent $m \to \infty$, obtenim $\sup_{k\in\mathbb{N}}|S_{m+k} - S_m| \xrightarrow[m\to\infty]{P} 0$. Això implica l'existència d'una successió estrictament creixent de nombres naturals m_i tals que per certa parcial $\{S_{m_i}\}_{i\in\mathbb{N}}$,

$$\sup_{k\in\mathbb{N}}|S_{m_i+k} - S_{m_i}| \xrightarrow[i\to\infty]{q.s} 0,$$

pel que la successió $\{S_n\}_{n\in\mathbb{N}}$ és de Cauchy quasi segurament i, per tant, convergent quasi segurament. □

El següent resultat és un enunciat sobre sèries de nombres reals, no hi ha cap probabilitat involucrada.

Lema 4.0.5 (Lema de Kronecker). *Sigui $\{x_n\}_{n\in\mathbb{N}} \subseteq \mathbb{R}$ una successió de nombres reals i $\{a_n\}_{n\in\mathbb{N}} \subseteq \mathbb{R}$ una successió creixent de nombres reals positius amb límit $+\infty$. Suposem que la sèrie $\sum_{n=1}^\infty \frac{x_n}{a_n}$ és convergent. Aleshores,*

$$\lim_{n\to\infty} \frac{1}{a_n} \sum_{k=1}^n x_k = 0.$$

Demostració. Una prova detallada d'aquest resulat la podem trobar al llibre [5, Capítol 7. Secció 8. Lema 1]. $\square$

Lema 4.0.6 (Llei forta per variables de quadrat integrable). *Sigui $\{X_n\}_{n\in\mathbb{N}}$ una successió de variables aleatòries independents, centrades i de quadrat integrable. Sigui $S_n = X_1 + ... + X_n$, amb $n \in \mathbb{N}$ i $\{a_n\}_{n\in\mathbb{N}}$ una successió creixent de nombres reals positius amb límit $+\infty$. Suposem que $\sum_{n=1}^{\infty} \frac{\mathrm{Var}[X_n]}{a_n^2} < \infty$. Aleshores,*

$$\frac{S_n}{a_n} \xrightarrow[n\to\infty]{q.s} 0.$$

Demostració.

$$\sum_{n=1}^{\infty} \mathrm{Var}\left[\frac{X_n}{a_n}\right] = \sum_{n=1}^{\infty} \frac{\mathrm{Var}[X_n]}{a_n^2} < \infty$$

Pel que $\sum_{n=1}^{\infty} \frac{X_n}{a_n}$ convergeix quasi segurament i, pel Lema 4.0.5, $\lim_{n\to\infty} \frac{1}{a_n} \sum_{k=1}^{n} X_k = 0$ quasi segurament. $\square$

Lema 4.0.7 (Criteri d'integrabilitat de variables aleatòries). *Per a tota variable aleatòria X, es compleix*

$$\sum_{n=1}^{\infty} P(\{|X| \geq n\}) \leq E[|X|] \leq 1 + \sum_{n=1}^{\infty} P(\{|X| \geq n\})$$

En particular, X és integrable si i només si $\sum_{n=1}^{\infty} P(\{|X| \geq n\}) < \infty$.

Demostració. Provem primer la primera desigualtat:

$$\begin{aligned}
\sum_{n=1}^{\infty} P(\{|X| \geq n\}) &= \sum_{n=1}^{\infty}\sum_{k=n}^{\infty} P(\{k \leq |X| < k+1\}) = \sum_{k=1}^{\infty}\sum_{n=1}^{k} P(\{k \leq |X| < k+1\}) \\
&= \sum_{k=0}^{\infty} k \cdot P(\{k \leq |X| < k+1\}) \leq \sum_{k=0}^{\infty} \int_{\{k\leq|X|<k+1\}} |X| dP \\
&= \int_{\Omega} |X| dP = E[|X|],
\end{aligned}$$

on hem usat el Teorema de Fubini aplicat a sèries, que ens diu que es pot intercanviar l'ordre de sumació ja que els termes de la sèrie són positius. Passem a veure la segona desigualtat:

$$\begin{aligned}
E[|X|] = \int_{\Omega} |X| dP &= \sum_{k=0}^{\infty} \int_{\{k\leq|X|<k+1\}} |X| dP \leq \sum_{k=0}^{\infty} (k+1) \cdot P(\{k \leq |X| < k+1\}) \\
&= \sum_{k=0}^{\infty} k \cdot P(\{k \leq |X| < k+1\}) + \sum_{k=0}^{\infty} P(\{k \leq |X| < k+1\}) \\
&= \sum_{k=1}^{\infty} P(\{|X| \geq k\}) + 1.
\end{aligned}$$

$\square$

Passem finalment a demostrar La llei forta dels grans nombres (Teorema 4.0.2):

Demostració Teorema 4.0.2. Per demostrar (i) definim $Y_n := X_n \cdot \mathbb{1}_{\{|X_n|<n\}}$ per tot $n \in \mathbb{N}$ i, amb l'intenció d'usar el Lema 4.0.7, escrivim:

$$\frac{1}{n}\sum_{i=1}^{n} X_i = \frac{1}{n}\sum_{i=1}^{n}(X_i - Y_i) + \frac{1}{n}\sum_{i=1}^{n}(Y_i - E[Y_i]) + \frac{1}{n}\sum_{i=1}^{n} E[Y_i].$$

Veiem que quan $n \to \infty$, els dos primers sumands tenen límit zero quasi segurament i el tercer té límit m.

(a) Primer sumand, aplicant el criteri d'integrabilitat de variables aleatòries (Lema 4.0.7),

$$\sum_{n=1}^{\infty} P(\{X_n \neq Y_n\}) = \sum_{n=1}^{\infty} P(\{|X_n| \geq n\}) = \sum_{n=1}^{\infty} P(\{|X_1| \geq n\}) \leq E[|X_1|] < \infty.$$

Pel primer Lema de Borel-Cantelli tindrem $P(\{\lim_n \sup\{X_n \neq Y_n\}\}) = 0$, és a dir tindrem $P(\{\lim_n \inf\{X_n = Y_n\}\}) = 1$. Per tant, $\frac{1}{n}\sum_{i=1}^{n}(X_i - Y_i) \xrightarrow[n\to\infty]{q.s} 0$.

(b) Segon sumand aplicant el Lema 4.0.6:

$$\begin{aligned}
\sum_{n=1}^{\infty} \frac{\text{Var}[Y_n - E[Y_n]]}{n^2} &= \sum_{n=1}^{\infty} \frac{\text{Var}[Y_n]}{n^2} \leq \sum_{n=1}^{\infty} \frac{E[Y_n^2]}{n^2} = \sum_{n=1}^{\infty} \frac{1}{n^2} E[X_n^2 \cdot \mathbb{1}_{\{|X_n|<n\}}] \\
&= \sum_{n=1}^{\infty} \frac{1}{n^2} \sum_{k=1}^{n} E[X_1^2 \cdot \mathbb{1}_{\{k-1 \leq |X_1| < k\}}] \\
&= \sum_{k=1}^{\infty} E[X_1^2 \cdot \mathbb{1}_{\{k-1 \leq |X_1| < k\}}] \sum_{n=k}^{\infty} \frac{1}{n^2} \\
&\leq \sum_{k=1}^{\infty} k \cdot E[|X_1| \cdot \mathbb{1}_{\{k-1 \leq |X_1| < k\}}] \cdot \frac{2}{k} = 2E[|X_1|] < \infty.
\end{aligned}$$

Hem usat el Teorema de Fubini que ens diu que podem intercanviar l'ordre de sumació perquè tots els termes són positius. També hem fet servir l'acotació següent:

$$\sum_{n=k}^{\infty} \frac{1}{n^2} \leq \frac{1}{k^2} + \int_k^{\infty} \frac{1}{x^2} dx = \frac{1}{k^2} + \frac{1}{k} \leq \frac{2}{k}.$$

(c) Tercer sumand:

$$E[Y_n] = E[X_n \cdot \mathbb{1}_{\{|X_n|<n\}}] = E[X_1 \cdot \mathbb{1}_{\{|X_1|<n\}}] \xrightarrow[n\to\infty]{(1)} E[X_1] = m$$

(1): Pel Teorema de la Convergencia Dominada [3] (Teorema 2.2.16). Aleshores pel Criteri de la mitjana aritmètica de convergència de successions de nombres reals [4]: $\frac{1}{n}\sum_{i=1}^{n} E[Y_i] \xrightarrow[n\to\infty]{} m$.

Per provar (ii) fixem $K \in \mathbb{N}$ i usem el Lema 4.0.7 a la desigualtat:

$$\sum_{n=1}^{\infty} P\left(\left\{\frac{|X_n|}{n} \geq K\right\}\right) = \sum_{n=1}^{\infty} P\left(\left\{\frac{|X_1|}{K} \geq n\right\}\right) \geq \frac{E[|X_1|]}{K} - 1 = +\infty$$

[3] $|X_1|$ domina la successió

[4] Sigui $\{a_n\}_{n\in\mathbb{N}}$ una successió de reals tals que $\lim_{n\to\infty} a_n = A \in (\mathbb{R} \cup \{\infty\})$ aleshores, la successió de les seves mitjanes aritmètiques convergeix també a A, és a dir: $\lim_{n\to\infty} \frac{a_1 + \ldots + a_n}{n} = A$

Pel segon Lema de Borel-Cantelli tenim, $P\Big(\limsup_{n\to\infty}\left\{\frac{|X_n|}{n}\geq K\right\}\Big)=1$.
Donat que la intersecció numerable d'esdeveniments de probabilitat 1 té probabilitat 1, obtenim $P(A)=1$ amb,

$$A := \bigcap_{K=1}^{\infty} \limsup_{n\to\infty}\left\{\frac{|X_n|}{n}\geq K\right\}$$

A és el conjunt de $\omega\in\Omega$ tals que, per cada $K\in\mathbb{N}$, existeixen infinits $n\in\mathbb{N}$ complint $\frac{|X_n(\omega)|}{n}\geq K$. Podem usar $|S_n|$ en comptes de $|X_n|$ per la desigualtat següent:

$$K\leq\frac{|X_n|}{n}=\frac{|S_n-S_{n-1}|}{n}<\frac{|S_n|}{n}+\frac{|S_{n-1}|}{n-1},$$

i tenim $\frac{|S_n|}{n}\geq\frac{K}{2}$ o bé $\frac{|S_{n-1}|}{n-1}\geq\frac{K}{2}$. En conclusió, $\limsup_{n\to\infty}\frac{|S_n(\omega)|}{n}=+\infty$, per tot ω d'un conjunt de probabilitat 1. $\square$

Capítol 5

Funció característica i Teorema central del límit

5.1 Convergència feble de probabilitats i convergència en llei

Volem definir un nou concepte de convergència per a successions de probabilitats $\{\mu_n\}_{n\in\mathbb{N}}$ en $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ o, més en general, $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$. Sembla natural establir que $\mu_n \xrightarrow[n\to\infty]{} \mu$ si

$$\lim_{n\to\infty} \mu_n(B) = \mu(B), \quad \text{per a tot} \quad B \in \mathcal{B}(\mathbb{R}^n).$$

però aquesta exigència resulta ser massa forta i no és útil. Definim el concepte de convergència feble en probabilitats.

Definició 5.1.1. *Diem que una successió de probabilitats $\{\mu_n\}_{n\in\mathbb{N}}$ en $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ convergeix feblement a una probabilitat μ si i només si per a tota funció $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ contínua i acotada es compleix:*

$$\lim_{n\to\infty} \int_{\mathbb{R}^n} f d\mu_n = \int_{\mathbb{R}^n} f d\mu.$$

Ho abreviarem $\mu_n \xrightarrow[n\to\infty]{\omega} \mu$.

La convergència feble té una caracterització fàcil en termes de funcions de distribució. L'estudiarem només per el cas real, però és vàlida també per $\mathbb{R}^n$.

Teorema 5.1.2. *Sigui $\{\mu_n\}_{n\in\mathbb{N}}$ una successió de probabilitats en $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, amb funcions de distribució associades $\{F_n\}_{n\in\mathbb{N}}$. Sigui μ una probabilitat amb funció de distribució F. Aleshores, $\mu_n \xrightarrow[n\to\infty]{\omega} \mu$ si i només si $F_n(x) \xrightarrow[n\to\infty]{} F(x)$ per tot punt x de continuïtat de F.*

Demostració.

(i) Suposem que $\mu_n \xrightarrow[n\to\infty]{\omega} \mu$. Fixem $x \in \mathbb{R}$ un punt de continuïtat de la funció F,

$$F_n(x) = \mu_n((-\infty, x]) = \int_{\mathbb{R}} \mathbb{1}_{(-\infty,x]}(y)\mu_n(dy).$$

L'integrand no és una funció contínua, i per tant, no podem aplicar aquí la hipòtesi directament. L'aproximarem inferiorment i superiorment per funcions contínues. Considerem $\varepsilon > 0$ i les funcions contínues i afitades següents:

$$f_\varepsilon^+(y) = \mathbb{1}_{(-\infty,x]}(y) + \Big(1 - \frac{y-x}{\varepsilon}\Big)\mathbb{1}_{(x,x+\varepsilon)}(y),$$
$$f_\varepsilon^-(y) = \mathbb{1}_{(-\infty,x-\varepsilon]}(y) + \Big(\frac{x-y}{\varepsilon}\Big)\mathbb{1}_{(x-\varepsilon,x)}(y).$$

Tindrem les desigualtats: $\mathbb{1}_{(-\infty,x-\varepsilon]} \leq f_\varepsilon^- \leq \mathbb{1}_{(-\infty,x]} \leq f_\varepsilon^+ \leq \mathbb{1}_{(-\infty,x+\varepsilon]}$. Definim,

$$l = \lim_{n\to\infty} \inf F_n(x) \quad \text{i} \quad L = \lim_{n\to\infty} \sup F_n(x),$$

aleshores,

$$\begin{aligned} F(x-\varepsilon) = \mu((-\infty, x-\varepsilon]) &\leq \int_{\mathbb{R}} f_\varepsilon^- d\mu = \lim_n \int_{\mathbb{R}} f_\varepsilon^- d\mu_n \leq \lim_{n\to\infty} \inf \mu_n((-\infty, x]) \\ &= l \leq L = \lim_{n\to\infty} \sup \mu_n((-\infty, x]) \leq \lim_n \int_{\mathbb{R}} f_\varepsilon^+ d\mu_n = \int_{\mathbb{R}} f_\varepsilon^+ d\mu \\ &\leq \mu((-\infty, x+\varepsilon]) = F(x+\varepsilon). \end{aligned}$$

Com que $\varepsilon > 0$ és arbitrari i F és contínua en el punt x, fent tendir ε cap a 0 obtenim $l = L = F(x)$.

(ii) Per veure la segona implicació fixem x un punt de continuïtat i suposem $F_n(x) \longrightarrow F(x)$ quan n tendeix a infinit. Considerem $f : \mathbb{R} \longrightarrow \mathbb{R}$ una funció contínua i acotada i fixem $\varepsilon > 0$. Existeix un nombre $k \in \mathbb{N}$ tal que $\mu((-k,k]^c) < \varepsilon$, ja que $\cap_{k=1}^\infty (-k,k]^c = \emptyset$. Siguin a, b dos punts de continuïtat de F tals que $-a \leq k$ i $b \geq k$, aleshores $\mu((-a,b]^c) < \varepsilon$. Com que f és uniformement contínua en $[a,b]$, existeix un δ tal que si $x, y \in [a,b]$ i $|x-y| \leq \delta$ aleshores $|f(x) - f(y)| \leq \varepsilon$. Descomposem l'interval $(a,b]$ en un nombre finit d'intervals $I_i = (a_i, b_i]$ per $i = 1, ..., r$ de longitud més petita o igual a δ tals que els seus extrems siguin punts de continuïtat de F [1]. Podem escriure,

$$\begin{aligned} \Big| \int_{\mathbb{R}} f d\mu_n - \int_{\mathbb{R}} f d\mu \Big| &= \Big| \int_{(a,b]^c} f d\mu_n + \sum_{i=1}^r \int_{I_i} f d\mu_n - \int_{(a,b]^c} f d\mu - \sum_{i=1}^r \int_{I_i} f d\mu \Big| \\ &\leq \|f\|_\infty \Big(\mu_n((a,b]^c) + \mu((a,b]^c) \Big) + \sum_{i=1}^r \Big| \int_{I_i} f d\mu_n - \int_{I_i} f d\mu \Big|. \end{aligned}$$

Sabem que,

- $\mu_n((a,b]^c) = F_n(a) + 1 - F_n(b) \xrightarrow[n\to\infty]{} F(a) + 1 - F(b) = \mu((a,b]^c)$.
- $\lim_{n\to\infty} \mu_n(I_i) = \mu(I_i)$ per a tot $i = 1, .., r$.

D'altra banda tenim,

$$\begin{aligned} \Big| \int_{I_i} f d\mu_n - \int_{I_i} f d\mu \Big| &\leq \Big| \int_{I_i} \big(f - f(a_i)\big) d\mu_n \Big| + \Big| f(a_i) \big[\mu_n(I_i) - \mu(I_i)\big] \Big| \\ &+ \Big| \int_{I_i} \big(f - f(a_i)\big) d\mu \Big| \leq \varepsilon \big(\mu_n(I_i) + \mu(I_i)\big) + \Big| f(a_i) \big[\mu_n(I_i) - \mu(I_i)\big] \Big|. \end{aligned}$$

Recuperant l'expressó anterior obtenim,

$$\begin{aligned} \lim_n \sup \Big| \int_{\mathbb{R}} f d\mu_n - \int_{\mathbb{R}} f d\mu \Big| &\leq \|f\|_\infty \Big(\mu_n((a,b]^c) + \mu((a,b]^c) \Big) + \\ + \sum_{i=1}^r \Big| \int_{I_i} f d\mu_n - \int_{I_i} f d\mu \Big| &\leq 2\|f\|_\infty \mu((a,b]^c) + \sum_{i=1}^r 2\varepsilon \mu(I_i) \leq 2\varepsilon(\|f\|_\infty + 1). \end{aligned}$$

Com que ε és arbitrari, tindrem que $\lim_n \int_{\mathbb{R}} f d\mu_n = \int_{\mathbb{R}} f d\mu$.

[1] per fer això descomposem l'interval $(a,b]$ en intervals de longitud $\delta/2$ i després prenem un punt que sigui de continuïtat de F en cadascun d'aquests intervals. Notem que tots aquests punts a, b, a_i, b_i, existeixen gràcies al fet que F té límits 0 i 1 en els infinits i al fet de que a tot interval no degenerat és possible trobar punts de continuïtat de F

□

Corol·lari 5.1.3. *El límit feble, si existeix, és únic.*

Demostració. Segons el teorema anterior, les funcions de distribució de dos possibles límits han de coincidir en tot punt de continuïtat. En un resultat previ (Teorema 1.5.9) hem vist que el conjunt de punts de discontinuïtat és numerable, per tant, els punts de continuïtat són densos a $\mathbb{R}$. Com que les funcions de distribució han de ser contínues per la dreta, de manera que les funcions han de coincidir. □

L'unicitat del límit feble també és certa en $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$.

És possible introduïr una distància en el conjunt $\mathfrak{B}$ de totes les probabilitats que metritzi la convergència feble. En $\mathbb{R}$ es pot definir de la manera següent: Si μ i ν són dues probabilitats amb funcions de distribució F i G respectivament, la distància entre μ i ν és:

$$d_\omega(\mu, \nu) := \inf\{\varepsilon > 0 : G(x-\varepsilon) - \varepsilon \leq F(x) \leq G(x+\varepsilon) + \varepsilon\}. \tag{5.1.1}$$

Tota variable aleatòria $X : \Omega \longrightarrow \mathbb{R}$ indueix una probabilitat en $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Podem transportar el concepte de convergència feble de probabilitats a un concepte de convergència de variables aleatòries per tal de simplificar la notació.

Definició 5.1.4. *Una successió $\{X_n\}_{n\in\mathbb{N}}$ de variables aleatòries definides en un espai de probabilitat $(\Omega, \mathcal{A}, P)$ convergeix en llei a una variable aleatòria X si i només si les lleis de X_n convergeixen feblement a la llei de X, és a dir:*

$$P \circ X_n^{-1} \xrightarrow[n\to\infty]{\omega} P \circ X^{-1},$$

o equivalentment, si per tota funció $f : \mathbb{R} \longrightarrow \mathbb{R}$ contínua i afitada, es compleix,

$$\lim_{n\to\infty} E[f(X_n)] = E[f(X)].$$

Ho abreviarem $X_n \xrightarrow[n\to\infty]{\mathcal{L}} X$.

Notem que en lloc de X podem posar qualsevol variable que tingui la mateixa llei que X, i per tant, el límit no és únic. Per això si μ és la llei límit, té sentit també escriure $X_n \xrightarrow[n\to\infty]{\mathcal{L}} \mu$.

No donarem moltes explicacions sobre altres tipus de convergència de successions de variables aleatòries, ja que molts resultats estan fets a l'assignatura de Probabilitats i les definicions les podem trobar a l'annex.

5.2 Funcions característiques

En aquest capítol introduïrem una tècnica per tractar la convergència feble de probabilitats: les funcions característiques de Paul Lévy [2].

Definició 5.2.1. *Sigui μ una probabilitat en $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Definim la funció característica de μ de la següent manera:*

$$\begin{aligned} \varphi_\mu : \mathbb{R} &\longrightarrow \mathbb{C} \\ t &\longmapsto \varphi_\mu(t) = \int_{\mathbb{R}} e^{itx}\mu(dx) = \int_{\mathbb{R}} \cos tx\mu(dx) + i\int_{\mathbb{R}} \sin tx\mu(dx) \end{aligned}$$

[2] [8]

Si $X : \Omega \longrightarrow \mathbb{R}$ és una variable aleatòria en un cert espai de probabilitat $(\Omega, \mathcal{A}, P)$, la funció característica de X és la funció característica de la seva llei, és a dir:

$$\varphi_X(t) = \int_{\mathbb{R}} e^{itx} P_X(dx) = E[e^{itX}].$$

Anàlogament, si μ és una probabilitat en $\mathbb{R}^n$, la funció característica de μ es defineix com

$$\begin{aligned} \varphi_\mu : \mathbb{R}^n &\longrightarrow \mathbb{C} \\ t \longmapsto \varphi_\mu(t) &= \int_{\mathbb{R}_n} e^{i<t,x>} \mu(dx), \end{aligned}$$

on $< t, x >$ és el producte escalar habitual en $\mathbb{R}^n$. La funció caracterítica d'un vector aleatori $X = (X_1, ..., X_n)$ serà la funció característica de la seva llei.

Podem consultar a l'annex algunes de les propietats fonamentals de les funcions característiques.

La funció caraterística d'una probabilitat μ en $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ està relacionada de forma senzilla amb el comportament a l'origen de la funció característica de μ i dels seus moments, com veurem en els següents teoremes. Establirem primer un resultat tècnic de derivació sota el signe integral.

Lema 5.2.2. *Sigui $f(t, x) : \mathbb{R}^2 \longrightarrow \mathbb{R}$ una funció contínua amb derivada parcial contínua $\frac{\partial f}{\partial t}$. Suposem que $|f(t, x)| + |\frac{\partial f}{\partial t}(t, x)| \leq g(x)$ on g és una funció integrable respecte una probabilitat μ en $\mathbb{R}$. Aleshores, la funció $F(t) = \int_{\mathbb{R}} f(t, x)\mu(dx)$ és derivable i es compleix,*

$$F'(t) = \int_{\mathbb{R}} \frac{\partial f}{\partial t}(t, x)\mu(dx).$$

Demostració. Una prova detallada d'aquesta demostració la podem trobar al llibre [10, Lema 12.1]. □

Recordem que una funció $\varphi : \mathbb{R} \longrightarrow \mathbb{C}$ és derivable si i només si ho són la seva part real i imaginària. El lema anterior és igualment cert per funcions que prenen valors en $\mathbb{C}$.

Teorema 5.2.3. *Sigui μ una probabilitat en $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ amb moment d'ordre $n \geq 1$ finit. Aleshores, φ_μ és n vegades derivable i*

$$\varphi_\mu^{(k)}(t) = i^k \int_{\mathbb{R}} x^k e^{itx} \mu(dx), \quad \text{per a tot} \quad k \in \{1, ..., n\}.$$

En particular, si designem per m_i amb $i \in \{1, ..., n\}$ als moments de la probabilitat μ, es compleix $\varphi_\mu^{(k)}(0) = i^k m_k$, per $k \in \{1, ..., n\}$.

Demostració. Apliquem el Lema 5.2.2 a la funció e^{itx} que compleix $|\frac{\partial}{\partial t}(e^{itx})| = |ixe^{itx}| = |x|$. Sabem que $\int_{\mathbb{R}} |x|\mu(dx) < \infty$, en conseqüència, $\varphi_\mu(t) = \int_{\mathbb{R}} e^{itx}\mu(dx)$ és derivable i $\varphi_\mu^{'}(t) = i \int_{\mathbb{R}} xe^{itx}(dx)$. Aplicant iterativament el Lema 5.2.2 a les funcions $\varphi_\mu^{'}, \varphi_\mu^{''}, ..., \varphi_\mu^{(n-1)}$ arribariem al resultat desitjat. □

El resultat següent estableix un tipus de recíproc d'aquest teorema.

Teorema 5.2.4. *Sigui μ una probabilitat sobre $\mathbb{R}$ Suposem que la funció característica φ_μ és k vegades diferenciable en un entorn del zero. Aleshores μ té moments d'ordre $2n$, amb $2n \leq k$.*

Demostració. Volem veure que $\int_{\mathbb{R}} x^{2n}\mu(dx) < \infty$. Suposem primer que $k = 2$ i $n = 1$, tindrem,

$$\begin{aligned}\varphi''_\mu(0) &= \lim_{h\to 0} \frac{\varphi_\mu(h) - 2\varphi_\mu(0) + \varphi_\mu(-h)}{h^2} = \lim_{h\to 0}\int_{\mathbb{R}} \frac{e^{ihx} - 2 + e^{-ihx}}{h^2}\mu(dx)\\ &= -2\lim_{h\to 0}\int_{\mathbb{R}} \frac{1-\cosh hx}{h^2}\mu(dx).\end{aligned}$$

Pel Lema de Fatou (Lema 2.2.14),

$$\int_{\mathbb{R}} x^2\mu(dx) = 2\int_{\mathbb{R}} \lim_{h\to 0}\frac{1-\cosh hx}{h^2}\mu(dx) \le 2\liminf_{h\to 0}\int_{\mathbb{R}}\frac{1-\cosh hx}{h^2}\mu(dx) = -\varphi''_\mu(0) < \infty.$$

El cas general es demostra per inducció: Suposem que es compleix per $\int_{\mathbb{R}} x^{2(n-1)}\mu(dx) < \infty$ i sabem que $2n \le k$ pel que φ_μ és $2n$ derivable en un entorn del zero. Pel Teorema 5.2.3 sabem que,

$$\varphi_\mu^{(2n-2)}(t) = (-1)^{n-1}\int_{\mathbb{R}} x^{2n-2}e^{itx}\mu(dx).$$

Per tot $A \in \mathcal{B}(\mathbb{R})$ posem $\nu(A) = \frac{\int_A x^{2n-2}\mu(dx)}{\int_{\mathbb{R}} x^{2n-2}\mu(dx)}$. Aleshores, ν és una probabilitat en $\mathbb{R}$ que té per funció característica:

$$\varphi_\nu(t) = \frac{1}{K}(-1)^{n-1}\varphi_\mu^{(2n-2)}(t), \quad \text{on} \quad K = \int_{\mathbb{R}} x^{2n-2}\mu(dx).$$

Notem que hem suposat $K > 0$, si tinguessim $K = 0$, aleshores $\mu = \delta_0$, $\varphi_\mu = 1$ i el teorema és trivial. Tenim per hipòtesi que $\varphi''_\nu(t)$ existeix en un entorn del zero i, utilitzant el resultat en el cas $k = 2$ obtenim,

$$-\varphi''_\nu(t) = \int_{\mathbb{R}} x^2\nu(dx) = \frac{1}{K}\int_{\mathbb{R}} x^{2n}\mu(dx),$$

és a dir, μ té un moment finit d'ordre $2n$. □

El següent resultat estableix una fòrmula per calcular la funció de distribució a partir de la seva funció característica. Per tant, la funció característica té aquest nom perquè *caracteritza* la probabilitat [3]. Això és deduirà d'una fòrmula que ens determina la funció de distribució a partir de la funció característica.

Teorema 5.2.5 (Fòrmula d'inversió). *Sigui μ una probabilitat en $\mathbb{R}, \mathcal{B}(\mathbb{R})$, amb funció característica φ_μ i funció de distribución F. Aleshores, per a tots $a < b$ punts de continuïtat de F [4],*

$$\mu((a,b]) = F(b) - F(a) = \lim_{M\to\infty}\frac{1}{2\pi}\int_{-M}^{M}\frac{e^{-ita} - e^{-itb}}{it}\varphi_\mu(t)dt.$$

Demostració. Una prova detallada d'aquesta demostració la podem trobar al llibre [3, Capítol 5. Secció 26. Teorema 26.2]. □

Corol·lari 5.2.6. *Si μ i ν són dues probabilitats en $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ amb la mateixa funció característica, aleshores $\mu = \nu$.*

[3]mesures diferents no poden tenir una mateixa funció característica

[4]$\mu\{a\} = \mu\{b\} = 0$

Demostració. La funció característica determina la funció de distribució en els seus punts de continuïtat, i per tant, a tot arreu. □

Observació 5.2.7. Com a aplicació d'aquest teorema es pot demostrar la propietat d'injectvitat (vi) de la Proposició A.3.1. En efecte si F_1 i F_2 són les funcions de distribució de certes probabilitats μ_1, μ_2 de $\mathbb{R}$ que tenen la mateixa funció característica tindrem que $F_1(\beta) - F_2(\alpha) = F_2(\beta) - F_1(\alpha)$ per a tota parella $\alpha < \beta$ de punts on F_1, F_2 siguin contínues. Llavors fent tendir $\alpha \longrightarrow -\infty$ obtenim $F_1(\beta) = F_2(\beta)$ en un conjunt dens i per tant, $F_1 = F_2$.

El resultat següent ens diu que quan la funció característica és integrable, la fórmula d'inversió esdevé més simple i ens podem estalviar el pas al límit. En particular, tenim també que μ és contínua i podem trobar la seva densitat com a transformada de Fourier inversa de la funció característica.

Proposició 5.2.8. *Si la funció característica φ_μ d'una probabilitat μ sobre $\mathbb{R}$ és integrable respecte la mesura de Lebesgue, aleshores μ és absolutament contínua, i la seva densitat és,*

$$f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-itx} \varphi_\mu(t) dt.$$

Demostració. Tenint en compte que,

$$\Big| \mathbb{1}_{[-M,M]}(t) \frac{e^{-ita} - e^{-itb}}{it} \varphi_\mu(t) \Big| \le \Big| \int_a^b e^{itx} dx \cdot \varphi_\mu(t) \Big| \le (b-a)|\varphi(t)|,$$

pel teorema de la convergència dominada (Teorema 2.2.16) i el teorema de Fubini tenim,

$$\begin{aligned} F(b) - F(a) &= \lim_{M\to\infty} \frac{1}{2\pi} \int_{-M}^{M} \frac{e^{-ita} - e^{-itb}}{it} \varphi_\mu(t) dt = \frac{1}{2\pi} \int_{\mathbb{R}} \varphi_\mu(t) \Big(\int_a^b e^{-itx} dx \Big) dt \\ &= \frac{1}{2\pi} \int_a^b \Big(\int_{\mathbb{R}} e^{-itx} \varphi_\mu(t) dt \Big) dx = \frac{1}{2\pi} \int_a^b f(x) dx, \end{aligned}$$

per a tots els punts $a, b \in \mathbb{R}$ de continuïtat de F. A més, la integral respecte dx és una funció contínua dels límits d'integració, pel que F és contínua. □

Com a últim punt d'aquest apartat considerem, com a cas interessant, el càlcul de la funció característica de la *Llei normal* que usarem per el resultat del Teorema central del límit de Lévy-Lindeberg (Teorema 5.4.1).

Exemple 5.2.9. Sigui μ la llei $N(0,1)$, $\alpha(t) = \int_{\mathbb{R}} e^{-x^2/2} \cos tx dx$ i $\beta(t) = \int_{\mathbb{R}} e^{-x^2/2} \sin tx dx$. Tindrem,

$$\varphi_\mu(t) = \Big(\alpha(t) + i\beta(t) \Big) \frac{1}{\sqrt{2\pi}}$$

Clarament, $\beta(t) = 0$. Pel Lema 5.2.2 aplicat a la funció $f(x,t) = e^{-x^2/2} \sin tx$, la funció $\alpha(t)$ és derivable. En efecte,

$$\Big| \frac{\partial f}{\partial t} \Big| = \Big| -x \cos tx e^{\frac{-x^2}{2}} \Big| \le |x| e^{\frac{-x^2}{2}} \quad \text{i} \quad \int_{\mathbb{R}} |x| e^{\frac{-x^2}{2}} dx < \infty.$$

Integrant per parts obtenim,

$$\alpha'(t) = -\int_{\mathbb{R}} x e^{\frac{-x^2}{2}} \sin tx dx = \Big[e^{\frac{-x^2}{2}} \sin tx dx \Big]_{-\infty}^{+\infty} - \int_{\mathbb{R}} e^{\frac{-x^2}{2}} t \cos tx dx = -t\alpha(t).$$

Com que $\alpha(0) = \sqrt{2\pi}$, això ens diu que $\alpha(t) = \sqrt{2\pi} e^{-t^2/2}$ i, en conclusió $\varphi(t) = e^{-t^2/2}$.

Observació 5.2.10. La funció característica és un exemple del que s'anomenen *transformades*, en particular, en el context de l'Anàlisi matemàtica general rep el nom de *transformada de Fourier*. A l'inici d'aquesta secció comentem que va ser Paul-Lévy qui va introduïr el concepte de funció característica, ja que va ser el que va adaptar la transformada de Fourier per tal de desenvolupar, tal i com veurem a la secció 5.4 el problema del límit central. Prèviament, idees similars havien estat utilitzades al llarg del segle XIX, com per exemple, el que avui anomenem la *transformada de Laplace*, ψ_μ, que fou inventada per Laplace amb el mateix objectiu:

$$\begin{aligned}\psi_\mu :&\mathbb{R} \longrightarrow \mathbb{R}\\ t \longmapsto& \int_{\mathbb{R}} e^{tx}\mu(dx)\end{aligned}$$

Quan μ és una probabilitat, el nom tradicional és *funció generatriu de moments*, a causa de que $\psi_\mu(0)$ és el moment d'ordre n de μ (si aquest existeix i és finit). El problema amb la transformada de Laplace és que no esta definida per a tot $t \in \mathbb{R}$, si $t = 0$, aquest no està a l'interior del seu domini i la fòrmula anterior pels moments no val.

És interessant comentar que quan es treballa amb lleis de probabilitat discretes, és a vegades més interessant utilitzar la funció generatriu en comptes de la funció característica: Considerem una probabilitat μ sobre $\mathbb{N}$, caracteritzada per una successió $p_i = \mu(\{i\})$ tal que $p_i \geq 0$ i $\sum_{i=0}^{\infty} p_i = 1$. Tenint en compte la mateixa notació,

Definició 5.2.11. *Anomenem funció generatriu de μ, a la funció G_μ definida en el disc $\{z \in \mathbb{C} : |z| \leq 1\}$ de la següent manera:*

$$G_\mu(z) = \sum_{n=0}^{\infty} p_n z^n.$$

En particular, les probabilitats p_n són els coeficients del desenvolupament en sèrie de potències de G_μ [5] a l'origen, $p_n = \frac{1}{n!}G_\mu^{(n)}(0)$. De manera anàloga a les funcions característiques, si X és una variable aleatòria a valors naturals, anomenem funció generatriu de X a la funció generatriu de la seva llei, és a dir $G_\mu(z) = E(z^X)$.

La relació entre la funció generatriu i la funció característica és,

$$G_\mu(e^{it}) = \varphi_\mu(t), \quad \text{per a tot} \quad t \in \mathbb{R}.$$

5.3 El teorema de continuïtat

En aquest apartat veurem que si una successió de probabilitats convergeix feblement, aleshores les funcions característiques corresponents convergeixen puntualment a la funció caraterística del límit, i recíprocament. Aquest resultat es coneix com el Teorema de continuïtat de Lévy i per demostrar una de les implicacions ens calen uns resultats preliminars que enunciarem seguidament.

Teorema 5.3.1 (Teorema de Helly). *Sigui $\{\mu_n\}$ una succesió de probabilitats de $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, amb funcions de distribució associades $\{F_n\}_{n\in\mathbb{N}}$. Aleshores existeix una successió parcial*

[5] Notem que la sèrie $\sum_{n=0}^{\infty} p_n z^n$ és convergent per a tot complex z tal que $|z| \leq 1$, és a dir, el radi de convergència d'aquesta sèrie serà més gran o igual a 1 i la sèrie convergeix uniformement en tot el disc $\{z \in \mathbb{C} : |z| \leq r\}$ per $r < 1$

$\{F_{n_k}\}_{k\in\mathbb{N}}$ de $\{F_n\}_{n\in\mathbb{N}}$ i una funció de distribució F fitada per 1, d'una mesura finita (no necessariament una probabilitat) tals que,

$$\lim_{k\to\infty} F_{n_k} = F(x) \quad \text{per a tot } x \text{ punt de continuïtat de } F.$$

Demostració. Una prova detallada d'aquest resultat la podem trobar al llibre [3, Capítol 5. Apartat 25. Helly's Theorem] □

Considerem el següent exemple: sigui $\mu_n = \delta_n$ (Delta de Dirac en el punt n), aleshores les funcions de distribució associades $\{F_n\}_{n\in\mathbb{N}}$ convergeixen puntualment a $F \equiv 0$. Aquest cas mostra que la funció límit del Teorema de Helly no té per què ser la funció de distribució d'una probabilitat. El concepte que veurem a continuació defineix la propietat que es necessita per assegurar que la convergència sigui a una probabilitat.

Definició 5.3.2. *Sigui $\mathcal{M}$ un conjunt de probabilitats sobre $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Es diu que $\mathcal{M}$ és un conjunt ajustat si i només si per tot $\varepsilon > 0$ existeixen $a, b \in \mathbb{R}$ amb $a < b$ tals que per tota probabilitat $\mu \in \mathcal{M}$ es té $\mu((a, b]^c) < \varepsilon$.*

Notem que si a un conjunt ajustat li afegim un nombre finit de probabilitats, el conjunt resultant torna a ser ajustat.

Teorema 5.3.3 (Teorema de Prokhorov). *Un conjunt de probabilitats $\mathcal{M}$ sobre $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ és ajustat si i només és un conjunt relativament compacte en l'espai mètric $(\mathfrak{B}, d_\omega)$ definit per (5.1.1). És a dir, si i només si tota successió $\{\mu_n\}_{n\in\mathbb{N}} \subseteq \mathcal{M}$ té una parcial convergent en $(\mathfrak{B}, d_\omega)$.*

Demostració. Una prova detallada d'aquest teorema la podem trobar al llibre [4, Volume II, Teorema 8.6.2]. □

Lema 5.3.4 (Desigualtat de truncació). *Sigui μ una probabilitat en $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, amb funció característica φ_μ. Aleshores, per tot $a > 0$,*

$$\mu(\{x : |x| \geq \frac{2}{a}\}) \leq \frac{1}{a}\int_{-a}^{a} (1 - \varphi_\mu(t))dt.$$

Demostració. Utilitzant el Teorema de Fubini, ja que $|1 - e^{itx}| \leq 2$, la definició de la exponencial complexa i el fet que la funció sinus és una funció imparella,

$$\begin{aligned}
\frac{1}{a}\int_{-a}^{a}(1-\varphi_\mu(t))dt &= \frac{1}{a}\int_{-a}^{a}\left(\int_{\mathbb{R}}(1-e^{itx})\mu(dx)\right)dt = \frac{1}{a}\int_{\mathbb{R}}\left(\int_{-a}^{a}(1-e^{itx})dt\right)\mu(dx) \\
&= \frac{1}{a}\int_{\mathbb{R}}\left(\int_{-a}^{a}(1-\cos tx)dt\right)\mu(dx) = \frac{1}{a}\int_{\mathbb{R}}2\left(a - \frac{\sin ax}{x}\right)\mu(dx) \\
&\geq \int_{\{x:|ax|\geq 2\}} 2\left(1 - \frac{\sin ax}{ax}\right)\mu(dx) \\
&\geq 2\inf_{|u|\geq 2}\left(1 - \frac{\sin u}{u}\right)\cdot \mu(\{x : |ax| \geq 2\}) \geq \mu\left(\left\{x : |x| \geq \frac{2}{a}\right\}\right),
\end{aligned}$$

ja que $|\frac{\sin u}{u}| \leq \frac{1}{2}$ si $|u| \geq 2$ i queda demostrat el lema. □

Teorema 5.3.5 (Teorema de continuïtat de Lévy). *Sigui $\{\mu_n\}_{n\in\mathbb{N}}$ una successió de probabilitats en $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ amb funcions característiques associades $\{\varphi_n\}_{n\in\mathbb{N}}$. Aleshores,* *(i) Si μ_n convergeix feblement a una probabilitat μ i φ és la seva funció característica, es té que $\lim_n \varphi_n(t) = \varphi(t)$ per a tot $t \in \mathbb{R}$.*

(ii) Si $\lim_n \varphi_n(t) = \varphi(t)$ és una funció contínua en el zero, aleshores φ és la funció característica d'una probabilitat μ i $\mu_n \xrightarrow[n\to\infty]{\omega} \mu$.

Demostració.

(i) És una conseqüència de la definició de convergència feble, ja que les funcions $\cos tx$ i $\sin tx$, part real i imaginària de la funció e^{itx}, són contínues i acotades:

$$\varphi_n(t) = \int_{\mathbb{R}} e^{itx}\mu_n(dx) \xrightarrow[n\to\infty]{} \int_{\mathbb{R}} e^{itx}\mu(dx) = \varphi(t).$$

(ii) Veurem primer que els elements de la successió $\{\mu_n\}_{n\in\mathbb{N}}$ constitueixen un conjunt ajustat de probabilitats. Usem la desigualtat de truncació (Lema 5.3.4) i el teorema de convergència dominada (Teorema 2.2.16), ja que $|1-\varphi(t)| \le 2$. Per $a > 0$ obtenim,

$$\mu_n\Big(\Big[\frac{-2}{a}, \frac{2}{a}\Big]^c\Big) = \mu_n(\{x : |x| \ge \frac{2}{a}\}) \le \frac{1}{a}\int_{-a}^{a}\Big(1-\varphi_n(t)\Big)dt \xrightarrow[n\to\infty]{} \frac{1}{a}\int_{-a}^{a}\Big(1-\varphi(t)\Big)dt.$$

D'altra banda, com que φ és contínua en el zero, el teorema fonamental del càlcul implica,

$$\lim_{a\to 0}\frac{1}{a}\int_{-a}^{a}\big(1-\varphi(t)\big)dt = 2(1-\varphi(0)) = 0.$$

Sigui $a > 0$, fixem un $\varepsilon > 0$ tal que $0 \le \frac{1}{a}\int_{-a}^{a}\Big(1-\varphi(t)\Big)dt \le \frac{\varepsilon}{2}$. Aleshores a partir d'un cert n_0 tindrem,

$$\mu_n\Big(\Big[\frac{-2}{a}, \frac{2}{a}\Big]^c\Big) \le \frac{1}{a}\int_{-a}^{a}\Big(1-\varphi_n(t)\Big)dt \le \frac{1}{a}\int_{-a}^{a}\Big(1-\varphi(t)\Big)dt + \frac{\varepsilon}{2} \le \varepsilon.$$

Per tant, $\{\mu_n : n \ge n_0\}$ és ajustat, el que implica que $\{\mu_n : n \in \mathbb{N}\}$ també ho és. El Teorema de Prokhorov (Teorema 5.3.3) aplicat a aquest últim conjunt ens diu,en particular, que tota successió parcial $\{\mu_{n_k}\}_{k\in\mathbb{N}}$ té una parcial convergent $\{\mu_{n_{k_l}}\}_{l\in\mathbb{N}}$ a un cert límit μ. Per la part (i) del teorema,

$$\varphi_\mu(t) = \lim_{l\to\infty}\varphi_{n_{k_l}}(t) = \varphi(t).$$

Hem vist que tota parcial de $\{\mu_n\}_{n\in\mathbb{N}}$ té una parcial convergent a un mateix límit μ. Com en tot espai mètric, això és equivalent al fet que $\{\mu_n\}_{n\in\mathbb{N}}$ convergeix a μ.

□

5.4 Teorema central del límit i llei normal multidimensional

La llei forta dels grans nombres (Teorema 4.0.2) ens diu que la successió $\{\frac{S_n}{n}\}_{n\in\mathbb{N}}$ [6], amb moment de primer ordre finit m, convergeix quasi segur a m. Per tant, les lleis de $\frac{S_n}{n}$ convergeixen feblement a la llei δ_m. Si les variables tenen variància finita σ^2, podem

[6] on S_n és la suma de n variables aleatòries identicament distribuïdes

impedir aquesta degeneració de les lleis cap a δ_m: Restant a $\frac{S_n}{n}$ la seva esperança i dividint per la desviació típica σ, obtenim variables d'esperança 0 i de variància 1:

$$\frac{\frac{S_n}{n} - E\left[\frac{S_n}{n}\right]}{\sqrt{\mathrm{Var}\left[\frac{S_n}{n}\right]}} = \frac{\frac{S_n}{n} - m}{\frac{\sigma}{\sqrt{n}}} = \frac{S_n - nm}{\sigma\sqrt{n}}.$$

No podem esperar convergència quasi segura, L^p o en probabilitat de la successió $\frac{S_n - nm}{\sigma\sqrt{n}}$ a cap variable aleatòria concreta, mirem si pot convergir en llei: Abraham De Moivre en el llibre *The Doctrine of Chances* [7] va formular el següent resultat:

> Si $\{X_n\}_{n\in\mathbb{N}}$ és una successió de variables aleatòries independents que prenen els valors 0 i 1 amb probabilitat $\frac{1}{2}$, aleshores les lleis de $\frac{S_n - \frac{1}{2}n}{\frac{1}{2}\sqrt{n}}$ convergeixen feblement a la llei $N(0,1)$.

El resultat de De Moivre va ésser generalitzat per Laplace (1810) al cas de variables discretes i simètriques, i constitueix la versió més senzilla del teorema central del límit. A continuació veurem una versió de Lévy-Lindeberg que generalitza el Teorema de De Moivre-Laplace.

Teorema 5.4.1 (Teorema central del límit de Lévy-Lindeberg). *Sigui $\{X_n\}_{n\in\mathbb{N}}$ una successió de variables aleatòries independents, identicament distribuïdes i de quadrat integrable, amb mitjna m i variància σ^2. Si $S_n = X_1 + ... + X_n$, aleshores*

$$\frac{S_n - nm}{\sigma\sqrt{n}} \xrightarrow[n\to\infty]{\mathcal{L}} N(0,1).$$

Demostració. Volem veure que les funcions característiques de $Y_n := \frac{S_n - nm}{\sigma\sqrt{n}}$ convergeixen a la funció característica de la llei $N(0,1)$ que és $e^{-t^2/2}$ (Exemple 5.2.9). Observem que $E(Y_n) = 0$ i $\sigma^2(Y_n) = 1$. Sigui φ_n la funció característica de $\frac{X_1 - m}{\sigma\sqrt{n}}$. Per la propetat (ix) de la Proposició A.3.1, i tenint en compte que les variables són indepedents i tenen la mateixa llei,

$$\varphi_{Y_n}(t) = E[e^{itY_n}] = E\left[\exp\left\{it\frac{S_n - mn}{\sigma\sqrt{n}}\right\}\right] = \left(E\left[\exp\left\{it\frac{X_1 - m}{\sigma\sqrt{n}}\right\}\right]\right)^n = \big(\varphi_n(t)\big)^n.$$

Com que X_1 és de quadrat integrable, φ_n és dues vegades derivable i tenim,

$$\varphi_n^{'}(t) = iE\left[\frac{X_1 - m}{\sigma\sqrt{n}}\exp\left\{it\frac{X_1 - m}{\sigma\sqrt{n}}\right\}\right],$$
$$\varphi_n^{''}(t) = -E\left[\left(\frac{X_1 - m}{\sigma\sqrt{n}}\right)^2\exp\left\{it\frac{X_1 - m}{\sigma\sqrt{n}}\right\}\right].$$

Aplicant la Fòrmula de Taylor a la funció $\varphi_n(t)$, en el punt $t = 0$,

$$\varphi_n(t) = \varphi_n(0) + \varphi_n^{'}(0)t + \frac{1}{2}\varphi_n^{''}(\xi)t^2 = 1 + 0 + \frac{1}{2}\varphi_n^{''}(\xi)t^2, \quad \text{per algun} \quad |\xi| \in (0,t).$$

Usem que si $\{a_n\}_{n\in\mathbb{N}}$ és una successió creixent amb límit $+\infty$ i existeix el límit de la successió $\{b_n\}_{n\in\mathbb{N}}$, aleshores es compleix $\lim\limits_{n\to\infty}\big(1 + \frac{1}{a_n}\big)^{a_n b_n} = \exp\big(\lim\limits_{n\to\infty} b_n\big)$. Escrivim,

$$\left(1 + \frac{t^2}{2}\varphi_n^{''}(\xi)\right)^n = \left(1 + \frac{1}{\left(\frac{t^2}{2}\varphi_n^{''}(\xi)\right)^{-1}}\right)^{n\left(\frac{t^2}{2}\varphi_n^{''}(\xi)\right)^{-1}\left(\frac{t^2}{2}\varphi_n^{''}(\xi)\right)}$$

[7] [9]

Si comprovem que $n\varphi_n''(\xi) \xrightarrow[n\to\infty]{} -1$ estem en condicions d'aplicar el resultat anterior i tindrem convergència precisament a $e^{-t^2/2}$. Gràcies al Teorema de Convergència Dominada (Teorema 2.2.16), tenim

$$n\varphi_n''(\xi) = -E\left[\left(\frac{X_1 - m}{\sigma}\right)^2 \exp\left\{i\xi \frac{X_1 - m}{\sigma\sqrt{n}}\right\}\right] \xrightarrow[n\to\infty]{} -E\left[\left(\frac{X_1 - m}{\sigma}\right)^2\right] = -1.$$

□

És interessant, especialment per diferents aplicacions en estadística, poder disposar d'una versió n-dimensional del teorema del límit de Lévy-Lindeberg. Ens cal introduïr prèviament el concepte de *llei normal multidimensional.*

Proposició 5.4.2. *Sigui $m \in \mathbb{R}^n$ i Λ una matriu simètrica d'ordre n definida positiva. Exiteix una probabilitat en $\mathbb{R}^n$, que designarem per $N(m, \Lambda)$, i anomenarem llei normal n-dimensional, que té per funció característica*

$$\varphi_n(t) = e^{it^*m - \frac{1}{2}t^*\Lambda t}, \quad \text{per a tot } t \in \mathbb{R}^n.$$

Demostració. Sigui C una matriu ortogonal tal que $C\Lambda C^* = D$, on D és una matriu diagonal. Els elements de la diagonal de D, que designarem per $\lambda_1, ..., \lambda_n$, coincideixen amb els valors propis de Λ i, en conseqüència, són no negatius. Tindrem $\Lambda = C^*DC$. Considerem un vector aletori $Y = (Y_1, ..., Y_n)$ amb les components independents i amb lleis $N(0, \lambda_i)$ si $\lambda_i \neq 0$ i $Y_i = 0$ si $\lambda_i = 0$ per $i \in \{1, ..., n\}$. Definim $X = C^*Y + m$. La funció característica del vector Y val, per $t \in \mathbb{R}^n$,

$$\varphi_Y(t) = E\left[\exp i \sum_{j=1}^{n} t_j Y_j\right] = \prod_{j=1}^{n} e^{\frac{-1}{2}t_j^2\lambda_j} = e^{\frac{-1}{2}t^*Dt}.$$

Per tant, la funció característica del vector X serà,

$$\varphi_X(t) = e^{it^*m}\varphi_Y(Ct) = e^{it^*m}e^{\frac{-1}{2}(Ct)^*DCt} = e^{it^*m - \frac{1}{2}t^*\Lambda t},$$

i la llei del vector X és la probabilitat que busquem. □

A l'annex podem trobar enunciades de manera esquematitzada algunes de les propietats de les lleis normals multidimensionals més rellevants. Acabem el capítol amb la versió del Teorema central del límit multidimensional.

Teorema 5.4.3 (Teorema central del límit multidimensional). *Sigui $\{X_n\}_{n\in\mathbb{N}}$ una successió de vectors aleatoris k-dimensionals, independents i identicament distribuïts. Sigui $S_n = X_1 + ... + X_n$, suposem que les components de X_1 són de quadrat integrable i posem $E[X_1] = m$ i $E[(X_1 - m)(X_1 - m)^*] = \Lambda$ aleshores,*

$$\frac{S_n - nm}{\sqrt{n}} \xrightarrow[n\to\infty]{\mathcal{L}} N(0, \Lambda)$$

Demostració. Definim $Y_n := \frac{S_n - nm}{\sqrt{n}}$ i fixem $t \in \mathbb{R}^k$. Aleshores $\{t^*X_n\}_{n\in\mathbb{N}}$ és una successió de variables aleatòries independents, idènticament distribuïdes amb mitjana t^*m i variància

$$\text{Var}[t^*X_1] = \sigma^2(t^*X_1) = E[t^*(X_1 - m)(X_1 - m)^*t] = t^*\Lambda t.$$

El Teorema central del límit (Teorema 5.4.1) ens diu que,

$$t^*Y_n = \frac{1}{\sqrt{n}}\left[\sum_{j=1}^{n} t^*X_j - nt^*m\right] \xrightarrow[n\to\infty]{\mathcal{L}} N(0, t^*\Lambda t).$$

Per tant,

$$\varphi_{Y_n}(t) = E[e^{it^*Y_n}] = \varphi_{t^*Y_n}(1) \xrightarrow[n\to\infty]{\mathcal{L}} \varphi_{N(0,t^*\Lambda t)}(1) = e^{\frac{-1}{2}t^*\Lambda t} = \varphi_{N(0,\Lambda)}(t).$$

Tenim que la funció característica de Y_n convergeix cap a la funció característica de la llei $N(0,\Lambda)$ en tot punt $t \in \mathbb{R}^k$. La versió multidimensional del teorema central del límit de Lévy implica aleshores que $Y_n = \frac{S_n - nm}{\sqrt{n}} \xrightarrow[n\to\infty]{\mathcal{L}} N(0,\Lambda)$, i el teorema queda provat. $\square$

Conclusió

A nivell personal, la probabilitat és una de les branques de les matemàtiques que més m'ha cridat l'atenció durant la carrera. Crec que es tracta d'un domini molt interessant i aplicable en molts sectors, ja que, amb l'aleatorietat present en gairebé tots els àmbits, la probabilitat et dóna les eines per analitzar, interpretar i fer prediccions que completen la informació i enriqueixen moltíssim la presa de decisions.

Per diferents motius, aquest últim any no vaig poder cursar l'optativa de probabilitats avançades i em vaig quedar amb ganes d'avançar en aquesta àrea. És per això que a l'hora d'escollir el tema del treball de final de grau, vaig decidir precisament això: aprofundir més en la teoria de la probabilitat. Inicialment doncs, no tenia clar on havia d'arribar, en el sentit que aquest treball no tenia una direcció clara més enllà de la d'aprofundir en aquest domini. Podríem dir que, des d'un inici, l'objectiu d'aquest treball ha sigut anar construïnt de manera gradual i per passos, resultats fonamentals. Com que les demostracions de molts teoremes principals s'aconsegueixen a través d'introduïr definicions i resultats amb interès en si mateix, l'evolució del treball ha sigut agraïda i essencialment natural.

Barcelona, Juny 2022.

Apèndix A

Dediquem aquest afegit a presentar alguns resultats que hem fet servir al llarg del treball i algunes ampliacions. En particular veurem definicions destacables de la convergència de variables aletòries, de l'esperança condicionada i algunes propietats de la funció característica i de les lleis multidimensionals.

A.1 Convergència de successions de variables aleatòries

Comencem per alguns desigualtats que hem usat en diferents raonaments del treball i recordarem al final de l'apartat algunes convergències de variables aleatòries. En general no farem les demostracions perquè estan fetes a l'assignatura de Probabilitats.

A.1.1 Desigualtat de Txebixev i Desigualtat de Schwarz

Proposició A.1.1 (Desigualtat de Markov). *Sigui $(\Omega, \mathcal{A}, P)$ un espai de probabilitat, X una variable aleatòria no negativa amb esperança finita. Per $a \in \mathbb{R}^+$ es verifica,*

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

Proposició A.1.2 (Desigualtat de Txebixev). *Sigui $(\Omega, \mathcal{A}, P)$ un espai de probabilitat, X una variable aleatòria no negativa i $f : \mathbb{R} \longrightarrow [0, +\infty]$ una funció creixent tal que $E(f(X)) < \infty$. Sigui $a \in \mathbb{R}$ tal que $0 < f(a) < +\infty$, aleshores es compleix,*

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

Un cas particular de la Desigualtat de Txebixev és $f(a) = x^k$, per $x \in \mathbb{R}^+$ i $k \in \mathbb{N}$. Agafant valor absolut d'una variable aleatòria i $a > 0$ tenim [1],

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

Proposició A.1.3 (Desigualtat de Jensen). *Sigui $(\Omega, \mathcal{A}, P)$ un espai de probabilitat, X una variable aleatòria i $f : \mathbb{R} \longrightarrow \mathbb{R}$ una funció real convexa [2] tal que es existeixen $E(X)$ i $E[f(X)]$, aleshores es verifica la següent desigualtat:*

$$f(E(X)) \leq E(f(X)).$$

[1] el cas $k = 2$, és com a vegades es formula la Desigualtat de Txebixev

[2] Una funció real f definida en un interval $[a, b] \subseteq \mathbb{R}$ és una funció convexa si per tot $x_1, x_2 \in [a, b]$ compleix la desigualtat: $f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2)$, per tor $t \in [0, 1]$

Lema A.1.4 (Desigualtat de Schwarz). *Si $X, Y : \Omega \longrightarrow \mathbb{R}$ són funcions mesurables tals que X^2 i Y^2 són integrables, aleshores XY és integrable i*

$$\left(\int_\Omega XY d\mu\right)^2 \leq \left(\int_\Omega X^2 d\mu\right) \cdot \left(\int_\Omega Y^2 d\mu\right).$$

Demostració. Tenim les següents desigualtats:

$$\begin{aligned} 0 \leq (X - Y)^2 \quad &\text{pel que} \quad XY \leq \frac{1}{2}(X^2 + Y^2), \\ 0 \leq (X + Y)^2 \quad &\text{pel que} \quad XY \geq -\frac{1}{2}(X^2 + Y^2). \end{aligned}$$

si considerem les dues expressions obtenim $|XY| \leq \frac{1}{2}(X^2 + Y^2)$ i per de ser X^2 i Y^2 integrables obtenim que XY també ho és.
Per veure la desigualtat suposem el cas no trivial $\int_\Omega Y^2 d\mu \neq 0$ (notem que si $\int_\Omega Y^2 d\mu = 0$ aleshores $Y = 0$ q.p.t i les dues bandes de l'expressió són zero). Per qualsevol nombre real γ tenim $(X + \gamma Y)^2 \geq 0$, pel que:

$$\int_\Omega (X + \gamma Y)^2 d\mu = \int_\Omega X^2 d\mu + 2\gamma \int_\Omega XY d\mu + \gamma^2 \int_\Omega Y^2 d\mu \geq 0.$$

Prenem,

$$\gamma = \frac{-\int_\Omega XY d\mu}{\int_\Omega Y^2 d\mu},$$

si subsituïm per aquest valor l'expressió anterior i multipliquem per $\int_\Omega Y^2 d\mu$ s'obté la desigualtat. □

A.1.2 Lemes de Borel-Cantelli

Passem a veure els Lemes de Borel-Cantelli.

Proposició A.1.5 (Primer lema de Borel-Cantelli). *Sigui $\{A_n\}_{n\in\mathbb{N}}$ una successió qualsevol d'esdeveniments. Aleshores es té la implicació següent:*

$$\sum_{i=1}^{\infty} P(A_n) < +\infty, \quad \text{aleshores}, \quad P(\lim_{n\to\infty} \sup A_n) = 0.$$

Demostració.

$$P(\lim_{n\to\infty} \sup A_n) = P(\bigcap_{n=1}^{\infty} \bigcap_{m=n}^{\infty} A_m) = \lim_{n\to\infty} P(\bigcap_{m=n}^{\infty} A_m) \leq \lim_{n\to\infty} \sum_{m=n}^{\infty} P(A_m) = 0$$

On hem usat la desigualtat seqüencial i la subadditivitat. □

També podem expressar el Primer Lema de Borel en termes del complementari, és a dir, amb les mateixes hipòtesis, es compleix la següent implicació:

$$\sum_{i=1}^{\infty} P(A_n) < +\infty, \quad \text{aleshores}, \quad P(\lim_{n\to\infty} \inf A_n^c) = 1.$$

Tant la desigualtat de Txebitxev com el primer lema de Borel-Cantelli valen per a mesures qualsevols, no necessàriament probabilitats.

Proposició A.1.6 (Segon lema de Borel-Cantelli). *Sigui* $\{A_n\}_{n\in\mathbb{N}}$ *una successió d'esdeveniments independents. Aleshores es té la implicació següent:*

$$\sum_{i=1}^{\infty} P(A_n) = +\infty, \quad \text{aleshores}, \quad P(\lim_{n\to\infty}\sup A_n) = 1.$$

Demostració. Tenim que $P(\lim_{n\to\infty}\sup A_n) = 1$ si, i només si, $P(\cap_{n=1}^{\infty}\cap_{k=n}^{\infty} A_k^c) = 0$, i per tant, és suficient veure que $P(\cap_{k=n}^{\infty} A_k^c) = 0$ per a tot n.

$$P(\bigcap_{k=n}^{m+n} A_k^c) = \prod_{k=n}^{m+n}(1 - P(A_k)) \leq \prod_{k=n}^{m+n} e^{-P(A_k)} = \exp\Big\{ - \sum_{k=n}^{m+n} P(A_k)\Big\}.$$

On hem usat la independència dels esdeveniments i la desigualtat $1-x \leq e^{-x}$. Fent tendir $m \to \infty$, com que $P(\cap_{k=n}^{\infty} A_k^c) = \lim_{m} P(\cap_{k=n}^{m+n} A_k^c) = 0$. □

A.1.3 Convergència de variables aleatòries

Suposarem a partir d'ara que les variables aleatòries X amb què tractarem estan totes definides en un mateix espai $(\Omega, \mathcal{A}, P)$. Quan parlem de successions de variables aleatòries hi ha diverses definicions naturals de convergència a considerar, en particular:

Definició A.1.7. *Direm que una successió de variables aleatòries* $\{X_n\}_{n\to\infty}$ *convergeix quasi segurament a una variable* X *si, i només si, existeix un conjunt* $N \in \mathcal{A}$ *de probabilitat zero tal que per tot* $\omega \notin N$,

$$\lim_{n\to\infty} X_n(\omega) = X(\omega).$$

Ho denotarem com $X_n \xrightarrow[n\to\infty]{q.s} X$.

Definició A.1.8. *Direm que una successió de variables aleatòries* $\{X_n\}_{n\to\infty}$ *convergeix en probabilitat a una variable* X *si, i només si, per tot* $\varepsilon > 0$ *es compleix,*

$$\lim_{n\to\infty} P(\{|X_n - X| > \varepsilon\}) = 0.$$

Ho denotarem com $X_n \xrightarrow[n\to\infty]{P} X$.

Definició A.1.9. *Direm que una successió* $\{X_n\}_{n\to\infty}$ *de variables aleatòries amb moment d'ordre* $p \geq 1$ *finit convergeix en* L^p *a una variable* $X \in L^p$ *si, i només si, es compleix,*

$$\lim_{n\to\infty} E[|X_n - X|^p] = 0.$$

Ho denotarem com $X_n \xrightarrow[n\to\infty]{L^p} X$.

A.2 Esperança Condicionada

De manera genèrica, l'esperança condicionada d'una variable aleatòria és el valor respecte d'una probabilitat modificada que inclou una informació *a priori*. El cas més simple correspon a condicionar respecte d'un esdeveniment $B \in \mathcal{A}$. En aquest cas, l'esperança condicionada és l'esperança matemàtica calculada a l'espai de probabilitat modificat $(\Omega, \mathcal{A}, P(\cdot|B))$. Veurem que en general, informació addicional no pot descriure's tant fàcilment.

En aquest apartat considerarem $\mathcal{G}$ una σ-àlgebra arbitrària inclosa en $\mathcal{A}$ i X una variable amb esperança finita. Recordem que donada una variable aleatòria X, l'esperança $E(X)$, està ben definida si X és una variable positiva (pot ser infinita) o bé si és integrable, és a dir, si $E(|X|) < \infty$. Suposarem en tota la secció que X és una variable integrable.

L'objectiu d'aquest apartat és donar una definició de l'esperança condicionada de X respecte de $\mathcal{G}$.

Definició A.2.1. *Sigui $B \in \mathcal{A}$ tal que $P(B) > 0$. L'esperança condicionada de X sobre B és el nombre real següent:*

$$E(X|B) = \frac{E(\mathbb{1}_B \cdot X)}{P(B)}. \tag{A.2.1}$$

És fàcil veure que,

- $E(X|\Omega) = E(X)$.
- $E(\mathbb{1}_A|B) = P(A|B)$, per tot $A \in \mathcal{A}$ [3].

Notem que l'esperança condicionada respecte d'un esdeveniment B, $E(X|B)$, coincideix amb l'esperança de la probabilitat condicionada $P(\cdot|B) := P_B$. Veiem-ho per una variable aleatòria discreta: Sigui $X = \sum_{i=1}^{\infty} x_i \cdot \mathbb{1}_{A_i}$ amb $A_i = \{X = x_i\}$ partició de Ω i complint que $P(A_i) > 0$. En efecte,

$$\begin{aligned} E_{P_B}(X) &= \sum_{i=1}^{\infty} x_i P_B(A_i) = \sum_{i=1}^{\infty} x_i \frac{P(A_i \cap B)}{P(B)} = \frac{1}{P(B)} \sum_{i=1}^{\infty} x_i E(\mathbb{1}_{A_i \cap B}) \\ &= \frac{1}{P(B)} E(\sum_{i=1}^{\infty} x_i \mathbb{1}_{A_i} \cdot \mathbb{1}_B) = \frac{E(\mathbb{1}_B \cdot X)}{P(B)} = E(X|B). \end{aligned}$$

Passem a veure l'esperança condicionada d'una variable.

Definició A.2.2. *Sigui Y una variable discreta, $Y = \sum_{i=1}^{\infty} y_i \mathbb{1}_{B_i}$, amb $B_i = \{Y = y_i\}$ partició de Ω i complint que $P(B_i) > 0$. L'esperança condicionada de X per Y, és la variable aleatòria,*

$$E(X|Y) = \sum_{i=1}^{\infty} E(X|Y = y_i) \mathbb{1}_{B_i} = \sum_{i=1}^{\infty} E(X|B_i) \mathbb{1}_{B_i}.$$

La definició ens diu que Y i $E(X|Y)$ són dues variables aleatòries discretes i constants sobre els mateixos conjunts B_i. Denotem per $\sigma(Y)$ la σ-àlgebra generada per la variable Y [4]. Es compleixen les següents propietats:

(i) $E(X|Y)$ és mesurable respecte $\sigma(Y)$.

(ii) Si les variables aleatòries X i Y són independents, tenim $E(X|Y) = E(X)$. És a dir, $E(X|Y)$ és una variable constant igual a la seva esperança.

(iii) Per a tot $B \in \sigma(Y)$ es té $E\big[E(X|Y) \cdot \mathbb{1}_B\big] = E(X \cdot \mathbb{1}_B)$. En particular, escollint $B = \Omega \in \sigma(Y)$ tenim $E[E(X|Y)] = E(X)$.

En la definició d'esperança condicionada respecte d'una variable discreta Y, l'important no són els valors que pren la variable, y_i, sino els conjunts on pren valors constants, $B_i = \{Y = y_i\}$.

Passem a veure la definició de l'esperança condicionada respecte d'una σ-àlgebra $\mathcal{G}$ qualsevol.

---

[3] $E(\mathbb{1}_A) = P(A)$

[4] generada per la partició formada pels $B_i = \{Y = y_i\}$

Teorema A.2.3. *Donada una σ-àlgebra $\mathcal{G} \subseteq \mathcal{A}$ i una variable aleatòria integrable X, existeix una única variable aleatòria definida quasi segurament que denotem Z tal que,*

(i) és mesurable respecte $\mathcal{G}$.

(ii) compleix que per qualsevol conjunt $G \in \mathcal{G}$, $E(Z \cdot \mathbb{1}_G) = E(X \cdot \mathbb{1}_G)$.

La variable aleatòria introduïda al teorema anterior és l'esperança condicionada de X respecte $\mathcal{G}$. L'existència i unicitat de Z, que denotarem $Z := E(X|\mathcal{G})$ està assegurada per el Teorema de Radon-Nikodym (tret d'un conjunt de probabilitat zero). No farem la interpretació en general d'aquesta variable, descriure'm com calcular-la en dues situacions particulars:

(a) Sigui $\mathcal{G}$ una σ-àlgebra generada per una partició finita o numerable $\{A_i\}_{i\in I}$. Aleshores,

$$E(X|\mathcal{G}) = \sum_{i\in I} E(X|A_i)\mathbb{1}_{A_i} = \sum_{i\in I} \frac{E(X \cdot \mathbb{1}_{A_i})}{P(A_i)}\mathbb{1}_{A_i}.$$

(b) Sigui $\mathcal{G}$ una σ-àlgebra generada per les variables aleatòries $Y_1, ..., Y_n$, [5]. Suposant que la distribució conjunta del vector aleatori $(X, Y_1, ..., Y_n)$ té densitat f, aleshores:

$$E(X|Y_1, ..., Y_n) = \int_{-\infty}^{\infty} xf(x|y_1, ..., y_n)dx, \tag{A.2.2}$$

A la igualtat (A.2.2) primer calculem l'esperança condicionada $E(X|Y_1 = y_1, ..., Y_n = y_n)$ i finalment, substituïm els valors reals $y_1, ..., y_n$ de les variables aleatòries $Y_1, ..., Y_n$.

Per comprovar que les definicions són consistents cal veure que compleixen les dues condicions del Teorema A.2.3. Ho veurem únicament per (a), on $Z := \sum_{i\in I} E(X|A_i)\mathbb{1}_{A_i}$.

(i) Veure que $\sum_{i\in I} E(X|A_i)\mathbb{1}_{A_i}$ és mesurable respecte $\mathcal{G}$ és equivalent a veure que ho és $\sum_{i\in I} E(X \cdot \mathbb{1}_{A_i})\mathbb{1}_{A_i}$; com que els conjunts A_i són mesurables respecte $\mathcal{G}$, és compleix la primera condició.

(ii) Falta veure que per tot $G \in \mathcal{G}$, es compleix $E\big(\sum_{i\in I} E(X|A_i)\mathbb{1}_{A_i}\mathbb{1}_G\big) = E(X \cdot \mathbb{1}_G)$.

$$\begin{aligned} E\big(\sum_{i\in I} E(X|A_i)\mathbb{1}_{A_i}\mathbb{1}_G\big) &= E\Big(\sum_{i\in I} \frac{E(X \cdot \mathbb{1}_{A_i})}{P(A_i)}\mathbb{1}_{A_i}\mathbb{1}_G\Big) = \sum_{i\in I} \frac{E(X \cdot \mathbb{1}_{A_i})}{P(A_i)}P(A_i \cap G) \\ &= \sum_{j\in J} E(X \cdot \mathbb{1}_{A_j}) = E(X \cdot \mathbb{1}_G), \end{aligned}$$

on hem usat que per ser G mesurable, $A_i \cap G$ és el buit o és A_i, i podem escriure $G = \cup_{j\in J} A_j$.

Passem a indicar diferents propietats de l'esperança de X condicionada a una σ-àlgebra. Considrem X, Y variables aleatòries integrables i $a, b \in \mathbb{R}$, es compleix:

(i) Linealitat: $E(aX + bY|\mathcal{G}) = aE(X|\mathcal{G}) + bE(Y|\mathcal{G})$.

(ii) Monotonia: Si $X \leq Y$ aleshores $E(X|\mathcal{G}) \leq E(Y|\mathcal{G})$.

[5] és a dir la σ-àlgebra està generada per els esdeveniments de la forma $Y_1^{-1}(B_1), ..., Y_n^{-1}(B_n)$ amb $B_1, ...B_n$ conjunts de la σ-àlgebra de Borel B

(iii) Si X és una variable $\mathcal{G}$-mesurable, aleshores $E(X|\mathcal{G}) = X$.

(iv) Si X és una variable aleatòria independent de $\mathcal{G}$, és a dir que qualsevol conjunt de la forma $X^{-1}(B) \in \mathcal{B}$ és independent de $\mathcal{G}$. Aleshores, $E(X|\mathcal{G}) = X$.

(v) Factorització: Si Y és una variable acotada i $\mathcal{G}$-mesurable, aleshores:

$$E(Y \cdot X|\mathcal{G}) = YE(X|\mathcal{G}).$$

(vi) Si $\mathcal{G}_i$ per $i = 1, 2$, són σ-àlgebres tals que $\mathcal{G}_1 \subseteq \mathcal{G}_2$, aleshores,

$$E[E(X|\mathcal{G}_1)|\mathcal{G}_2] = E[E(X|\mathcal{G}_2)|\mathcal{G}_1] = E(X|\mathcal{G}_1).$$

(vii) Si X és una variable aleatòria independent de $\mathcal{G}$, i Z una variable aleatòria $\mathcal{G}$-mesurable es compleix que per qualsevol funció mesurable $h(x, z)$, la variable aleatòria $h(X, Z)$ té esperança finita i es compleix,

$$E[h(X, Z)|\mathcal{G}] = E[h(X, z)]\Big|_{z=Z}.$$

A.3 Propietats fonamentals de les funcions característiques

Proposició A.3.1 (**Propietats fonamentals de les funcions característiques**). *Sigui μ una probabilitat en $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$. La funció característica de μ compleix les propietats següents:*

(i) $\varphi_\mu(0) = 1$.

(ii) $|\varphi_\mu(t)| \leq 1$ *per tot* $t \in \mathbb{R}^n$.

(iii) $\varphi_\mu(-t) = \overline{\varphi_\mu(t)}$ *per tot* $t \in \mathbb{R}^n$.

(iv) φ_μ *és una funció uniformement contínua.*

(v) Sigui X un vector aleatori n-dimensional, A una matriu $m \times n$ i $b \in \mathbb{R}^m$. Aleshores per tot $t \in \mathbb{R}^m$,

$$\varphi_{AX+b}(t) = e^{i<t,b>}\varphi_X(A^*t).$$

(vi) Propietat fonamental d'injectivitat: Si μ_1, μ_2 són dues probabilitats en $\mathbb{R}^n$ tals que $\varphi_{\mu_1} = \varphi_{\mu_2}$, necessàriament $\mu_1 = \mu_2$.

(vii) Direm que la probabilitat μ sobre $\mathbb{R}^n$ és simètrica si $\mu(B) = \mu(-B)$ per a tot borelià $B \in \mathcal{B}(\mathbb{R}^n)$. La propietat de simetria és equivalent a que la seva funció característica φ_μ sigui real.

(viii) Sigui $X = (X_1, ..., X_n)$ un vector aleatori. Les variables aleatòries $X_1, ..., X_n$ són independents si i només si

$$\varphi_X(t_1, ..., t_n) = \varphi_{X_1}(t_1) \times \cdots \times \varphi_{X_n}(t_n).$$

(ix) Si $X_1, ..., X_n$ són variables aleatòries independents es té,

$$\varphi_{X_1+...+X_n}(t) = \varphi_{X_1}(t) \times \cdots \times \varphi_{X_n}(t).$$

Demostració. Donarem alguna indicació de la demostració de les diferents propietats. Per més detalls consultar el llibre [1, Proposició 5.2.11] i el llibre [6, Proposició 9.3].

(i) $\varphi_\mu(0) = \int_{\mathbb{R}^n} e^{i<0,x>}\mu(dx) = \int_{\mathbb{R}^n} 1\mu(dx) = 1.$

(ii) $|\varphi_\mu(t)| = \Big|\int_{\mathbb{R}^n} e^{i<t,x>}\mu(dx)\Big| \leq \int_{\mathbb{R}^n} |e^{i<t,x>}|\mu(dx) = \int_{\mathbb{R}^n} 1\mu(dx) = 1.$

(iii) $\varphi_\mu(-t) = \int_{\mathbb{R}^n} e^{-i<t,x>}\mu(dx) = \int_{\mathbb{R}^n} \overline{e^{i<t,x>}}\mu(dx) = \overline{\int_{\mathbb{R}^n} e^{i<t,x>}\mu(dx)} = \overline{\varphi_\mu(t)}.$

(iv) Considerem $s, t \in \mathbb{R}^n$,

$$\begin{aligned}|\varphi_\mu(t) - \varphi_\mu(s)| &= \Big|\int_{\mathbb{R}^n} \big(e^{i<t,x>} - e^{i<s,x>}\big)\mu(dx)\Big| \leq \int_{\mathbb{R}^n} \big|e^{i<t,x>} - e^{i<s,x>}\big|\mu(dx)\\ &\leq \int_{\mathbb{R}^n} \big|e^{i<s,x>}\big| \cdot \big|(e^{i<t-s,x>} - 1)\big|\mu(dx) = \int_{\mathbb{R}^n} \big|e^{i<t-s,x>} - 1\big|\mu(dx).\end{aligned}$$

Pel Teorema de la convergència dominada (Teorema 2.2.16), com que el límit puntual de la integral és zero i $|e^{i<t-s,x>} - 1| \leq 2$, tenim que:

$$\int_{\mathbb{R}^n} |e^{i<t-s,x>} - 1|\mu(dx) \xrightarrow[|t-s|\to 0]{} 0.$$

(v) $\varphi_{AX+b}(t) = E\big(e^{i<t,AX+b>}\big) = e^{i<t,b>}E\big(e^{i(A^*t)^*X}\big) = e^{i<t,b>}\varphi_X(A^*t).$

(vi) Aquest resultat es pot deduir de la fórmula d'inversió (Teorema 5.2.5) que veurem més endevant [6].

(vii) Considerem l'aplicació $\tau : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ definida per $\tau(x) = -x$. Pel Teorema de la mesura imatge (Teorema 2.2.19), per a tota funció $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ mesurable i acotada es compleix,

$$\int_{\mathbb{R}^n} f(-x)\mu(dx) = \int_{\mathbb{R}^n} (f \circ \tau)(x)\mu(dx) = \int_{\mathbb{R}^n} f(x)(\mu \circ \tau^{-1})(dx).$$

Aleshores,

$$\overline{\varphi_\mu(t)} = \varphi_\mu(-t) = \int_{\mathbb{R}^n} e^{-i<t,x>}\mu(dx) = \int_{\mathbb{R}^n} e^{i<t,x>}(\mu \circ \tau^{-1})(dx) = \varphi_{\mu\circ\tau^{-1}}(t).$$

Tenint en compte la propietat (vi), tindrem que μ és simètrica si i només si $\mu = \mu \circ \tau^{-1}$, i això equival a $\varphi_\mu = \varphi_{\mu\circ\tau^{-1}} = \overline{\varphi_\mu}$, és a dir, que φ_μ sigui real.

(viii) Sigui $X = (X_1, ..., X_n)$ un vector aleatori. Sabem que les variables $X_1, ..., X_n$ són independents si i només si $(P \circ X^{-1}) = (P \circ X_1^{-1}) \times \cdots \times (P \circ X_n^{-1})$ i, per la propietat (vi) això equival a la igualtat de les funcions característiques d'aquestes probabilitats. La funció característica de $(P \circ X^{-1})$ és φ_X i la de $(P \circ X_1^{-1}) \times \cdots \times (P \circ X_n^{-1})$ és,

$$\begin{aligned}&\int_{\mathbb{R}^n} e^{i<t,x>}(P \circ X_1^{-1})(dx_1) \times \cdots \times (P \circ X_n^{-1})(dx_n) =\\ &\Big(\int_{\mathbb{R}} e^{it_1X_1}(P \circ X_1^{-1})(dx_1)\Big) \times \cdots \times \Big(\int_{\mathbb{R}} e^{it_nX_n}(P \circ X_n^{-1})(dx_n)\Big) =\\ &\varphi_{X_1}(t_1) \times \cdots \times \varphi_{X_n}(t_n).\end{aligned}$$

(ix) $\varphi_{X_1+...+X_n}(t) = E[e^{it(X_1+...+X_n)}] = E[e^{itX_1}] \times \cdots \times E[e^{itX_n}] = \varphi_{X_1}(t) \times \cdots \times \varphi_{X_n}(t).$

$\square$

[6] veure Obsevacio 5.2.7

A.4 Propietats de les lleis normals multidimensionals

Enunciem de manera esquematitzada algunes de les propietats de les lleis normals multidimensionals més rellevants.

(i) Sigui X un vector amb llei $N(m,\Lambda)$. Si C és una matriu ortogonal que compleix $\Lambda = C^*DC$ [7], aleshores el vector $Y := C(X-m)$ té les components independents i amb lleis $N(0,\lambda_i)$ si $\lambda_i \neq 0$, o bé zero si $\lambda_i = 0$. En efecte,

$$\varphi_Y(t) = e^{it^*(-Cm)}\varphi_X(C^*t) = \exp\big(-it^*Cm + it^*Cm - \frac{1}{2}t^*C\Lambda C^*t\big) = \exp\big(\frac{-1}{2}t^*Dt\big).$$

(ii) Si X és un vector aleatori amb llei $N(m,\Lambda)$, aleshores m respresenta el vector de mitjanes i Λ respresenta la matriu de variàncies i covariàncies de X. Usant les notacions d'abans,

$$\begin{aligned} &E(X) = C^*E(Y) + m = m, \\ &E[(X-m)(X-m)^*] = E[C^*YY^*C] = C^*DC = \Lambda. \end{aligned}$$

(iii) Sigui X un vector aleatori n-dimensional amb llei $N(m,\Lambda)$. Aleshores, si B és una matriu d'ordre $r \times n$, el vector BX té llei $N(Bm, B\Lambda B^*)$. Veiem-ho,

$$\begin{aligned} \varphi_{BX}(t) = \varphi_X(B^*X) &= \exp\big(it^*Bm - \frac{1}{2}(B^*t)^*\Lambda(B^*t)\big) \\ &= \exp\big(it^*(Bm) - \frac{1}{2}(B\Lambda B^*)t\big), \quad \text{per a tot} \quad t \in \mathbb{R}^r. \end{aligned}$$

En particular tot vector aleatori de la forma $(X_{i_1}, ..., X_{i_m})$, amb $m \leq n$, té llei normal, i tota combinació lineal $\sum_{i=1}^n a_iX_i$ té llei normal. Recíprocament, si $X = (X_1, ..., X_n)$ és un vector alatori tal que tota combinació lineal $\sum_{i=1}^n a_iX_i$ és normal, alshores X té una llei normal multidimensional. Veiem es compleix,

$$E(e^{it^*X}) = \varphi_{t^*X}(1) = e^{-\frac{1}{2}\sigma^2(t^*X) + iE(t^*X)},$$

tindrem que es compeix $E(t^*X) = t^*E(X)$ i, per tant,

$$\begin{aligned} \sigma^2(t^*X) &= E[\big(t^*(X - E(X))\big)^2] \\ &= E[t^*(X-E(X))(X-E(X))^*t] = t^*\Lambda t, \end{aligned}$$

on Λ és la matriu de variàncies i covariàncies del vector X. Per tant, X té una llei normal multidimensional amb paràmetres $m = E(X)$ i Λ.

(iv) Sigui $X = (X_1, ..., X_n)$ un vector amb llei normal n-dimensional $N(m,\Lambda)$. La independència de les variables $X_1, ..., X_n$ és equivalent a que la matriu Λ sigui diagonal, és a dir a que les variables siguin incorrelacionades $cov(X_i, X_j) = 0$ per a tot $i \neq j$ amb $i, j \in \{1, ..., n\}$. Recíprocament, si la matriu Λ és diagonal i designem per $\lambda_1, ..., \lambda_n$ els elements de la diagonal, tindrem que les variables $X_1, ..., X_n$ són independents:

$$\varphi_X(t) = e^{it^*m - \frac{1}{2}t^*\Lambda t} = \prod_{j=1}^n e^{it_jm_j - \frac{1}{2}\lambda_j t_j^2} = \prod_{j=1}^n \varphi_{X_j}(t_j).$$

[7] D és la matriu de la Proposició 5.4.2 que té $\lambda_1, ..., \lambda_n$ com a elements de la diagonal, que coincideixen amb els valors propis de A

Bibliografia

[1] Alabert, A. *Mesura i Probabilitat.* Primera edició. Barcelona: Servei de Publicacions de la UAB, 1996.

[2] Amir, D. *Probability Theory.* Department of Mathematics, Stanford University, Stanford, CA 94305.

[3] Billingsley, Patrick. *Probability and Measure.* Tercera Edició. New York: John Wiley & Sons, 1995. ISBN 0-471-00710-2.

[4] Bogachev, V. *Measure Theory.* Volume I i II. Berlin: Springer, 2000. ISBN 3-540-34513-2.

[5] Feller, William. *An Introduction to Probability Theory and Its Applications.* Volume II. Second Edition. New York: John Wiley & Sons, 1965. ISBN 0-471-25708-7.

[6] Garcia, A. *Teorias de la Medida y de la Probabilidad.* Cáceres: Universidad de Extremadura Servicio de Publicaciones, 2008. Colección manuales uex-57. ISBN 978-84-691-6411-2.

[7] Jacod, J. Protter, P. *Probability Essentials.* Second Edition. Heidelberg: Die Deutsche Bibliothek, 2004. ISBN 978-3-540-43871-7.

[8] Lévy, P. *Calcul des probabilités.* Paris: Gauthier-Villars, 1925.

[9] De Moivre, A. *The Doctrine of Chances.* Londres: A Millar, 1756.

[10] Nualart, D. Sanz, M. *Curs de Probabilitats.* Primera edició. Barcelona: Promociones y Publicaciones Universitarias, 1990. ISBN 84-7665-718-8.

[11] Ore, O. *Pascal and the invention of probability theory*, Rhode Island: Amer. Math, 1960. Monthly 67.

[12] Ross, S. *A first course in probability.* Eight edition. New Jersey: Pearson, 2010. ISBN 978-0-13-603313-4

[13] Kolmogorov, A. N. *Sulla Determinazione Empirica di Una Legge di Distribuzione.* Giornale dell'Istituto Italiano degli Attuari, 1933.