PROBABILITATS

Marta Sanz Solé
Facultat de Matemàtiques
Universitat de Barcelona

Al Roger, tot evocant el plaer de l'aprenentatge.

Índex

Índex de figures

Introducció

La teoria de la probabilitat ha estat, des de l'inici del seu desenvolupament, molt relacionada amb aplicacions de la matemàtica. Citem, per exemple, problemes d'assegurances que apareixen a partir del segle XIV amb l'establiment de companyies d'assegurances navals a Itàlia i Holanda, dissenys de loteries, particularment popularitzades al segle XVIII, taules de mortalitat i altres elements per a l'estudi de qüestions demogràfiques, etc. Més recentment la probabilitat és utilitzada com a ingredient valuós de la modelització matemàtica en problemes de la física, l'economia financera, l'evolució de poblacions, entre d'altres. Actualment, en el públic no especialitzat, la popularitat de la probabilitat prové de la seva gran incidència en l'estadística. De fet, els models de l'estadística matemàtica són models probabilístics i bona part de les tècniques emprades en estadística provenen de la probabilitat.

És per això que la teoria de la probabilitat pot resultar atractiva a un públic amb interessos ben diversos. El científic, o futur científic, motivat per la matemàtica, hi pot trobar l'al·licient de la modelització de l'atzar amb tècniques pròpies i d'altres branques de la matemàtica que van des d'un nivell elemental a un de molt sofisticat. Al tecnòleg l'ajudarà en l'elecció i la comprensió de models i marcs d'estudi que responen de manera força acurada a la realitat. L'usuari de l'estadística podrà entendre quins són els mecanismes que el porten a prendre decisions davant de la incertesa. I, en fi, qualsevol persona amb curiositat intel·lectual podrà intentar resoldre problemes atractius de jocs i d'estratègia.

Aquest llibre va adreçat de manera prioritària a estudiants de la llicenciatura en matemàtiques. Els temes que s'hi desenvolupen abasten uns continguts bàsics en teoria de la probabilitat que haurien de ser adquirits per l'estudiant en el primer cicle de la llicenciatura. Per coherència amb aquest objectiu, he defugit conscientment utilitzar elements avançats de la teoria de la mesura, tot ubicant aquest text en un nivell no especialitzat. Sóc plenament conscient dels riscs que això comporta, especialment pel que fa al rigor, concepte que, en el col·lectiu dels matemàtics, forma part d'una ètica irrenunciable. He intentat, doncs, un equilibri entre la necessitat de comunicar ingredients rellevants i àmpliament utilitzats de la matemàtica als joves estudiants i la de contribuir a la formació en l'ofici del matemàtic amb les característiques que tal aprenentatge porta implícites.

La presentació es fa seguint un nivell de dificultat progressiu, la qual cosa permet una lectura en dues etapes i estén sensiblement el públic destinatari. Per a un curs marcadament elemental es recomana incidir en els capítols 1–3.

L'elaboració d'aquest llibre ha estat possible gràcies a l'ajut concedit per la Universitat de Barcelona a *Projectes de formació del professorat, innovació docent i avaluació* de la convocatòria de 1996. El Dr. José Manuel Corcuera, amb la seva acurada lectura de versions prèvies a aquest manual, m'ha transmès moltes idees i comentaris que han contribuït a millorar-ne el resultat. En Roger Trias Sanz ha tingut cura eficient de la composició del text i els gràfics, suportant amb paciència tants de canvis i noves versions. A tots ells, el meu agraïment.

Capítol 1

El model probabilístic

En aquest capítol comencem introduint l'axiomàtica de Kolmogorov, que proporciona un model matemàtic per al tractament de l'atzar. En el cas particular del model finit, expliquem la relació entre combinatòria i càlcul de probabilitats. Introduïm també els conceptes bàsics de la independència i el condicionament des d'una perspectiva elemental. En capítols posteriors aquests conceptes seran tractats amb més profunditat.

La creació d'un model matemàtic per al tractament de l'atzar no ha estat un problema senzill. Com en altres àrees de la matemàtica, la seva fonamentació rigorosa s'ha produït quan, a nivell de resultats, la probabilitat ja havia adquirit un desenvolupament notable, encara que no exempt de certs traumatismes (vegeu més endavant les paradoxes de Bertrand).

L'axiomàtica de Kolmogorov, concretada en la noció d'*espai de probabilitat*, fou introduïda per aquest autor en una monografia publicada l'any 1933. Els conceptes matemàtics que s'utilitzen provenen de la teoria de la mesura, desenvolupada a partir de l'any 1900 per H. Lebesgue, M. Fréchet, E. Borel, J. Radon, C. Carathéodory, entre d'altres.

1.1 Fenòmens aleatoris

La teoria de la probabilitat analitza fenòmens dels quals, tot i conèixer les condicions de la seva realització, no és possible deduir-ne exactament el resultat. S'anomenen fenòmens aleatoris, en contraposició als deterministes, per als quals hi ha una relació clara causa-efecte. Els jocs d'atzar ens en proporcionen exemples típics: el llançament d'una moneda, l'extracció de boles d'una urna, etc. Altres problemes més lligats a la realitat, com un procés de producció o el pas d'automòbils per un peatge, tenen també un component de fenomen aleatori a causa de la manca de control sobre totes les causes possibles que poden afectar el resultat.

En els fenòmens aleatoris suposem, però, conegut el conjunt dels resultats possibles, així com també l'assignació d'una *versemblança* a cadascun d'aquests resultats. Vegem-ne uns quants exemples.

Exemple 1.1 Llancem una moneda tres vegades seguides. El conjunt de resultats possibles és $\Omega = \{(+{+}+), (+{+}C), (+C+), (C{+}+), (+CC), (CC+), (C+C), (CCC)\}$, on $+$ indica *creu* i C indica *cara*. Podem fer una assignació de $\frac{1}{8}$ a cadascun dels resultats, ja que res fa pensar que hi hagi resultats privilegiats.

Exemple 1.2 Llancem dues monedes idèntiques i observem el nombre de cares que surten. El conjunt de resultats possibles és $\Omega = \{0, 1, 2\}$. Aquí és raonable assignar als resultats 0 i 2 un pes de $\frac{1}{4}$, mentre que al 1 li assignem un pes de $\frac{1}{2}$.

Exemple 1.3 En el joc de *cara o creu*, ens pot interessar tenir informació sobre *la primera vegada que sortirà cara.* Aquí ens convindrà considerar un conjunt Ω que, a diferència dels exemples anteriors, no és finit. En efecte, Ω és l'espai de les successions $\{(a_n)_{n\geq 1}, a_n \in \{C, +\}\}$.

1.2 Espais de probabilitat

Un *espai de probabilitat* és un tern $(\Omega, \mathcal{A}, P)$.

Ω és un conjunt que correspon al dels resultats possibles de l'experiència aleatòria. S'anomena *espai mostral.*

$\mathcal{A}$ és una família de parts d'Ω que té estructura de σ-*àlgebra*; és a dir,

1. $\Omega \in \mathcal{A}$
2. $\mathcal{A}$ és estable per pas al complementari: Si $A \in \mathcal{A}$, també $A^c \in \mathcal{A}$.
3. $\mathcal{A}$ és estable per unions numerables: Si $\{A_n, n \geq 1\} \subset \mathcal{A}$, es compleix $\cup_{n\geq 1} A_n \in \mathcal{A}$.

La σ-àlgebra $\mathcal{A}$ serveix per descriure tots els *esdeveniments* possibles relacionats amb l'experiència aleatòria.

El darrer element de l'espai de probabilitat, P, s'anomena *probabilitat* i determina l'assignació de *versemblança* als esdeveniments. És una aplicació

$$P : \mathcal{A} \longrightarrow [0,1]$$

que té les propietats següents:

1. $P(\Omega) = 1$.
2. (σ-*additivitat*) Si $\{A_n, n \geq 1\}$ és una successió de conjunts de $\mathcal{A}$ disjunts dos a dos, aleshores

$$P(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n).$$

En el llenguatge de la teoria de la mesura, un espai de probabilitat és un espai de mesura finita, amb mesura total igual a 1.

Donem, tot seguit, una breu llista d'algunes de les equivalències més importants entre el llenguatge conjuntista i probabilista, basades en el model que acabem d'introduir.

Notacions	**Conjunts**	**Probabilitats**
ω	element de Ω	resultat possible
A	element de $\mathcal{A}$	esdeveniment
$A \cup B$	A reunió amb B	A o B
$A \cap B$	A intersecció amb B	A i B
A^c	complementari de A	no A
$\emptyset$	conjunt buit	esdeveniment impossible
Ω	conjunt total	esdeveniment segur

Observació 1.1 *Si el conjunt Ω és finit, que $\mathcal{A}$ sigui una σ-àlgebra és equivalent a dir que sigui una* àlgebra*(ço és, $\Omega \in \mathcal{A}$, és estable per pas al complementari i per reunions finites). En aquest cas, la σ-additivitat (propietat (2) de la probabilitat) és equivalent a l'*additivitat finita*: si A i B són conjunts de $\mathcal{A}$, i $A \cap B = \emptyset$, aleshores*

$$P(A \cup B) = P(A) + P(B).$$

1.2.1 Propietats de la probabilitat

Les propietats que demostrem a continuació són una conseqüència més o menys evident dels axiomes (1) i (2) de la probabilitat. Totes són molt útils en el càlcul de probabilitats d'esdeveniments.

1. $P(\emptyset) = 0$.

   En efecte, fixem $A \in \mathcal{A}$ i apliquem la propietat de σ-additivitat a la successió $A, \emptyset, \emptyset, \ldots$. Tindrem que

   $$P(A) = P(A) + \sum P(\emptyset).$$

   En conseqüència, $P(\emptyset) = 0$.

2. La propietat de σ-additivitat implica l'additivitat finita.

   En efecte, sigui $\{A_j, j = 1, \ldots, n\}$ una col·lecció finita de conjunts disjunts de $\mathcal{A}$. Clarament $\{A_1, \cdots, A_n, \emptyset, \cdots,\}$ és una successió de conjunts disjunts de $\mathcal{A}$. Per la σ-additivitat de P i la propietat 1 que acabem de demostrar, tenim que

   $$\begin{aligned} P\left(\bigcup_{j=1}^{n} A_j\right) &= P\Big(\big(\cup_{j=1}^{n} A_j\big) \cup \big(\cup \emptyset\big)\Big) \\ &= \sum_{j=1}^{n} P(A_j) + \sum P(\emptyset) = \sum_{j=1}^{n} P(A_j). \end{aligned}$$

3. Per a tot $A \in \mathcal{A}$, $P(A^c) = 1 - P(A)$.

   Aquesta igualtat és conseqüència de la descomposició $\Omega = A \cup A^c$ i la propietat d'additivitat finita.

4. Si $A, B \in \mathcal{A}$, $A \subset B$, aleshores $P(A) \le P(B)$.

   És una conseqüència de la desigualtat següent:

   $$P(B) = P\big(A \cup (B \cap A^c)\big) = P(A) + P(B \cap A^c) \ge P(A).$$

5. Siguin A i B de $\mathcal{A}$, aleshores,

$$P(A \cup B) + P(A \cap B) = P(A) + P(B).$$

En efecte, considerem la descomposició $A = (A \cap B) \cup (A \cap B^c)$. Per la propietat d'additivitat finita,

$$\begin{aligned} P(A) + P(B) &= P(A \cap B) + P(A \cap B^c) + P(B) \\ &= P(A \cap B) + P(A \cup B). \end{aligned}$$

6. Tota probabilitat P és *subadditiva*; és a dir,

$$P(\cup_{i=1}^n A_i) \leq \sum_{i=1}^n P(A_i),$$

per a qualsevol família de conjunts $\{A_i, i = 1, \ldots, n\}$ de $\mathcal{A}$.

La demostració d'aquesta propietat es fa per inducció sobre n. Per a $n = 2$ és una conseqüència immediata de la propietat anterior. Suposem la propietat certa per a qualsevol col·lecció de conjunts de $\mathcal{A}$, $\{A_1, \ldots, A_j\}$, amb $1 \leq j \leq n - 1$. Sigui $A_n \in \mathcal{A}$. Aleshores, si $A = \cup_{i=1}^{n-1} A_i$, tenim que

$$P\big(\cup_{i=1}^n A_i\big) = P\big(A \cup A_n\big) \leq P(A) + P(A_n) \leq \sum_{i=1}^n P(A_i).$$

7. Una generalització de l'argument de la demostració de la propietat 5 permet de donar una expressió exacta de $P(\cup_{i=1}^n A_i)$ per a qualsevol família de conjunts $\{A_i, i = 1, \ldots, n\}$ de $\mathcal{A}$. Més precisament, es compleix que

$$\begin{aligned} P(\cup_{i=1}^n A_i) &= \sum_{i=1}^n P(A_i) - \sum_{\substack{i<j \\ 1 \leq i,j \leq n}} P(A_i \cap A_j) \\ &+ \sum_{\substack{i<j<k \\ 1 \leq i,j,k \leq n}} P(A_i \cap A_j \cap A_k) - \cdots + (-1)^{n-1} P(\cap_{i=1}^n A_i). \end{aligned} \tag{1.1}$$

La demostració de (1.1) es fa per inducció sobre el nombre de conjunts n. Per a $n = 2$ és la propietat 5. Suposem (1.1) certa per a $n-1$ esdeveniments, és a dir,

$$\begin{aligned} P(\cup_{i=1}^{n-1} A_i) = \sum_{i=1}^{n-1} P(A_i) - \sum_{\substack{i<j \\ 1\le i,j\le n-1}} P(A_i \cap A_j) \\ + \sum_{\substack{i<j<k \\ 1\le i,j,k\le n-1}} P(A_i \cap A_j \cap A_k) - \cdots + (-1)^{n-2} P(\cap_{i=1}^{n-1} A_i). \end{aligned} \tag{1.2}$$

Descomponem $\cup_{i=1}^{n} A_i$ en la reunió dels conjunts $B := \cup_{i=1}^{n-1} A_i$ i A_n. Aplicant primer la fórmula (1.1) per a $n = 2$ i després la fórmula (1.2), hom obté que

$$\begin{aligned} P(\cup_{i=1}^{n} A_i) &= P(\cup_{i=1}^{n-1} A_i) + P(A_n) - P\big(\cup_{i=1}^{n-1} (A_i \cap A_n)\big) \\ &= \sum_{i=1}^{n} P(A_i) - \sum_{\substack{i<j \\ i,j=1,\ldots,n-1}} P(A_i \cap A_j) \\ &+ \sum_{\substack{i<j<k \\ i,j,k=1,\ldots,n-1}} P(A_i \cap A_j \cap A_k) \\ &- \cdots + (-1)^{n-2} P(\cap_{i=1}^{n-1} A_i) - P\big(\cup_{i=1}^{n-1} (A_i \cap A_n)\big). \end{aligned} \tag{1.3}$$

El primer terme del segon membre de la segona desigualtat de (1.3) és un dels que hem d'obtenir. Apliquem novament la fórmula (1.2) als conjunts

$$B_i = A_i \cap A_n, \quad i = 1, \ldots, n-1.$$

Resulta

$$\begin{aligned} P\big(\cup_{i=1}^{n-1} (A_i \cap A_n)\big) = \sum_{i=1}^{n-1} P(A_i \cap A_n) - \sum_{\substack{i<j \\ i,j=1,\ldots n-1}} P(A_i \cap A_j \cap A_n) \\ + \sum_{\substack{i<j<k \\ i,j,k=1,\ldots,n-1}} P(A_i \cap A_j \cap A_k \cap A_n) - \cdots + (-1)^{n-2} P\big((\cap_{i=1}^{n-1} A_i) \cap A_n\big). \end{aligned}$$

En substituir aquesta expressió en (1.3) obtenim la fórmula desitjada.

El resultat següent s'anomena propietat de continuïtat seqüencial de la probabilitat.

Proposició 1.1 *Sigui* $\{A_n, n \geq 1\}$ *una successió de conjunts de* $\mathcal{A}$. *Aleshores,*

1. *Si la successió és creixent i denotem per* A *el conjunt* $\cup_{n=1}^{\infty} A_n$, *es compleix que*

$$P(A) = \lim_{n\to\infty} P(A_n).$$

2. *Si la successió és decreixent i denotem per* A *el conjunt* $\cap_{n=1}^{\infty} A_n$, *es té també que*

$$P(A) = \lim_{n\to\infty} P(A_n).$$

Demostració: Comencem per l'apartat (1); l'apartat (2) es dedueix fàcilment d'aquest per pas al complementari.
Com que la successió $\{A_n, n \geq 1\}$ és creixent, els conjunts de la successió $\{B_n = A_n \cap A_{n-1}^c, n \geq 1\}$, que són també de $\mathcal{A}$, són disjunts dos a dos i

$$\cup_{n=1}^{\infty} A_n = \cup_{n=1}^{\infty} B_n,$$

on hem fet el conveni $A_0 = \emptyset$. Aleshores, per la propietat de σ-additivitat de P, tenim que

$$\begin{aligned} P(A) = P(\cup_{n=1}^{\infty} A_n) &= P(\cup_{n=1}^{\infty} B_n) = \sum_{n=1}^{\infty} P(B_n) \\ &= \lim_{n\to\infty} \sum_{i=1}^{n} P(B_i) = \lim_{n\to\infty} P(A_n). \end{aligned}$$

Considerem ara la situació de (2). La successió $\{A_n^c, n \geq 1\}$ és creixent i $\cup_{i=1}^{\infty} A_n^c = A^c$, aleshores, per l'apartat (1),

$$P(A^c) = 1 - P(A) = \lim_{n\to\infty} P(A_n^c) = 1 - \lim_{n\to\infty} P(A_n).$$

□

1.2.2 Espais de probabilitat finits

Si el conjunt Ω és finit, l'espai de probabilitat $(\Omega, \mathcal{A}, P)$ s'anomena *espai de probabilitat finit.* Els exemples 1.1 i 1.2 corresponen a aquesta situació. En aquest cas l'àlgebra $\mathcal{A}$ és necessàriament finita i es pot donar una descripció dels seus elements.

Definició 1.1 *Una família de conjunts de $\mathcal{A}$, $\{A_1, \dots, A_n\}$, és una* partició *del conjunt Ω, si són disjunts dos a dos i*

$$\cup_{i=1}^n A_i = \Omega.$$

Definició 1.2 L'àlgebra *(respectivament,* σ-àlgebra*)* generada *per una classe de conjunts és la mínima àlgebra (respectivament σ-àlgebra) que conté els conjunts de la classe.*

Teorema 1.1 *Tota àlgebra finita $\mathcal{A}$ de parts de Ω és l'àlgebra generada per una partició finita.*

Demostració: Suposem que $\mathcal{A} = \{B_1, \dots, B_m\}$. Els conjunts

$$B_1 \cap \cdots \cap B_m, \quad B_1^c \cap \cdots \cap B_m, \quad \dots, \quad B_1^c \cap \cdots \cap B_m^c,$$

obtinguts a partir del $B_1 \cap \cdots \cap B_m$ posant de dos fins a m signes de complementari de totes les maneres possibles, formen una partició del conjunt Ω. En efecte, és clar, per la seva construcció, que són conjunts disjunts i és fàcil comprovar que la seva reunió és Ω. L'àlgebra generada per aquests conjunts, que denotarem per $\mathcal{B}$, coincideix amb $\mathcal{A}$. En efecte, cadascun dels conjunts anteriors ha estat obtingut a partir dels conjunts de $\mathcal{A}$ mitjançant operacions compatibles amb l'estructura d'àlgebra. En conseqüència, $\mathcal{B} \subset \mathcal{A}$. Recíprocament, tot element de $\mathcal{A}$ s'obté fent una reunió de conjunts generadors de $\mathcal{B}$ dels que hem descrit més amunt. Per tant, $\mathcal{A} \subset \mathcal{B}$. □

Sigui $\mathcal{P}$ una partició finita de Ω com la descrita en la definició 1.1. L'àlgebra generada per $\mathcal{P}$ té 2^n elements i admet la descripció següent:

$$\mathcal{A} = \big\{\emptyset, A_{i_1} \cup \cdots \cup A_{i_k}, \{i_1, \ldots, i_k\} \subset \{1, \ldots, n\}\big\}.$$

En efecte, denotem per $\mathcal{B}$ el conjunt del segon membre de la igualtat anterior. El nombre de conjunts que són reunió de i conjunts de la partició és $\binom{n}{i}$. D'on, $\operatorname{card} \mathcal{B} = \sum_{i=0}^{n} \binom{n}{i} = 2^n$. Clarament, $\mathcal{B} \subset \mathcal{A}$, car els conjunts de $\mathcal{B}$ han estat formats a partir d'una operació compatible amb l'estructura d'àlgebra. D'altra banda, $\mathcal{B}$ és una àlgebra que conté els generadors de $\mathcal{A}$; per tant, conté $\mathcal{A}$.

El problema més delicat en la determinació d'un model probabilístic associat a una experiència aleatòria és, en general, l'assignació de probabilitat, és a dir, la definició de l'aplicació P. Aquest problema, però, no es planteja en el cas finit. Si suposem que el cardinal del conjunt Ω és r, l'assignació d'una probabilitat P queda fixada per r nombres $p_1, \ldots, p_r \in [0,1]$ tals que $\sum_{i=1}^{r} p_i = 1$. En efecte, per a tot $A \in \mathcal{A}$, $A \neq \emptyset$, definim

$$P(A) = \sum_{j:\omega_j \in A} p_j, \tag{1.4}$$

i posem $P(\emptyset) = 0$. Això dóna una aplicació $P : \mathcal{A} \longrightarrow [0,1]$ que compleix els axiomes de la probabilitat.

Un cas particular important s'obté en prendre

$$p_1 = \cdots = p_r = \frac{1}{r}.$$

L'espai de probabilitat obtingut així s'anomena *espai de probabilitat finit i uniforme*. La probabilitat d'un esdeveniment A es calcula mitjançant la fórmula

$$P(A) = \frac{\sharp A}{\sharp \Omega}, \tag{1.5}$$

on la notació $\sharp A$ indica el cardinal del conjunt A.

Aquest model té una importància històrica. Abans que la teoria de la probabilitat es presentés amb el formalisme que hem introduït, s'estudiaven experiències aleatòries amb resultats igualment probables. La definició de probabilitat donada per Laplace (1812) en la seva obra *Théorie analytique des Probabilités* correspon a aquesta situació.

1.3 La combinatòria en el càlcul de probabilitats

En els espais de probabilitat finits i uniformes, el càlcul de probabilitats d'esdeveniments es redueix al càlcul del cardinal de conjunts, a causa de la fórmula (1.5). L'eina matemàtica per abordar aquest problema és la *Combinatòria*. En aquesta secció introduïm alguns elements relacionats amb aquest tema, basant-nos en problemes d'obtenció de mostres o en problemes d'ocupació.

Suposem que tenim una urna amb n boles diferents, numerades des de 1 fins a n, i que fem k extraccions sota condicions que seran especificades en cada cas. Volem determinar el nombre de resultats possibles.

De vegades resulta també útil considerar el model consistent en n caselles numerades en les quals volem col·locar k boles.

Donat un conjunt A, indicarem amb la lletra minúscula a el seu cardinal.

1. *Mostreig ordenat amb reposició.* Extraiem successivament k boles, i les tornem a l'urna després de cada extracció. Es tracta de comptar el nombre de k-tuples ordenades $(n_1, \dots, n_k)$, on $n_j \in \{1, \dots, n\}, j = 1, \dots, k$. El resultat és n^k.

   Aquest nombre és també el cardinal del conjunt de les aplicacions $f : K \longrightarrow N$.

2. *Mostreig ordenat sense reposició.* Suposem $k \leq n$. En aquest cas, les boles extretes no es tornen a l'urna, de manera que les k-tuples obtingudes estan formades per elements diferents. El nombre de resultats possibles és

$$n(n-1)(n-2)\cdots(n-k+1) = \frac{n!}{(n-k)!}.$$

   Aquest nombre coincideix amb el cardinal del conjunt de les aplicacions injectives $f : K \longrightarrow N$.

Observeu també que, quan $k = n$, obtenim $n!$ (*permutacions d'un conjunt de n elements*).

3. *Mostreig sense ordre i sense reposició.* Suposem també $k \leq n$. No tornem les boles extretes a l'urna, ni ens importa l'ordre en què han aparegut. És, doncs, com si haguéssim tret les k boles en una única extracció. Es tracta de comptar els subconjunts de N amb cardinal k. N'hi ha

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Relacionat amb aquest exemple, podem plantejar el càlcul del nombre de particions del conjunt N en r subconjunts $K_1, \ldots, K_r$ amb cardinals $k_1, \ldots, k_r$, respectivament. Òbviament $k_1 + \cdots + k_r = n$. Una manera d'enfocar el problema és comptar primer els subconjunts de N de k_1 elements; ja sabem que n'hi ha $\binom{n}{k_1}$. Tot seguit, comptar els subconjunts de $N - K_1$ de k_2 elements: $\binom{n-k_1}{k_2}$, ... Finalment, els subconjunts de $N - K_1 - K_2 - \cdots - K_{r-1}$ de k_r elements: $\binom{n-k_1-\cdots-k_{r-1}}{k_r}$. El resultat que busquem és

$$\binom{n}{k_1}\binom{n-k_1}{k_2}\cdots\binom{n-k_1-\cdots-k_{r-1}}{k_r},$$

que coincideix amb l'anomenat *coeficient multinomial*

$$\frac{n!}{k_1! \cdot k_2! \cdot \cdots \cdot k_r!}.$$

4. *Mostreig sense ordre i amb reposició.* Tornem la bola a l'urna després de cada extracció; per tant, després de les k extraccions hi haurà possiblement repeticions en el resultat.

   El problema consisteix a trobar el nombre de conjunts no ordenats, $\{i_1, \ldots, i_k\}$, on $i_j \in \{1, \ldots, n\}$, $j = 1, \ldots, k$, no necessàriament diferents. Una manera anàloga d'enfocar el problema és comptar el nombre de solucions $(r_1, \ldots, r_n)$ d'enters no negatius de l'equació

$$r_1 + \cdots + r_n = k.$$

Aquí $r_j, j = 1, \ldots, n$, denota el nombre de vegades que ha sortit la bola j en les k extraccions. L'equivalència entre ambdós plantejaments s'explica de la manera següent: De les boles extretes ens interessa el tipus $(j = 1, \ldots, n)$ i el nombre obtingut de cada tipus $(r_j, j = 1, \ldots, n)$.

Alternativament, el problema consisteix a comptar el nombre de maneres possible de col·locar k boles en n urnes, per la qual cosa cal fixar el nombre de boles que va a cada urna. Aquest nombre és

$$\binom{n+k-1}{k},$$

que coincideix amb $\binom{n+k-1}{n-1}$. En efecte, escollim els k llocs on van les boles entre $k+n-1$ possibilitats, on $n-1$ és el nombre de separacions entre les n urnes.

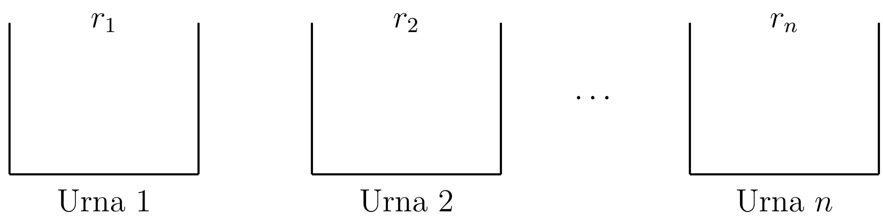

Figura 1.1: Una possible disposició de k boles en n urnes

En la terminologia de la combinatòria, el cas (1) correspon a les *variacions amb repetició* i el cas (2), a les *variacions*. Les dues possibilitats presentades en el cas (3) corresponen a les *combinacions* i les *permutacions amb repetició*, respectivament, i, finalment el cas (4) a les *combinacions amb repetició*.

Podríem estudiar altres situacions, però ens aturem aquí. Els exemples que hem presentat, ultra un interès didàctic, són útils com a models per a la física i l'enginyeria. Com a il·lustració presentem un exemple clàssic de la mecànica estadística.

Considerem un sistema mecànic format per k partícules indistingibles. En la mecànica estadística és usual subdividir l'espai de fases en un nombre gran n de subregions o cel·les, corresponents a diversos estats d'energia, de manera que cada partícula estigui assignada a una cel·la. L'estat del sistema queda, doncs, descrit (utilitzant el llenguatge d'un problema d'ocupació) en termes d'una distribució aleatòria de k boles en n caselles. Històricament s'han proposat diversos models d'ocupació que el físic anomena *estadístiques*.

L'*estadística de Maxwell-Boltzmann* es basa en el model (1) d'un mostreig ordenat amb reposició. Suposa que les n^k disposicions possibles de les partícules tenen la mateixa probabilitat. Noves teories de la física demostren que cap de les partícules conegudes es distribueix segons aquest model.

L'*estadística de Bose-Einstein* proposa el model (4) d'un mostreig no ordenat i amb reposició. Designem per $A_{k,n}$ el nombre combinatori $\binom{n+k-1}{k}$. Aleshores se suposa que hi ha $A_{k,n}$ maneres diferents de disposar les partícules i cadascuna d'aquestes disposicions és equiprobable. Es demostra que els fotons segueixen aquest model d'ocupació.

Finalment, l'*estadística de Fermi-Dirac* proposa el model (3) de mostreig no ordenat i sense reposició, que atribueix a cadascuna de les $\binom{n}{k}$ distribucions la mateixa probabilitat, és a dir, $\binom{n}{k}^{-1}$. Aquest model s'aplica als electrons, els neutrons i els protons.

1.4 Alguns exemples de models probabilístics

El primer exemple que presentem en aquest apartat pretén il·lustrar la necessitat d'una certa dosi d'imaginació per a descriure el model probabilístic, de manera que sigui fàcil calcular determinades probabilitats. El segon exemple, un clàssic en el càlcul de probabilitats, s'ha triat amb la intenció de posar en evidència com l'elecció del model condiciona els resultats dels càlculs i també l'ambigüitat que es presenta en abordar tota modelització.

1.4.1 Problema de l'escrutini

El dia 27 de juliol de 1997, es celebraren eleccions a la presidència del Barça. Hi havia únicament dos candidats, el senyor Fernández i el senyor Núñez. Va guanyar el senyor Núñez. Un soci amb inquietuds intel·lectuals es va fer la pregunta següent: Seria gaire estrany pensar que al llarg de l'escrutini el senyor Núñez ha anat sempre al davant del senyor Fernández?

Per simplificar l'exposició, designarem per A el senyor Núñez i per B el senyor

Fernández. Denotarem per a el nombre de vots obtinguts pel candidat A i per b els obtinguts pel candidat B. Sabem que $a > b$.

Representem el procés d'escrutini mitjançant un camí no decreixent que va del punt $(0,0)$ (que representa l'inici de l'escrutini) al punt (b,a) (que representa el resultat final). Ens desplacem una unitat a la dreta si el vot és favorable al candidat B, i una unitat cap amunt si ho és al candidat A. Observeu que tot camí com el descrit anteriorment es pot identificar amb una funció no decreixent

$$f : \{0,1,2,\dots,b\} \longrightarrow \{0,1,2,\dots,a\},$$

$f(0) = 0$, $f(b) = a$.

Sigui Ω el conjunt de tals funcions, que és un conjunt finit. Definim $\mathcal{A} = \mathcal{P}(\Omega)$ i P mitjançant la fórmula (1.5). La resposta al soci inquiet la donarem calculant la probabilitat del subconjunt de Ω format pels camins que no tallen la recta $y = x$ (diagonal).

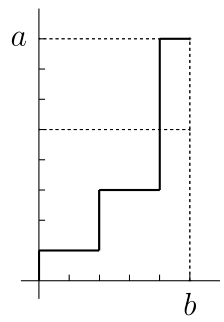

Figura 1.2: Un element del conjunt Ω

Considerem la partició de Ω, $\Omega = C \cup D \cup E$, on

- C és l'esdeveniment "Al llarg de l'escrutini, el candidat A ha anat sempre per davant del candidat B",
- D és el subconjunt d'elements de C^c tals que el primer vot ha estat favorable al candidat A,
- E és el subconjunt d'elements de C^c tals que el primer vot ha estat favorable al candidat B.

Els gràfics següents representen elements típics de cadascun dels conjunts descrits anteriorment.

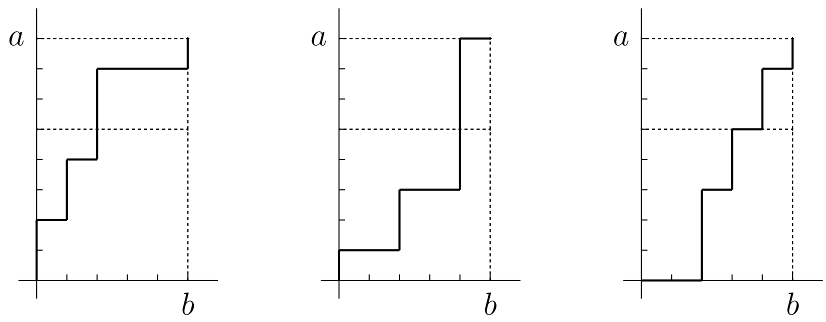

Figura 1.3: Elements dels conjunts C, D i E, respectivament

El càlcul de $P(C)$ es fa seguint les etapes següents:

1. *Càlcul del cardinal de* Ω

   Un element de Ω es pot identificar amb una successió de $a+b$ elements, $(x_1, \ldots, x_{a+b})$, amb a elements que designen el candidat A i b elements que designen el candidat B. Les possibles successions amb aquestes característiques s'obtenen comptant les maneres possibles de col·locar a vegades el nom del candidat A en $a+b$ caselles o, equivalentment, b vegades el nom del candidat B en $a+b$ caselles. En definitiva,

$$\sharp\Omega = \binom{a+b}{a} = \binom{a+b}{b}.$$

2. *Principi de reflexió*

   Es compleix que

$$\sharp D = \sharp E.$$

   En efecte, pot establir-se una correspondència bijectiva entre els conjunts D i E de la manera següent. Fixem un element $\omega \in D$ i sigui (u, u) el primer punt (d'acord amb l'ordre parcial de $\mathbb{R}^2$ definit coordenada a coordenada) on ω talla la diagonal. Associem a ω una trajectòria $\tilde{\omega}$ que coincideix amb ω des de (u, u) fins a (b, a) i, entre $(0, 0)$ i (u, u), s'obté per reflexió de ω respecte a la diagonal.

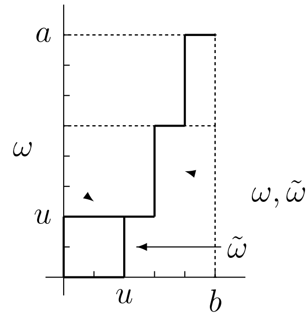

Figura 1.4: Bijecció entre els conjunts D i E

3. *Càlcul del cardinal d'E*

   Podem identificar les trajectòries del conjunt E amb camins com els que hem descrit abans, que surten de $(0,0)$ i van a parar a $(b-1,a)$. Per les mateixes consideracions que en l'etapa 1,

$$\sharp E = \binom{a+b-1}{a} = \binom{a+b-1}{b-1}.$$

Finalment, per la fórmula (1.5),

$$\begin{aligned} P(C) &= \frac{\sharp A}{\sharp\Omega} = \frac{\sharp\Omega - \sharp D - \sharp E}{\sharp\Omega} = 1 - 2\frac{\sharp E}{\sharp\Omega} \\ &= 1 - 2\frac{\binom{a+b-1}{a}}{\binom{a+b}{a}} \\ &= \frac{a-b}{a+b}. \end{aligned}$$

Els mitjant de comunicació varen donar les dades següents: El senyor Núñez va obtenir 24.025 vots i el senyor Fernández, 5.209. Per tant,

$$P(C) = 0,64.$$

1.4.2 Una paradoxa de Bertrand

Es considera una corda dibuixada aleatòriament en el cercle unitat. Quina és la probabilitat que la seva longitud sigui més gran que $\sqrt{3}$?

Aquest problema es troba en el llibre de Joseph Bertrand *Calcul des Probabilités*, editat per Gauthier Villars a París, l'any 1899. L'autor dóna diverses solucions que porten a resultats diferents. Com veurem, l'explicació de la contradicció aparent cal buscar-la en els diferents models probabilístics que s'adapten al problema. Noteu que el valor $\sqrt{3}$ és el de la longitud del costat d'un triangle equilàter inscrit en la circumferència unitat.

Proposarem tres models probabilístics diferents que porten a resultats diferents.

1. L'elecció aleatòria de la corda es basa en l'elecció d'un punt ω a l'atzar en el disc unitat. Una vegada triat aquest punt, dibuixem el radi que passa per ell i, tot seguit, la corda perpendicular a aquest radi que passa per ω.

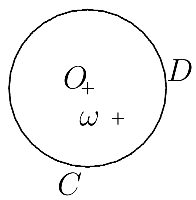

Figura 1.5: El punt ω és triat a l'atzar en el cercle

2. Fixem un punt de la circumferència unitat que serà un dels extrems de la corda. L'elecció aleatòria d'aquesta es basa en l'elecció a l'atzar d'un punt de la circumferència unitat que serà l'altre extrem de la corda.

3. Fixem un radi del disc unitat. Triem un punt a l'atzar en aquest radi i, tot seguit, dibuixem la recta perpendicular al radi que passa per aquest punt.

És fàcil de descriure els espais mostrals Ω_i, $i = 1, 2, 3$ que corresponen a cadascun dels models anteriors:

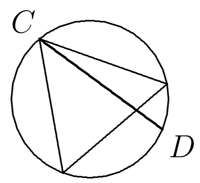

Figura 1.6: C és fix, D es tria a l'atzar

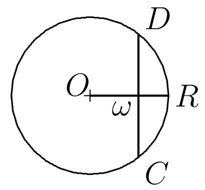

Figura 1.7: El punt ω és triat a l'atzar en el radi $\overline{OR}$.

$$\begin{aligned}
\Omega_1 &= \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}, \\
\Omega_2 &= \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}, \\
\Omega_3 &= [0,1].
\end{aligned}$$

L'assignació de probabilitat és, en canvi, una qüestió més delicada, i l'elecció del conjunt d'esdeveniments $\mathcal{A}_i$, $i = 1, 2, 3$, es veu condicionada per tal elecció.

La frase *triar a l'atzar* que hem escrit en la descripció dels tres models que proposem, suggereix una assignació de probabilitat uniforme. Hem introduït amb precisió aquest concepte en el cas d'un espai mostral Ω finit. Si Ω és un conjunt numerable, no és possible definir una probabilitat sobre una σ-àlgebra de parts de Ω que compleixi que

$$P(\{\omega\}) = c, \quad \forall \omega \in \Omega,$$

ja que, tal condició implica $P(\Omega) = \infty$, la qual cosa contradiu l'axioma (1) de la probabilitat.

Per als models que hem proposat, els espais mostrals corresponents són conjunts infinits, no numerables. La situació és, doncs, encara pitjor. Podríem, però, intentar exportar l'heurística de la fórmula (1.5) en termes de proporció d'àrea o de longitud d'arc de corba, segons els casos. En altres paraules, centrant-nos en el model 1, considerarem subconjunts A de Ω_1 pels quals tingui sentit parlar de la seva *àrea*, que denotarem per $|A|$, i, per a tals esdeveniments, definirem

$$P(A) = \frac{|A|}{|\Omega_1|}. \tag{1.6}$$

Per tant, l'espai de probabilitat corresponent al model 1 és $(\Omega_1, \mathcal{L}(\Omega_1), l)$, on $\mathcal{L}(\Omega_1)$ és la σ-àlgebra dels subconjunts de Ω_1 que són mesurables en el sentit de Lebesgue, i l és la mesura de Lebesgue en $\mathbb{R}^2$.

Amb la formulació anterior, podem donar una solució al problema plantejat de la manera següent.

- *Per al model 1*: Sigui x la longitud del segment $\overline{O\omega}$. La longitud de corda igual a $\sqrt{3}$ correspon a $x = \frac{1}{2}$. Per tant, l'esdeveniment A = "*la longitud de la corda és més gran que* $\sqrt{3}$" s'identifica amb el subconjunt de $\mathbb{R}^2$ format per la bola centrada en $(0,0)$ i radi $\frac{1}{2}$ i, en conseqüència,

$$P(A) = \frac{\pi/4}{\pi} = \frac{1}{4}.$$

- *Per al model 2*: L'esdeveniment A s'identifica amb l'arc de circumferència assenyalat en la figura 1.6, que té una longitud igual a la tercera part de la longitud de la circumferència. Aixi doncs, per aquest model,

$$P(A) = \frac{1}{3}.$$

- *Per al model 3*: L'elecció de ω favorable a l'esdeveniment A correspon a punts sobre el segment d'extrems O i R (vegeu figura 1.7) per als quals la longitud del segment $\overline{O\omega}$ és inferior o igual a $\frac{1}{2}$. Per tant,

$$P(A) = \frac{1}{2}.$$

1.5 Probabilitat condicionada

De vegades es pot disposar d'informació complementària que ens permeti de modificar el model probabilístic associat a una experiència aleatòria.

Exemple 1.4 ([Ba]) Juguem a la ruleta els números 3, 13 i 22. Els resultats possibles són els números del 0 al 36. Si la ruleta no està trucada, és natural considerar un model *uniforme*. Per tant, la probabilitat de guanyar és 3/37. Ara bé, si la ruleta està trucada de tal manera que sempre surt un nombre imparell, la probabilitat de guanyar ja no serà la mateixa d'abans.

Per tractar aquests tipus de qüestions introduïm el concepte de probabilitat condicionada. Sigui $(\Omega, \mathcal{A}, P)$ un espai de probabilitat associat a una certa experiència aleatòria. Fixem un conjunt $B \in \mathcal{A}$ de probabilitat no nul·la.

Definició 1.3 *La* probabilitat *d'un conjunt $A \in \mathcal{A}$* condicionada *per B es defineix per*

$$P(A/B) = \frac{P(A \cap B)}{P(B)}, \tag{1.7}$$

Es comprova fàcilment que $P(\,\cdot\,/B) : \mathcal{A} \longrightarrow [0,1]$ defineix també una probabilitat sobre $(\Omega, \mathcal{A})$. Aquesta probabilitat s'introdueix quan sabem *a priori* que l'esdeveniment B s'ha realitzat. És a dir, en rebre aquesta informació addicional sobre l'experiència que estem analitzant, cal modificar la probabilitat ja que, per exemple, l'esdeveniment B passa a tenir probabilitat 1.

En l'exemple 1.4 de més amunt tenim

$$\Omega = \{0, 1, \ldots, 36\}, \quad A = \{3, 13, 22\}, \quad B = \{1, 3, 5, \ldots, 35\},$$

i ens demanen $P(A/B)$. A partir de la fórmula (1.7) obtenim $P(A/B) = \frac{1}{9}$.

La proposició següent proporciona una fórmula útil per al càlcul de la probabilitat d'una intersecció d'esdeveniments, a partir de la noció de probabilitat condicionada.

Proposició 1.2 (Fórmula de les probabilitats compostes) *Siguin A_1, $\ldots, A_n$ elements de $\mathcal{A}$ tals que $P(A_1 \cap \cdots \cap A_{n-1}) > 0$. Aleshores,*

$$\begin{aligned} &P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1) \\ &\times P(A_2/A_1) P(A_3/A_1 \cap A_2) \cdots P(A_n/A_1 \cap A_2 \cap \cdots \cap A_{n-1}). \end{aligned} \tag{1.8}$$

Demostració: S'utilitza inducció. Per a $n = 2$ s'obté $P(A_1 \cap A_2) = P(A_1)$ $P(A_2/A_1)$, que és una conseqüència immediata de la definició de probabilitat condicionada. Suposant que la fórmula és certa per a $n-1$, escrivim

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1 \cap A_2 \cap \cdots \cap A_{n-1})P(A_n/A_1 \cap \cdots \cap A_{n-1})$$

i apliquem la hipòtesi d'inducció. □

Exemple 1.5 Volem calcular la probabilitat que en una població de n persones n'hi hagi almenys dues amb la mateixa data d'aniversari. Suposem que ningú no ha nascut el 29 de febrer i que tots els dies són equiprobables.

Per a $k = 1, 2, \ldots, n-1$, sigui A_k l'esdeveniment "*la* $(k+1)$*-èsima persona té una data d'aniversari diferent de les* k *que la precedeixen*". Aleshores, la probabilitat buscada és $p = 1 - P(A_1 \cap \cdots \cap A_{n-1})$. Es compleix que

$$\begin{aligned}
&P(A_1) = \frac{364}{365}, \quad P(A_2/A_1) = \frac{363}{365}, \\
&\cdots \\
&P(A_{n-1}/A_1 \cap \cdots \cap A_{n-2}) = \frac{365-(n-1)}{365},
\end{aligned}$$

d'on, a partir de (1.8),

$$p = 1 - \frac{364 \cdot 363 \cdot \cdots \cdot (365-n+1)}{365^{n-1}}.$$

Considerem una partició finita del conjunt Ω, $\mathcal{P}_1 = \{A_1, A_2, \ldots, A_n\} \subset \mathcal{A}$ on $P(A_i) > 0, 1 \leq i \leq n$. Donat un esdeveniment $A \in \mathcal{A}$, la coneixença de les probabilitats $P(A_i)$ i de les probabilitats condicionades $P(A/A_i)$ permet de calcular la probabilitat de A de la manera següent:

$$P(A) = \sum_{i=1}^{n} P(A \cap A_i) = \sum_{i=1}^{n} P(A_i)P(A/A_i). \tag{1.9}$$

La primera d'aquestes igualtats s'anomena fórmula de les probabilitats totals, i és conseqüència de la descomposició en conjunts disjunts $A = \bigcup_{i=1}^{n}(A \cap A_i)$. La segona igualtat surt de la fórmula (1.7).

(1.9) expressa que la probabilitat de A pot obtenir-se com una mitjana ponderada de les probabilitats condicionades.

Exemple 1.6 ([P]) Considerem una baralla de pòquer, que té 52 cartes, 26 de negres i 26 de vermelles. Barregem bé les cartes i en traiem dues, una després de l'altra. Quina és la probabilitat que la segona carta extreta sigui negra?

Podem esperar una resposta com la següent: Depèn; si la primera carta que hem extret ha estat vermella, la probabilitat demanada és $\frac{26}{51}$; si la primera ha estat negra, la probabilitat val ara $\frac{25}{51}$. Aquesta resposta no és gaire satisfactòria, ja que res ens fa suposar que tinguem coneixement del que ha passat en la primera extracció.

Sembla, doncs, més raonable, davant l'aparent desconeixement del resultat de la primera extracció, tenir en compte les dues possibilitats, negra o vermella, i fer una mitjana d'ambdues. En altres paraules, la resposta satisfactòria s'obté utilitzant la fórmula (1.9). Considerem com a espai mostral el determinat per les observacions simultànies de les dues extraccions, és a dir, $\Omega = \{(v_1, v_2), (v_1, n_2), (n_1, v_2), (n_1, n_2)\}$ on el subíndex fa esment a l'ordre d'extracció, v significa vermella i n negra. Considerem la partició $\mathcal{P}$ formada pels conjunts $V_1 = \{(v_1, v_2), (v_1, n_2)\}$, $N_1 = \{(n_1, v_2), (n_1, n_2)\}$ i sigui$N_2 = \{(v_1, n_2), (n_1, n_2)\}$. Aleshores, per (1.9)

$$P(N_2) = P(N_2/V_1)P(V_1) + P(N_2/N_1)P(N_1) = \frac{26}{51} \cdot \frac{1}{2} + \frac{25}{51} \cdot \frac{1}{2} = \frac{1}{2}.$$

1.5.1 Urnes de Pólya

Es considera una urna que conté b boles blanques i v boles vermelles. Es treu una bola de l'urna, es mira el seu color, es retorna a l'urna i s'afegeixen c boles del mateix color que la bola que ha estat extreta (c pot ser positiu o negatiu; un valor de c negatiu significa que es treuen boles). Per a $c = 0$ el procés anterior és el d'un mostreig amb reposició. Aquest model fou proposat per G. Pólya per a l'assignació de probabilitat en situacions de contagi per malalties, on cal tenir en compte que la presència d'individus malalts incrementa la probabilitat que un individu qualsevol de la població estigui malalt. Ens preguntem: *Quina és la probabilitat que les tres primeres extraccions tinguin colors* (b, b, v)*, en aquest ordre? Quina seria la resposta per a* (b, v, b)*,* (v, b, b)*?*

L'elecció a l'atzar d'una bola en una urna és una experiència aleatòria a la qual correspon un espai mostral finit perfectament determinat a partir de la composició de l'urna. Interpretem el terme *elecció a l'atzar* mitjançant una assignació uniforme de probabilitat. En el problema que ens ocupa, la composició de l'urna canvia després de cada extracció i, per tant, també ho fa al model probabilístic corresponent.

Denotem per B_i l'esdeveniment *"el resultat de la i-èsima extracció és bola blanca"* i per V_i l'esdeveniment *"el resultat de la i-èsima extracció és bola vermella"*. Hem de calcular $P(B_1 \cap B_2 \cap V_3)$. Per la fórmula (1.8),

$$P(B_1 \cap B_2 \cap V_3) = P(B_1)P(B_2/B_1)P(V_3/B_1 \cap B_2). \tag{1.10}$$

Per al càlcul de $P(B_1)$, utilitzem la composició inicial de l'urna. Tindrem que

$$P(B_1) = \frac{b}{b+v}.$$

En el càlcul del segon factor del segon membre de (1.10), utilitzem la composició de l'urna derivada del fet que la primera extracció ha resultat bola blanca. És a dir, en el moment de fer la segona extracció l'urna conté $b + c$ boles blanques i v de vermelles. Així, doncs,

$$P(B_2/B_1) = \frac{b+c}{b+c+v}.$$

Finalment, en el càlcul de $P(V_3/B_1 \cap B_2)$ cal tenir present que, en fer la tercera extracció l'urna conté v boles vermelles i $b + 2c$ boles blanques. Per tant,

$$P(V_3/B_1 \cap B_2) = \frac{v}{b+2c+v}.$$

Resulta, doncs,

$$P(B_1 \cap B_2 \cap V_3) = \frac{b(b+c)v}{(b+v)(b+c+v)(b+2c+v)}.$$

Anàlogament provaríem

$$P(B_1 \cap V_2 \cap B_3) = P(V_1 \cap B_2 \cap B_3) = \frac{b(b+c)v}{(b+v)(b+c+v)(b+2c+v)}.$$

Observem que l'ordre en què apareixen els diversos colors no afecta la probabilitat. D'on resulta que la probabilitat d'extreure 2 boles blanques i una de vermella en les tres primeres extraccions és

$$\binom{3}{1}\frac{b(b+c)v}{(b+v)(b+c+v)(b+2c+v)}.$$

No és difícil extrapolar els arguments anteriors per obtenir el resultat següent: La probabilitat d'extreure una successió fixada de k boles blanques i $n-k$ de vermelles és

$$p_{n,k} = \frac{b(b+c)\cdots(b+(k-1)c)v(v+c)\cdots(v+(n-k-1)c)}{(b+v)(b+v+c)(b+v+2c)\cdots(b+v+(n-1)c)},$$

$n \geq 1$, $c \geq 0$.

En conseqüència, la probabilitat d'extreure k boles blanques i $n-k$ vermelles en n extraccions és

$$\binom{n}{k}p_{n,k} = \frac{\binom{-b/c}{k}\binom{-v/c}{n-k}}{\binom{(-b+v)/c}{n}},$$

on, per a $m, n \in \mathbb{Z}_+$, hem posat

$$\binom{-m}{n} = \frac{-m(-m-1)\cdots(-m-n+1)}{n!}.$$

1.5.2 Inversió de les condicions: fórmula de Bayes

Si A i B són dos esdeveniments de probabilitat no nul·la, es compleix que

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = P(B/A)\frac{P(A)}{P(B)}, \tag{1.11}$$

fórmula que permet d'invertir les condicions en el càlcul de les probabilitats condicionades.

Més generalment, considerem dues particions del conjunt Ω, $\mathcal{P}_1 = \{A_1, A_2, \ldots, A_n\}$ i $\mathcal{P}_2 = \{B_1, B_2, \ldots, B_m\}$, donades per conjunts de $\mathcal{A}$ de probabilitat no nul·la. Coneguda la probabilitat dels conjunts de la primera partició, $P(A_i)$, i la probabilitat dels conjunts de la segona partició condicionada pels conjunts de la primera, $P(B_j/A_i)$, ens plantegem de calcular la probabilitat dels conjunts de la primera condicionada per la dels de la segona $P(A_i/B_j)$. Utilitzant la fórmula d'inversió (1.11) i la fórmula (1.9) s'obté la *fórmula de Bayes* , que resol aquest problema:

$$P(A_i/B_j) = \frac{P(A_i)P(B_j/A_i)}{P(B_j)} = \frac{P(A_i)P(B_j/A_i)}{\sum_{k=1}^{n} P(B_j/A_k)P(A_k)}. \tag{1.12}$$

Exemple 1.7 Tenim una urna amb 3 boles blanques i 2 de negres. Traiem dues boles successivament. No hi ha reposició. Volem calcular la probabilitat que la primera bola sigui blanca, sabent que la segona és negra.

Podem prendre com a conjunt Ω el producte cartesià $\{b, n\} \times \{b, n\}$, on b representa treure una bola blanca i n treure una bola negra. Considerem els esdeveniments següents:

- $B_1 =$ "primera bola blanca" $= \{(b, n), (b, b)\}$,
- $B_2 =$ "segona bola blanca" $= \{(b, b), (n, b)\}$,
- $N_1 =$ "primera bola negra" $= \{(n, b), (n, n)\}$,
- $N_2 =$ "segona bola negra" $= \{(b, n), (n, n)\}$.

Les particions són $\mathcal{P}_1 = \{B_1, N_1\}$, $\mathcal{P}_2 = \{B_2, N_2\}$ i sabem que

$$P(B_1) = \frac{3}{5}, \quad P(N_1) = \frac{2}{5},$$
$$P(N_2/B_1) = \frac{1}{2}, \quad P(N_2/N_1) = \frac{1}{4}, \quad P(B_2/B_1) = \frac{1}{2}, \quad P(B_2/N_1) = \frac{3}{4}.$$

Aleshores, per una aplicació de la fórmula de Bayes,

$$P(B_1/N_2) = \frac{P(B_1)P(N_2/B_1)}{P(B_1)P(N_2/B_1) + P(N_1)P(N_2/N_1)} = \frac{3}{4}.$$

Exemple 1.8 ([P]) Considerem un test sanguini per a determinar una certa malaltia amb dos resultats possibles: positiu i negatiu. Es coneix, a partir de l'experiència, que un 95% de persones que pateixen la malaltia donen positiu i un 2% que no la té dóna també positiu. D'altra banda, se sap que un 1% de la població pateix la malaltia. Es vol calcular la probabilitat que una persona elegida a l'atzar estigui malalta suposant que el test li ha donat positiu.

En aquest exemple tenim dos caràcters discriminants diferents, amb els seus respectius contraris: malalt (sa) i positiu (negatiu). Cadascun dóna lloc a una partició de l'espai mostral $\Omega = \{(m,+),(m,-),(s,+),(s,-)\}$, on hem indicat per m l'atribut *malalt*, per s l'atribut *sa*, per $+$ que el resultat del test ha estat *positiu*, i per $-$ que el resultat del test ha estat *negatiu*. Considerem, doncs, la partició $\mathcal{P}_1$ formada pels conjunts $M = \{(m,+),(m,-)\}$ i $S = \{(s,+),(s,-)\}$ i la $\mathcal{P}_2$ determinada pels conjunts $p = \{(m,+),(s,+)\}$ i $N = \{(m,-),(s,-)\}$. Aleshores, segons la fórmula (1.12),

$$P(M/p) = \frac{P(p/M)P(M)}{P(p/M)P(M) + P(p/S)P(S)} = \frac{0.95 \cdot 0.01}{0.95 \cdot 0.01 + 0.02 \cdot 0.99} = 0.32.$$

Observem que aquesta probabilitat és molt petita (en canvi $P(p/M)$ és gran). Aquest resultat pot explicar-se perquè hi ha molt pocs malalts i, per tant, la presència de falsos positius pot ser gran.

1.6 Esdeveniments independents

Definició 1.4 *Es diu que dos esdeveniments $A, B \in \mathcal{A}$ són* independents *si*

$$P(A \cap B) = P(A)P(B).$$

Si $P(A) > 0$, el fet que A i B siguin independents és equivalent a dir que $P(B/A) = P(B)$. És a dir, si sabem que A s'ha realitzat, la probabilitat de B no queda modificada. En altres paraules, la realització de l'esdeveniment A no influeix de cap manera sobre l'esdeveniment B. Si $P(B) > 0$, podem fer un raonament simètric i concloure que la definició d'independència que hem donat equival a la noció intuïtiva d'independència entre la realització dels esdeveniments.

1.6.1 Propietats dels esdeveniments independents

1. Els esdeveniments Ω i $\emptyset$ són independents de qualsevol esdeveniment A.

   En efecte, com que $A \cap \Omega = A$, $A \cap \emptyset = \emptyset$, $P(\Omega) = 1$ i $P(\emptyset) = 0$, es té

   $$\begin{aligned} P(A \cap \Omega) &= P(A) = P(A)P(\Omega), \\ P(A \cap \emptyset) &= P(\emptyset) = P(A)P(\emptyset). \end{aligned}$$

2. Si un esdeveniment A és independent de si mateix, aleshores $P(A) = 0$ o $P(A) = 1$.

   En efecte, per la definició 1.4, haurà de complir-se

   $$P(A) = \big(P(A)\big)^2,$$

   i això implica $P(A) = 0$ o $P(A) = 1$.

3. Les afirmacions següents són equivalents:

   (a) "A és independent de B",

   (b) "A^c és independent de B",

   (c) "A^c és independent de B^c".

   En efecte, vegem, per exemple, que "A independent de B" implica "A^c independent de B". Considerem la descomposició

   $$B = (A \cap B) \cup (A^c \cap B).$$

   Aleshores,

   $$\begin{aligned} P(A^c \cap B) &= P(B) - P(A \cap B) = P(B)\big(1 - P(A)\big) \\ &= P(B)P(A^c). \end{aligned}$$

   Arguments similars s'apliquen per a les proves de les altres implicacions.

En la definició següent el conjunt I és arbitrari.

Definició 1.5 *Un conjunt d'esdeveniments* $\{A_i, i \in I\} \subset \mathcal{A}$ *són dits independents si per a qualsevol subconjunt finit* $\{A_{i_1}, \ldots, A_{i_n}\}$ *se satisfà que*

$$P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_n}) = P(A_{i_1})P(A_{i_2}) \cdots P(A_{i_n}).$$

Aquesta definició d'independència és més forta que la independència entre cada parella d'esdeveniments del conjunt $\{A_i, i \in I\}$. Per exemple, en el cas de tres esdeveniments A_1, A_2 i A_3, pot ocórrer que

$$\begin{aligned}
P(A_1 \cap A_2) &= P(A_1)P(A_2), \\
P(A_1 \cap A_3) &= P(A_1)P(A_3), \\
P(A_2 \cap A_3) &= P(A_2)P(A_3),
\end{aligned}$$

però, en canvi,

$$P(A_1 \cap A_2 \cap A_3) \neq P(A_1)P(A_2)P(A_3).$$

Això voldria dir que els esdeveniments A_1, A_2 i A_3 són independents dos a dos, però $A_1 \cap A_2$ no és independent de A_3.

Per exemple, en el llançament simultani de dos daus, els esdeveniments $A =$ "el resultat del primer dau és 1, 2 o 3", $B =$ "el resultat del segon dau és 4, 5, o 6", i $C =$ "la suma dels resultats obtinguts en ambdós daus és 7" són independents dos a dos, però no conjuntament, com pot comprovar-se fàcilment.

Exemple 1.9 Llancem una moneda n vegades. Suposem que els diferents llançaments són independents, que la probabilitat que surti cara és p, i la probabilitat que surti creu, $q = 1 - p$. Volem determinar quina és la probabilitat d'obtenir una seqüència determinada de cares i creus.

En primer lloc especifiquem quin és l'espai de probabilitat associat a l'experiència aleatòria descrita en aquest exemple. L'espai mostral és

$$\Omega = \{\omega = (\omega_1, \ldots, \omega_n) : \omega_i = 0 \text{ o } \omega_i = 1; i = 1, \ldots, n\},$$

on hem indicat per 1 la cara i per 0 la creu.

Podem prendre $\mathcal{A}$ igual al conjunt de tots els subconjunts de Ω. Per a assignar una probabilitat als elements de Ω definim, per a $i = 1, \ldots, n$, el conjunt $A_i = \{\omega : \omega_i = 1\}$, que correspon a l'esdeveniment "*el resultat de la i-èsima tirada és cara*". Aleshores, $P(A_i) = p$. Tot element de Ω es pot representar mitjançant la intersecció de conjunts de la forma A_i i A_i^c. Per exemple,

$$\omega = A_1 \cap \cdots \cap A_k \cap A_{k+1}^c \cap \cdots \cap A_n^c$$

correspon a una seqüència on els k primers elements són 1 i els $n-k$ restants, 0. En conseqüència, la propietat d'independència ens porta a assignar

$$\begin{aligned} P(\{\omega\}) &= P(A_1) \cdot \cdots \cdot P(A_k) \cdot P(A_{k+1}^c) \cdot \cdots \cdot P(A_n^c) \\ &= p^k (1-p)^{n-k}. \end{aligned}$$

Observeu que l'ordre en què apareixen els diversos 1 i 0 en la seqüència no intervé en el resultat de $P(\{\omega\})$. Noteu també que hi ha exactament $\binom{n}{k}$ seqüències amb k uns i $n-k$ zeros. En conseqüència,

$$P(\Omega) = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = 1.$$

Exemple 1.10 Juguem repetidament un joc en què es té una probabilitat $\frac{1}{N}$ de guanyar, independentment del resultat dels jocs previs. Volem determinar el nombre n mínim de vegades que hem de jugar per tal que la probabilitat de guanyar superi $\frac{1}{2}$.

És clar que n haurà d'ésser proporcional a N. Denotem per p la probabilitat de guanyar almenys una vegada en les n repeticions de la jugada. Utilitzant la hipòtesi d'independència entre les jugades resulta

$$p = 1 - \left(1 - \frac{1}{N}\right)^n.$$

Per tant, cal trobar el nombre natural n més petit possible tal que

$$\left(1 - \frac{1}{N}\right)^n < \frac{1}{2}. \tag{1.13}$$

Si N és petit, és fàcil comprovar per càlculs directes que n ha de ser l'enter més proper a $\frac{2}{3}N$. Per a N gran escrivim la condició (1.13) en la forma equivalent

$$n \log\left(1 - \frac{1}{N}\right) < \log \frac{1}{2}.$$

Desenvolupem en sèrie de Taylor la funció $\log(1+x)$ en un entorn de $x = 0$ fins a l'ordre 1 i rebutgem la resta. Substituïm x per $-\frac{1}{N}$. Resulta

$$\bar{n} = \frac{-\log 2}{\log(1 - 1/N)} \simeq \frac{\log 2}{1/N} \simeq \frac{2}{3}N.$$

Així, doncs, també trobem que n ha de ser l'enter més proper a $\frac{2}{3}N$.

Exercicis

1.1 *En una xarxa de tres ordinadors la probablitat que, en un moment donat, l'ordinador k falli és p_k, $k = 1, 2, 3$. Calculeu, sota les hipòtesis excloents (1) i (2), la probabilitat que en aquest instant almenys un dels ordinadors funcioni.*

1. *Els ordinadors es comporten de manera independent.*
2. *Hi ha una probabilitat p que el corrent elèctric falli; en aquest cas, tots els ordinadors deixen de funcionar. Si el corrent elèctric no falla, cadascun dels ordinadors es comporta de manera independent respecte dels altres.*

1.2 *Es posa a la venda una col·lecció de cromos, a un preu de* 10 *pessetes la unitat, que consisteix en* 50 *exemplars diferents. Un petit co≪ eccionistaestàdisposatagastarelsseusestal ho*?

Indicació: Utilitzeu la fórmula (1.8).

1.3 *S'instal·la un programa antivirus en un ordinador. La probabilitat que l'ordinador tingui el virus detectable per l'antivirus és* 0,2. *Si l'ordinador té el virus, la probabilitat que l'antivirus el detecti val* 0,9. *Si l'ordinador no té el virus, la probabilitat que l'antivirus doni un missatge d'existència del virus és* 0,02. *Es vol conèixer:*

1. *La probabilitat que, si ha aparegut un missatge d'existència de virus, l'ordinador no tingui el virus.*

2. *La probabilitat que l'ordinador tingui el virus i l'antivirus no el detecti.*

3. *La probabilitat que, si no ha sortit cap missatge d'existència de virus, l'ordinador tingui el virus.*

1.4 *En una família de tres germans, A, B i C, un d'ells ha guanyat el primer premi de la loteria de Nadal, i cadascun d'ells té probabilitat $1/3$ d'haver-lo guanyat. El germà A pregunta al seu pare qui, de B i C, no ha guanyat tal premi. Si el pare respon, quin efecte té la informació sobre la probabilitat que A hagi guanyat? És que aquesta probabilitat és encara $1/3$ o ha augmentat?*

1.5 *Considerem un polígon convex de n costats. Un braç robòtic uneix a l'atzar dos vèrtexs del polígon. Quina és la probabilitat que dibuixi una diagonal?*

1.6 *Un grup format per $2n$ nois i $2n$ noies es separa en dos subgrups formats pel mateix nombre de persones. Calculeu la probabilitat que, en cada subgrup, hi hagi el mateix nombre de nois que de noies. Doneu un valor aproximat d'aquesta probabilitat per a $n = 6$.*

1.7 *Fa molts anys, vaig tenir un professor despistat que sovint portava els dos mitjons diferents. Vaig imaginar que cada matí, en vestir-se, obria un calaix on guardava els seus n parells diferents de mitjons desaparellats i n'agafava a l'atzar un grapat. Em vaig fer la pregunta següent:*

Suposant que el grapat contingués $2r$ mitjons ($r < n$), quina seria la probabilitat de no haver-ne agafat cap parella?

Vaig observar que al llarg de 30 classes, només en dues ocasions portava el mateix model de mitjó a cada peu. Suposant $r = 1$, quants mitjons, com a mínim, havia de tenir al calaix?

1.8 *Dues persones escriuen a l'atzar un nombre enter de dues xifres comprès entre 10 i 99.*

1. *Es repeteix l'experiència n vegades i es suposa que els resultats són independents entre si. Quina és la probabilitat $p(n)$ que les dues persones hagin escrit, almenys una vegada, el mateix nombre?*

2. *Quantes vegades cal repetir l'experiència per tal que $p(n)$ sigui superior o igual a $0,99$?*

1.9 *En una guarderia amb $n+1$ nens, un nen s'inventa un conte que explica a un altre nen. Aquest l'explica a un altre nen, i així successivament. A cada etapa el nen a qui se li explica el conte és triat aleatòriament. El conte s'explica r vegades. Es demana de calcular*

1. *la probabilitat que el conte no s'expliqui mai al nen que l'ha inventat,*
2. *la probabilitat que el conte no s'expliqui dues vegades al mateix nen.*

1.10 *Sigui* $(\Omega, \mathcal{A}, P)$ *un espai de probabilitat. Per a* $A, B \in \mathcal{A}$*, es defineix la* diferència simètrica de A, B *per*

$$A\Delta B = (A \cap B^c) \cup (A^c \cap B).$$

Demostreu

$$\begin{aligned} P(A\Delta C) &\leq P(A\Delta B) + P(B\Delta C), \\ P\big((A \cup B)\Delta(C \cup D)\big) &\leq P(A\Delta C) + P(B\Delta D), \end{aligned}$$

$A, B, C, D \in \mathcal{A}$.

Capítol 2

Variables i vectors aleatoris

Des del punt de vista matemàtic, l'espai mostral és incòmode de tractar, ja que pot ser un conjunt d'objectes bastant arbitrari. Habitualment, en voler estudiar un fenomen aleatori, estem interessats en algunes característiques numèriques associades als resultats. Per això, és convenient establir una correspondència entre els elements de l'espai mostral i conjunts de nombres, com $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$, o vectors, com $\mathbb{N}^n$, $\mathbb{Z}^n$, $\mathbb{R}^n$.

Un exemple ben familiar el proporcionen els jocs d'atzar, que han estat a l'origen del desenvolupament del càlcul de probabilitats. Aquests jocs consisteixen generalment en la repetició de fenòmens aleatoris, com poden ser llançar una o diverses monedes a l'aire, daus, fer girar ruletes, triar cartes, etc. El resultat de l'experiència —la realització de la jugada— té associat un guany (positiu o negatiu), és a dir, un resultat numèric. Ens pot interessar conèixer quin serà previsiblement el capital acumulat quan acabi el joc o al cap d'unes quantes jugades. En altres paraules, fer un estudi de l'assignació numèrica de la qual hem parlat abans. Dediquem aquest capítol, i també el següent, a aquestes qüestions. Distingirem el cas unidimensional (variables aleatòries) del multidimensional (vectors aleatoris) i ens centrarem en dos exemples importants, el cas discret i l'absolutament continu.

2.1 Variables aleatòries

Per a poder definir amb rigor la noció de variable aleatòria necessitem introduir la σ-àlgebra de Borel; la denotarem per $\mathcal{B}$ i és la σ-àlgebra generada pels conjunts oberts de $\mathbb{R}$, és a dir, la més petita de les σ-àlgebres de parts de $\mathbb{R}$ que contenen tots els conjunts oberts de $\mathbb{R}$ respecte a la topologia euclidiana.

Al llarg d'aquest capítol fixem un espai de probabilitat de referència $(\Omega, \mathcal{A}, P)$.

Definició 2.1 *Una* variable aleatòria *és una aplicació* $X : \Omega \to \mathbb{R}$ *que compleix la propietat següent:*

$$\forall B \in \mathcal{B}, X^{-1}(B) \in \mathcal{A}. \tag{2.1}$$

Una variable aleatòria proporciona, doncs, una assignació numèrica als elements de l'espai mostral. La condició (2.1), anomenada de *mesurabilitat*, significa que podrem calcular quantitats com ara $P\big(X^{-1}(B)\big) = P\{\omega : X(\omega) \in B\}$, qualsevol que sigui $B \in \mathcal{B}$.

A partir de la definició que acabem de donar sembla difícil comprovar que una aplicació $X : \Omega \to \mathbb{R}$ és una variable aleatòria. La dificultat és en el fet que no hi ha una descripció explícita de com són els elements de la σ-àlgebra $\mathbb{R}$. Semblaria, doncs, que la definició 2.1 no és gaire operativa. Però això no és així. En efecte, pot demostrar-se que la condició (2.1) d'aquesta definició pot substituir-se per alguna de les següents:

Per a tot $t \in \mathbb{R}$,

$$X^{-1}\Big((-\infty, t]\Big) \in \mathcal{A}, \tag{2.2}$$

$$X^{-1}\Big((-\infty, t)\Big) \in \mathcal{A}, \tag{2.3}$$

$$X^{-1}\Big((t, \infty)\Big) \in \mathcal{A}, \tag{2.4}$$

$$X^{-1}\Big([t, \infty)\Big) \in \mathcal{A}. \tag{2.5}$$

Aquesta afirmació es basa en els arguments següents: La condició (2.1) és equivalent si mateixa quan se substitueix "$\forall B \in \mathcal{B}$" per "$\forall B$ que pertanyi a una classe de conjunts que generi la σ-àlgebra $\mathcal{B}$". D'altra banda, la σ-àlgebra generada pels conjunts oberts de $\mathbb{R}$ coincideix amb la generada per les semirectes del tipus $(-\infty, t]$, o bé $(-\infty, t)$ o bé (t, ∞) o bé $[t, \infty)$, $t \in \mathbb{R}$. Aquí no entrarem en la demostració d'aquestes qüestions perquè corresponen a un curs més avançat de teoria de la probabilitat, però ho utilitzarem.

Clarament, l'aplicació $X : \Omega \longrightarrow \mathbb{R}$, $X(\omega) = c$, $c \in \mathbb{R}$, és una variable aleatòria, ja que, per a tot $B \in \mathcal{B}$, $X^{-1}(B)$ és igual a Ω o a $\emptyset$, segons que $c \in B$ o $c \notin B$.
Un altre exemple molt senzill de variable aleatòria és el que donem tot seguit.

Exemple 2.1 Sigui $A \in \mathcal{A}$. Definim $1\!\!1_A : \Omega \to \mathbb{R}$ per

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

És immediat comprovar que es tracta d'una variable aleatòria. En efecte, sigui $B \in \mathcal{B}$. Poden donar-se diverses possibilitats:

1. B no conté ni el 0 ni el 1,
2. B conté el 0 i el 1,
3. B conté el 0 però no el 1,
4. B conté el 1 però no el 0.

Aleshores, $X^{-1}(B)$ és, respectivament, $\emptyset$, Ω, A^c, A, que són conjunts de $\mathcal{A}$. La variable aleatòria $1\!\!1_A$ s'anomena la *funció indicatriu* del conjunt A. Observeu que, si $A = \Omega$, $1\!\!1_A \equiv 1$.

El conjunt de les variables aleatòries és tancat sota determinades operacions com la suma, el producte per escalars, el producte, el màxim i el mínim. Les proposicions següents precisen aquesta afirmació. Per simplificar les notacions, sovint no escriurem la dependència en ω. Així, $\{X \in B\}$ vol dir $\{\omega : X(\omega) \in B\}$.

Proposició 2.1 *Siguin X i Y dues variables aleatòries definides en l'espai de probabilitat $(\Omega, \mathcal{F}, P)$ i $a, b \in \mathbb{R}$. Aleshores,*

1. *$aX + bY$ i XY són també variables aleatòries.*
2. *$\max(X, Y)$ i $\min(X, Y)$ són variables aleatòries.*
3. *Si $Y : \Omega \longrightarrow (0, \infty)$, el quocient $\frac{X}{Y}$ és una variable aleatòria.*

Demostració: Utilitzarem alguna de les condicions (2.2) equivalents a la definició de variable aleatòria. En tot el que segueix fixem $t \in \mathbb{R}$.

És evident que aX és una variable aleatòria . En efecte, si $a = 0$, $aX \equiv 0$, que, segons hem vist abans, és una variable aleatòria. Suposem, doncs, $a \neq 0$; aleshores,

$$\{aX \leq t\} = \{X \leq t/a\}, \text{ si } a > 0,$$
$$\{aX \leq t\} = \{X \geq t/a\}, \text{ si } a < 0.$$

Com que X és una variable aleatòria, això dóna el resultat.

Per acabar la demostració de la primera part de (1) només falta veure que $X + Y$ és una variable aleatòria.

Es compleix que

$$\{X + Y < t\} = \cup_{q \in \mathbb{Q}} \big(\{X < q\} \cap \{Y < t - q\}\big).$$

Per tant, el conjunt $\{X + Y < t\}$ s'expressa com una reunió numerable de conjunts de la σ-àlgebra $\mathcal{A}$ i, en conseqüència, també és de $\mathcal{A}$.

Per a completar la demostració de l'apartat (1) utilitzem la identitat

$$XY = \frac{1}{2}\big((X + Y)^2 - X^2 - Y^2\big),$$

i provem que X^2 és una variable aleatòria. Això és una conseqüència de les igualtats

$$\{X^2 \leq t\} = \begin{cases} \emptyset, & \text{si } t < 0, \\ \{-\sqrt{t} \leq X \leq \sqrt{t}\}, & \text{si } t \geq 0. \end{cases}$$

Per a la demostració de (2) cal tenir en compte les igualtats següents:

$$\{\max(X, Y) \leq t\} = \{X \leq t\} \cap \{Y \leq t\},$$
$$\{\min(X, Y) \leq t\} = \{X \leq t\} \cup \{Y \leq t\}.$$

Deixem com a exercici al lector la demostració de l'apartat 3. □

De la proposició anterior es dedueix que el conjunt de les variables aleatòries té l'estructura d'espai vectorial sobre $\mathbb{R}$. D'altra banda, la relació d'ordre $X \le Y$ definida puntualment dota aquest conjunt d'una estructura d'ordre reticular compatible amb la d'espai vectorial. Resulta, doncs, que el conjunt de les variables aleatòries és un espai vectorial reticulat o *espai de Riesz*.

Proposició 2.2 *Sigui* $\{X_n, n \ge 1\}$ *una successió de variables aleatòries definides en el mateix espai de probabilitat* $(\Omega, \mathcal{A}, P)$. *Aleshores,* $\inf X_n$, $\sup X_n$, $\liminf X_n$, $\limsup X_n$ *són també variables aleatòries.*

Demostració: Els arguments són similars als de la demostració de l'apartat (2) de la proposició 2.1. Per a qualsevol $t \in \mathbb{R}$, es té que

$$\sup\{X_n \le t\} = \cap_n \{X_n \le t\}. \tag{2.6}$$

El segon membre de (2.6) és una intersecció numerable de conjunts de $\mathcal{A}$ i, per tant, pertany a $\mathcal{A}$.
La resta d'afirmacions són una conseqüència de les igualtats següents:

$$\begin{aligned} \inf_n X_n &= -\sup_n(-X_n), \\ \liminf_n X_n &= \sup_n \inf_{m\ge n} X_m, \\ \limsup_n X_n &= \inf_n \sup_{m \ge n} X_m. \end{aligned}$$

Això completa la demostració de la proposició. □

L'exemple 2.1 és un cas particular de *variable aleatòria simple*, que és la que té una representació del tipus

$$X(\omega) = \sum_{i=1}^{r} a_i \mathbb{1}_{A_i}(\omega), \tag{2.7}$$

on $a_i \in \mathbb{R}$ i $A_i \in \mathcal{A}$, $i = 1, \ldots, r$, formen una partició de Ω. Aquesta classe de variables aleatòries permeten d'entendre l'estructura de qualsevol variable aleatòria, d'acord amb la proposició següent.

Proposició 2.3 *Tota variable aleatòria no negativa, $X : \Omega \to [0, \infty)$, és límit creixent d'una successió de variables aleatòries simples positives $\{X_n, n \geq 1\}$.*

Demostració: Per a tot $n \geq 1$ definim

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

Com que X és una variable aleatòria, és clar que $\{X_n, n \geq 1\}$ és una successió de variables aleatòries simples positives. Comprovem que es tracta d'una successió creixent.

Fixem $\omega \in \Omega$, $n \geq 1$, i suposem que $X(\omega) \in [k2^{-n}, (k+1)2^{-n})$, per a algun $k = 1, \ldots, n2^n - 1$. Aleshores, $X_n(\omega) = k2^{-n}$ i

$$X_{n+1}(\omega) = \begin{cases} k2^{-n}, & \text{si } X(\omega) \in [k2^{-n}, (2k+1)2^{-(n+1)}), \\ (2k+1)2^{-(n+1)}, & \text{si } X(\omega) \in [(2k+1)2^{-(n+1)}, (k+1)2^{-n}). \end{cases}$$

Per tant, veiem que $X_n(\omega) \leq X_{n+1}(\omega)$. Es faria un raonament anàleg en el cas $X(\omega) \geq n$.

Finalment hem de provar que $\lim_{n\to\infty} X_n(\omega) = X(\omega)$ per a tot $\omega \in \Omega$. En efecte, fixat ω sigui $n_0 > X(\omega)$. Aleshores, per a tot $n \geq n_0$, es compleix $|X(\omega) - X_n(\omega)| \leq 2^{-n}$, que tendeix a zero quan n tendeix a infinit. □

Donada una variable aleatòria X, definim

$$X^+ = \sup(X, 0),$$
$$X^- = \sup(-X, 0),$$

que s'anomenen les parts *positiva* i *negativa* de X, respectivament. Per la proposició 2.1, apartat (2), tant X^+ com X^- són variables aleatòries. Clarament,

$$X = X^+ - X^-,$$

i

$$|X| = X^+ + X^-.$$

Aplicant la proposició 2.3 a X^+ i a X^- resulta el següent

Corol·lari 2.1 *Tota variable aleatòria X és límit d'una successió $\{X_n, n \geq 1\}$ de variables aleatòries simples que compleixen $|X_n| \leq |X|$ per a tot $n \geq 1$.*

2.2 Llei d'una variable aleatòria

En l'apartat anterior hem establert una assignació numèrica als elements de l'espai mostral o resultats de l'experiència aleatòria. En aquest apartat veurem com, a partir d'aquesta assignació, es pot transferir la probabilitat definida en $(\Omega, \mathcal{A})$.

Definició 2.2 *La* llei *d'una variable aleatòria X és la probabilitat sobre $(\mathbb{R}, \mathcal{B})$, que designarem per Q, definida de la manera següent:*

$$\forall B \in \mathcal{B}, Q(B) = P\big(X^{-1}(B)\big), \tag{2.8}$$

on $X^{-1}(B) = \{\omega, X(\omega) \in B\}$.

Veiem que, efectivament, la fórmula (2.8) defineix una probabilitat sobre $(\mathbb{R}, \mathcal{B})$. Clarament,

$$Q : \mathcal{B} \longrightarrow [0, 1].$$

D'altra banda, $Q(\mathbb{R}) = P\big(X^{-1}(\mathbb{R})\big) = 1$. Finalment, sigui $\{B_n, n \geq 1\}$ una co$\ll$ *ecciódeconjuntsdisjuntsde*B.*Aleshores*,

$$X^{-1}(\cup_{n \geq 1} B_n) = \cup_{n \geq 1} X^{-1}(B_n)$$

i els conjunts $X^{-1}(B_n)$ són disjunts. Per tant,

$$\begin{aligned}Q(\cup_{n\geq 1}B_n) &= P\big(X^{-1}(\cup_{n\geq 1}B_n)\big)\\ &= P\big(\cup_{n\geq 1} X^{-1}(B_n)\big) = \sum_{n\geq 1} P\big(X^{-1}(B_n)\big)\\ &= \sum_{n\geq 1} Q(B_n).\end{aligned}$$

En el llenguatge de la teoria de la mesura, la llei Q és la mesura imatge de P per la variable aleatòria X, és a dir, $P \circ X^{-1}$. Al llarg d'aquest llibre escriurem moltes vegades $P \circ X^{-1}$ en comptes de Q per a designar la llei d'una variable aleatòria X.

Així, doncs, la variable aleatòria X permet de passar de l'espai de probabilitat $(\Omega, \mathcal{A}, P)$ a l'espai de probabilitat *numèric* $(\mathbb{R}, \mathcal{B}, P \circ X^{-1})$.

En la resta d'aquest apartat estudiarem el concepte de funció de distribució, una noció que, com veurem, està fortament relacionada amb la llei d'una variable aleatòria.

Definició 2.3 *La* funció de distribució *associada a una variable aleatòria X és la funció $F : \mathbb{R} \to [0,1]$ definida per*

$$F(x) = P \circ X^{-1}\big((-\infty, x]\big). \tag{2.9}$$

És a dir, considerem la família de les semirectes $(-\infty, x]$, $x \in \mathbb{R}$, i el valor que pren $P \circ X^{-1}$ sobre aquesta família.

Donem, tot seguit, les propietats bàsiques de les funcions de distribució.

Proposició 2.4 *Sigui F la funció de distribució d'una variable aleatòria X. Es compleix que*

1. *F és creixent,*
2. *F és contínua per la dreta,*
3. *$\lim_{x\to\infty} F(x) = 1$ i $\lim_{x\to-\infty} F(x) = 0$.*

Demostració:

1. Siguin $x, y \in \mathbb{R}$, $x \leq y$. Aleshores $\{X \leq x\} \subset \{X \leq y\}$. Per la propietat de monotonia de la probabilitat, $P\{X \leq x\} \leq P\{X \leq y\}$, és a dir, $F(x) \leq F(y)$.

2. És equivalent demostrar la continuïtat per successions. Fixem $x \in \mathbb{R}$ i considerem una successió decreixent de nombres reals, $\{x_n, n \geq 1\}$, tal que $\lim_{n\to\infty} x_n = x$. La successió de conjunts $\{X \leq x_n\}$, $n \geq 1$, decreix i $\cap_{n=1}^{\infty}\{X \leq x_n\} = \{X \leq x\}$. Aleshores, utilitzant l'apartat (2) de la proposició 1.1 resulta que
$$\lim_{n\to\infty} F(x_n) = F(x),$$
que és el que volíem demostrar.

3. Considerem una successió creixent de nombres reals $\{x_n, n \geq 1\}$ tal que $\lim_{n\to\infty} x_n = \infty$. Considerem també la successió de conjunts associada, $\{X \leq x_n\}$, $n \geq 1$, que creix cap a $\{X < \infty\} = \Omega$. Aleshores, l'apartat (1) de la proposició 1.1 ens diu que
$$\lim_{x\to\infty} F(x) = P(\Omega) = 1.$$
Per a la demostració de l'altra propietat asimptòtica prenem una successió decreixent de nombres reals, $\{x_n, n \geq 1\}$ tal que $\lim_{n\to\infty} x_n = -\infty$. En aquest cas, els conjunts $\{X \leq x_n\}$, $n \geq 1$, formen una successió decreixent cap a $\emptyset$. Per tant, novament per la proposició 1.1 (2),
$$\lim_{x\to-\infty} F(x) = P(\emptyset) = 0.$$

Això acaba la demostració de la proposició. □

Tota funció monòtona té límits per l'esquerra i per la dreta en tot punt x. En particular, per a tota funció de distribució F i per a tot $x \in \mathbb{R}$ existeix $\lim_{y\to x^-} F(y)$ (el denotarem per $F(x-)$). Ara bé, en general, no es pot esperar que $F(x-) = F(x)$, per a tot $x \in \mathbb{R}$, és a dir, les funcions de distribució són funcions discontínues, amb discontinuïtats de primera espècie i, com que estan fitades, podem assegurar que el conjunt de punts de discontinuïtat és finit o *infinit numerable*.

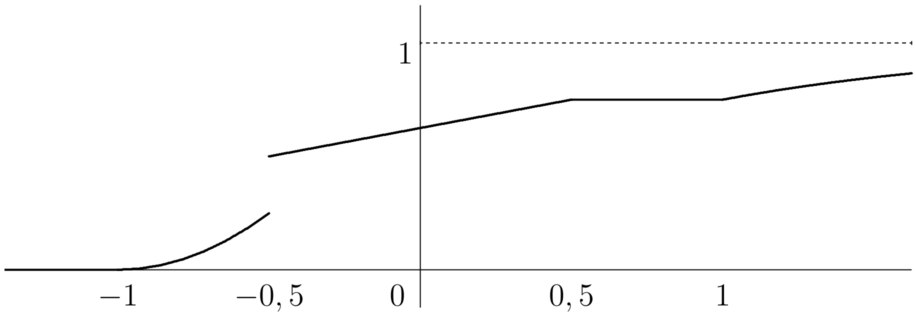

Figura 2.1: Gràfica d'una funció de distribució

La figura següent correspon a un exemple típic de funció de distribució:

Per a la variable aleatòria $\mathbb{1}_A$ definida en l'exemple 2.1, la seva funció de distribució té la gràfica que indiquem en la figura 2.2. En efecte, com que $\mathbb{1}_A$ no pren valors negatius, per a tot $x \in (-\infty, 0)$ es té $P\{X \leq x\} = 0$. Si $x \in [0, 1)$, aleshores $P\{X \leq x\} = P(A^c)$ i, finalment, si $x \in [1, \infty)$, $P\{X \leq x\} = P(A^c) + P(A) = 1$.

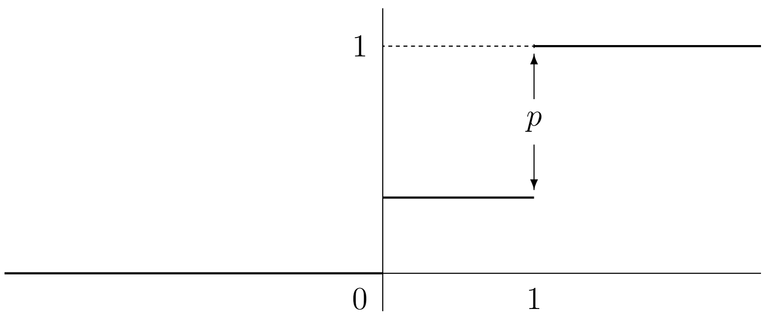

Figura 2.2: Gràfica de la funció de distribució d'una funció indicatriu

Amb el mateix tipus d'arguments justifiquem que la funció de distribució de la variable aleatòria simple definida per (2.7) té la gràfica que donem en la figura 2.3, on hem suposat $a_1 \leq \cdots \leq a_r$.

Mitjançant la funció de distribució es poden expressar les probabilitats que la variable aleatòria X prengui els seus valors en intervals diversos. Per

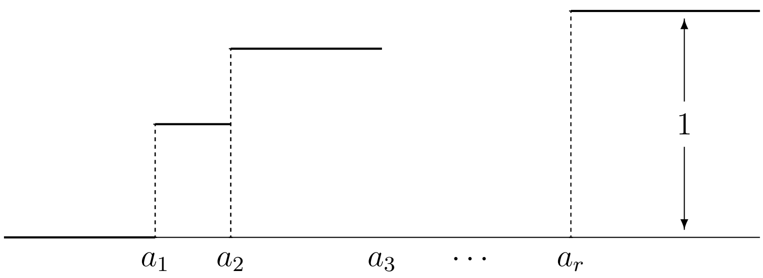

Figura 2.3: Gràfica de la funció de distribució d'una variable aleatòria simple

exemple, siguin $a, b \in \mathbb{R}$. Aleshores,

$$
\begin{aligned}
P\{X < a\} &= F(a-), \\
P\{a \leq X \leq b\} &= F(b) - F(a-), \\
P\{a < X \leq b\} &= F(b) - F(a), \\
P\{a \leq X < b\} &= F(b-) - F(a-), \\
P\{a < X < b\} &= F(b-) - F(a).
\end{aligned}
$$

Farem la demostració de les dues primeres d'aquestes identitats.
Sigui $\{a_n, n \geq 1\}$ una successió de nombres reals creixent cap a a. Aleshores, els conjunts $\{X \leq a_n\}$ formen una successió creixent cap a $\{X < a\}$. En conseqüència, per l'apartat (1) de la proposició 1.1, es té que

$$P\{X < a\} = \lim_{n \to \infty} P\{X \leq a_n\} = \lim_{n \to \infty} F(a_n) = F(a-).$$

Pel que fa a la segona identitat, cal considerar la descomposició del conjunt $\{X \leq b\}$ en una reunió de dos conjunts disjunts,

$$\{X \leq b\} = \{a \leq X \leq b\} \cup \{X < a\}.$$

Per la propietat d'additivitat de la probabilitat resulta que

$$F(b) = P\{a \leq X \leq b\} + F(a-),$$

que és el que es volia provar.

Les altres identitats es demostren de manera anàloga.

Per a les variables aleatòries simples hem vist que les seves funcions de distribució són funcions discontínues en els punts $a_1, \dots, a_r$. Més precisament, $F(a_k) - F(a_k-) = P\{X = a_k\}$, per a tot $k = 1, \dots, r$. Aquesta propietat és general, com pot veure's en la proposició següent.

Proposició 2.5 *Sigui X una variable aleatòria i F la seva funció de distribució. Per a tot $x \in \mathbb{R}$, es compleix que*

$$F(x) - F(x-) = P\{X = x\}.$$

Demostració: Sigui $\{x_n, n \geq 1\}$ una successió de nombres reals creixent cap a x. Aleshores, els conjunts $\{x_n < X \leq x\}$, $n \geq 1$, formen una successió decreixent i $\cap_{n=1}^{\infty}\{x_n < X \leq x\} = \{X = x\}$. En conseqüència,

$$\lim_{n\to\infty} P\big(\{x_n < X \leq x\}\big) = F(x) - F(x-) = P\{X = x\}.$$

Això acaba la demostració de la proposició. □

Observació 2.1 *Els comentaris que fem ara es basen en els teoremes d'extensió de mesures i no corresponen al nivell d'aquest llibre. Per tant, no en donem la demostració. Malgrat tot, creiem que és important per a entendre el paper que juga la funció de distribució i val la pena de remarcar-ho.*

Es pot demostrar que els valors de $P \circ X^{-1}(B), B \in \mathcal{B}$ queden determinats pels de $F(x), x \in \mathbb{R}$ en el sentit següent. Sigui $\mathcal{S}$ el conjunt de parts de $\mathbb{R}$ format pels intervals $(a, b]$, les semirectes (a, ∞), $(-\infty, a]$, $a, b \in \mathbb{R}$, $a < b$ i els conjunts $\emptyset$ i $\mathbb{R}$. Definim una funció de conjunt, μ sobre $\mathcal{S}$ posant

$$\begin{aligned}
&\mu\big((a, b]\big) = F(b) - F(a),\\
&\mu\big((a, \infty)\big) = 1 - F(a),\\
&\mu\big((-\infty, a]\big) = F(a),\\
&\mu(\emptyset) = 0, \mu(\mathbb{R}) = 1.
\end{aligned}$$

És fàcil veure que μ té la propietat d'additivitat finita i, per tant, la definició anterior permet d'estendre μ a l'àlgebra $\mathcal{C}$ formada per les reunions finites *de conjunts disjunts de $\mathcal{S}$. El que ja no és tan fàcil, però cert, és que μ és σ-additiva sobre $\mathcal{C}$. Existeix aleshores un procediment explícit (teorema d'extensió de Carathéodory) per a estendre μ a la σ-àlgebra generada per $\mathcal{C}$, que és $\mathcal{B}$. L'extensió de μ a $\mathcal{B}$ és la llei de X construïda a partir de la seva funció de distribució F.*

De fet, la definició de μ pot fer-se a partir d'una funció $F : \mathbb{R} \longrightarrow [0,1]$ que compleixi les propietats donades en la proposició 2.4, sense associar-la a cap variable aleatòria. El resultat serà una probabilitat definida en $(\mathbb{R}, \mathcal{B})$ que compleix $\mu(-\infty, a] = F(a)$, per a tot $a \in \mathbb{R}$ i, per aquesta raó, direm que F és la funció de distribució de la probabilitat μ.

Així, doncs, el procediment que hem descrit anteriorment ens proporciona una manera de construir probabilitats sobre $(\mathbb{R}, \mathcal{B})$ a partir d'una funció de distribució i, clarament, si dues probabilitats sobre $(\mathbb{R}, \mathcal{B})$, μ_1, μ_2, tenen la mateixa funció de distribució, també coincidiran sobre els conjunts de $\mathcal{S}$.

Acabem aquesta secció amb un resultat senzill sobre l'estructura de les funcions de distribució.

Proposició 2.6 *Sigui $F : \mathbb{R} \to [0,1]$ una funció creixent. Denotem per $\mathcal{D}$ el conjunt finit o numerable dels seus punts de discontinuïtat. Aleshores, val la descomposició següent:*

$$F(x) = F_c(x) + F_d(x),$$

per a tot $x \in \mathbb{R}$. On,

$$F_d(x) = \sum_{\{y \in \mathcal{D}, y \leq x\}} \big(F(y) - F(y-)\big),$$

i

$$F_c(x) = F(x) - F_d(x).$$

Les funcions F_c i F_d s'anomenen les parts contínua i discontínua de F, respectivament.

Observeu que la funció F_d és esglaonada i té el mateix aspecte que la funció de distribució d'una variable aleatòria simple. A més, se satisfà que

$$\lim_{x\to\infty} F_d(x) \leq 1,$$

però no podem assegurar que el límit anterior sigui 1.

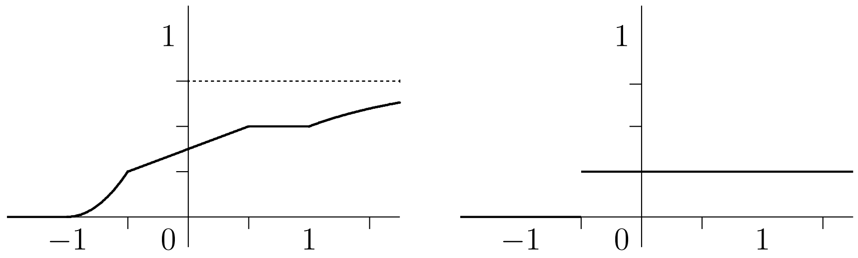

Figura 2.4: Gràfiques de les parts contínua i discontínua de la funció de distribució de la figura 2.1

2.3 Variables aleatòries discretes

La proposició 2.6 suggereix l'estudi de variables aleatòries de dos tipus: aquelles per a les quals la seva funció de distribució és, com F_d, esglaonada i les que tenen funció de distribució contínua. En aquest apartat estudiem les del primer tipus.

Definició 2.4 *Una variable aleatòria X direm que és discreta si la seva llei està concentrada en un conjunt numerable $B_0 \in \mathcal{B}$.*

Les variables aleatòries simples (vegeu l'equació (2.7)) són discretes. En aquest cas, el conjunt B_0 és $\{a_1, \ldots, a_r\}$, que és finit.
Més generalment, sigui $\{a_i, i \in \mathbb{N}\}$ una successió de nombres reals i $\{A_i, i \in \mathbb{N}\}$ conjunts de $\mathcal{A}$ que formen una partició de Ω. Aleshores,

$$X = \sum_{i\in\mathbb{N}} a_i \mathbb{1}_{A_i} \tag{2.10}$$

és una variable aleatòria discreta i $B_0 = \{a_i, i \in \mathbb{N}\}$.

La llei d'una variable aleatòria discreta queda determinada pels valors

$$P \circ X^{-1}(\{a_i\}), \quad i \in \mathbb{N},$$

ja que, per a qualsevol conjunt $B \in \mathcal{B}$, $P \circ X^{-1}(B) = \sum_{i:a_i \in B} P \circ X^{-1}(\{a_i\})$. L'aplicació $p : B_0 \longrightarrow [0,1]$, $p(a_i) = P \circ X^{-1}(\{a_i\})$ s'anomena funció de probabilitat.

La funció de distribució d'una variable aleatòria discreta és similar a la de les variables aleatòries simples: presenta salts en els punt $a_i, i \in \mathbb{N}$, d'amplitud igual a $P \circ X^{-1}(a_i)$ i és constant entre els dos instants de salt.

Exemple 2.2 Considerem el joc consistent en el llançament de dos daus. Ens interessa estudiar la suma dels punts obtinguts. Suposem que els dos llançaments són independents. Proposem com a espai mostral

$$\Omega = \{\omega = (\omega_1, \omega_2) : 1 \leq \omega_1, \omega_2 \leq 6\},$$

com a $\mathcal{A}$, la σ-àlgebra de parts de Ω, i, finalment, com a P, la probabilitat que assigna a cada element de l'espai mostral el valor $\frac{1}{36}$. Definim les variables aleatòries $X_1(\omega) = \omega_1$ i $X_2(\omega) = \omega_2$. La variable aleatòria d'interès és $S = X_1 + X_2$. Per trobar la llei de S cal descriure tots els conjunts $\{S = n\}$, $n = 2, \ldots, 12$. Per exemple, $\{\omega : S(\omega) = 2\} = \{(1,1)\}$ i $Q\{2\} = \frac{1}{36}$, $\{\omega : S(\omega) = 3\} = \{(1,2),(2,1)\}$ i $Q\{3\} = \frac{2}{36}, \ldots, \{\omega : S(\omega) = 12\} = \{(6,6)\}$ i $Q\{12\} = \frac{1}{36}$.

2.4 Exemples de variables aleatòries discretes

En aquest apartat presentem alguns exemples molt habituals de variables aleatòries que, com veurem al llarg del curs, serveixen per modelitzar una gamma àmplia de fenòmens aleatoris.

2.4.1 Variable aleatòria de Bernoulli

Considerem una experiència aleatòria amb dos resultats possibles a_1 i a_2 amb probabilitats respectives p i $q = 1 - p$, $p \in (0,1)$. Per exemple, el llançament d'una moneda. Associem a aquesta experiència la variable aleatòria X definida per

$$X(a_1) = 1, \quad X(a_2) = 0.$$

La llei d'aquesta variable aleatòria és una probabilitat concentrada en els punts 1 i 0 amb intensitats p i q, respectivament. S'anomena *llei de Bernoulli de paràmetre* p, i es denota per $B(1, p)$.

2.4.2 Variable aleatòria binomial

Considerem aquí la repetició, n vegades de manera independent, d'una experiència aleatòria com la descrita en l'apartat anterior. En altres paraules, reconsiderem l'exemple 1.9. Ens interessa estudiar el nombre de vegades que hem obtingut el resultat a_1. L'espai mostral és

$$\Omega = \{\omega = (\omega_1, \cdots, \omega_n) : \omega_i = a_1 \text{ o } \omega_i = a_2; i = 1, 2, \ldots, n\}.$$

Fixem $\omega \in \Omega$ i sigui $n_1(\omega)$ el nombre de components a_1 en ω. Aleshores, els arguments explicitats en l'exemple 1.9 ens porten a assignar la probabilitat següent als elements de Ω,

$$P\big(\{\omega\}\big) = p^{n_1(\omega)}(1-p)^{n-n_1(\omega)}.$$

Definim $X : \Omega \to \{0, 1, \ldots, n\}$, $X(\omega) = n_1(\omega)$. La llei d'aquesta variable aleatòria és una probabilitat Q concentrada en $0, 1, \ldots, n$ tal que

$$Q\big(\{k\}\big) = P\{\omega : n_1(\omega) = k\} = \binom{n}{k} p^k (1-p)^{n-k}. \tag{2.11}$$

En efecte, el conjunt $\{\omega : n_1(\omega) = k\}$ té cardinal $\binom{n}{k}$ i cada element del conjunt té una probabilitat $p^k(1-p)^{n-k}$.

És evident, a partir del desenvolupament del binomi de Newton, que $\sum_{k=0}^{n} Q\big(\{k\}\big) = 1$. La llei de la variable aleatòria X s'anomena *llei binomial* amb paràmetres (n, p) i es denota per $B(n, p)$.

La llei binomial és utilitzada, per exemple, en models probabilístics per al *control de qualitat*. L'observació de la qualitat d'un producte porta a dos resultats: bo o defectuós. L'estimació del nombre de productes defectuosos serveix en el control del procés de producció.

2.4.3 Variable aleatòria geomètrica

Considerem una successió de repeticions independents d'una experiència aleatòria com la de l'apartat 2.4.1. Ens aturem la primera vegada que obtenim el resultat a_1 (o *cara* en el llançament de la moneda).

Designem per X la variable aleatòria que descriu el nombre de jugades fins que ha aparegut a_1 per primera vegada. L'espai mostral associat a aquesta experiència és

$$\Omega = \{\omega = (\omega_1, \ldots, \omega_n, \ldots) : \omega_i = a_1 \text{ o } \omega_i = a_2; i = 1, 2, \ldots\}.$$

La llei de X és una probabilitat Q concentrada en $\{1, 2, \ldots\}$ tal que

$$Q\big(\{k\}\big) = p(1-p)^{k-1}. \tag{2.12}$$

S'anomena *llei geomètrica* de paràmetre p. Observeu que $\sum_{k=1}^{\infty} Q\big(\{k\}\big) = 1$ car, pels resultats sobre la suma de la sèrie geomètrica,

$$\sum_{k=1}^{\infty} p(1-p)^{k-1} = p \sum_{k=0}^{\infty} (1-p)^k = 1.$$

2.4.4 Variable aleatòria binomial negativa

Considerem el mateix esquema que en l'apartat 2.4.3. Ens aturem la primera vegada que obtenim el resultat a_1 per n-èsima vegada. Denotem per X_n la variable aleatòria que descriu aquest instant, és a dir, el nombre de *jugades* necessàries per obtenir a_1 per n-èsima vegada.

La llei d'aquesta variable aleatòria és una probabilitat Q_n concentrada en el conjunt $\{n, n+1, \ldots\}$, donada per

$$Q_n\big(\{k\}\big) = \binom{k-1}{n-1} p^n (1-p)^{k-n}. \tag{2.13}$$

En efecte, una seqüència $(\omega_1, \ldots, \omega_{k-1}, \omega_k)$ favorable a l'esdeveniment que estudiem ha de ser de la forma següent: $\omega_k = a_1$ i en $(\omega_1, \ldots, \omega_{k-1})$ hi ha exactament "$n-1$"a_1, que poden distribuir-se de $\binom{k-1}{n-1}$ maneres possibles, i "$k-n$"a_2.

La probabilitat Q_n s'anomena *llei binomial negativa*. Expliquem tot seguit la raó d'aquest nom.

Es defineix el nombre combinatori $\binom{-n}{j}$ mitjançant la relació següent

$$\binom{-n}{j} = \frac{(-n)(-n-1)\cdots(-n-j+1)}{j!},$$

amb el conveni que $\binom{-n}{0} = 1$. Aleshores,

$$Q_n(\{k\}) = \binom{-n}{k-n} p^n (-q)^{k-n}.$$

Vegem que es tracta d'una distribució de probabilitat. Posem

$$\sum_{k=n}^{\infty} Q_n(\{k\}) = \sum_{k=0}^{\infty} \binom{-n}{k} p^n (-q)^k = p^n \sum_{k=0}^{\infty} \binom{-n}{k} (p-1)^k.$$

Es té que

$$\sum_{k=0}^{\infty} \binom{-n}{k} (p-1)^k = p^{-n}.$$

En efecte, sigui $f(x) = \frac{1}{(1-x)^n}$, $|x| < 1$. Fent un desenvolupament per Taylor en un entorn de $x = 0$ obtenim:

$$\frac{1}{(1-x)^n} = \sum_{k=0}^{\infty} \frac{n(n+1)\cdots(n+k-1)}{k!} x^k.$$

Per a $x = 1 - p$, això és

$$p^{-n} = \sum_{k=0}^{\infty} \frac{n(n+1)\cdots(n+k-1)}{k!} (1-p)^k = \sum_{k=0}^{\infty} \binom{-n}{k} (p-1)^k.$$

En conseqüència,

$$\sum_{k=n}^{\infty} Q_n(\{k\}) = 1.$$

2.4.5 Variable aleatòria hipergeomètrica

Considerem una urna que conté N boles de les quals D són negres i $N-D$, blanques. Això podria ser un model per a l'estudi de poblacions de N individus amb dues característiques diferents. Fem un mostreig sense reposició, ço és, extraiem simultàniament n boles de l'urna. Ens interessa descriure la llei de la variable aleatòria X que dóna el nombre de boles negres que han sortit. Observeu que els valors de $X(\omega)$ han de complir les restriccions següents

$$X(\omega) \leq (D \wedge n), \quad n - X(\omega) \leq (N-D) \wedge n,$$

o, equivalentment,

$$(n - N + D) \vee 0 \leq X(\omega) \leq (D \wedge n).$$

Denotem per $(n-N+D)^+$ el nombre $(n-N+D) \vee 0$. Aleshores, la variable aleatòria X pren valors enters k, amb $(n-N+D)^+ \leq k \leq (D \wedge n)$, i amb probabilitats

$$Q(\{k\}) = P\{X = k\} = \frac{\binom{D}{k}\binom{N-D}{n-k}}{\binom{N}{n}}. \tag{2.14}$$

En efecte, suposem que totes les possibles mostres de n elements tenen la mateixa probabilitat. Aleshores, $\binom{N}{n}$ és el nombre total de mostres amb n boles i $\binom{D}{k}\binom{N-D}{n-k}$ és el nombre total de mostres amb exactament k boles negres.

Comprovem, tot seguit, que els nombres $P\{X = k\}$ donats en (2.14) defineixen una probabilitat sobre el conjunt $\{(n-N+D)^+, (n-N+D)^+ + 1, \ldots, D \wedge n\}$. Una manera senzilla de veure això és la següent: si $t \in \mathbb{R}$, tenim que

$$\begin{aligned}
\sum_{n=0}^{N} t^n \binom{N}{n} &= (1+t)^N = (1+t)^D (1+t)^{N-D} \\
&= \left[\sum_{k=0}^{D} t^k \binom{D}{k}\right] \left[\sum_{\ell=0}^{N-D} t^\ell \binom{N-D}{\ell}\right] = \sum_{k=0}^{D} \sum_{\ell=0}^{N-D} t^{k+\ell} \binom{D}{k}\binom{N-D}{\ell}.
\end{aligned}$$

En conseqüència, igualant coeficients s'obté que

$$\binom{N}{n} = \sum_{\substack{\{(k,\ell):k+\ell=n\\ 0\le k\le D\\ 0\le \ell\le N-D\}}} \binom{D}{k}\binom{N-D}{\ell} = \sum_{k=(n-N+D)^+}^{D\wedge n} \binom{D}{k}\binom{N-D}{n-k}.$$

Per tant,

$$\sum_{k=(n-N+D)^+}^{D\wedge n} \frac{\binom{D}{k}\binom{N-D}{n-k}}{\binom{N}{n}} = 1,$$

que és el que volíem demostrar.

La variable aleatòria X amb llei Q definida per (2.14) s'anomena *hipergeomètrica*. Observeu que per descriure (2.14) necessitem els valors dels paràmetres n, N i D.

Si l'extracció de les boles s'hagués fet amb reposició, la nova variable aleatòria, Y, que dóna el nombre de boles negres que han sortit, té una distribució binomial $B(n, \frac{D}{N})$ i, en conseqüència,

$$P\{Y = k\} = \binom{n}{k}\left(\frac{D}{N}\right)^k \left(\frac{N-D}{n}\right)^{n-k}.$$

És fàcil veure que, quan N, D i $N - D$ tendeixen a infinit, el segon membre d'aquesta desigualtat i el de (2.14) són equivalents.

2.4.6 Variable aleatòria de Poisson

Fixem un valor $\lambda \in (0, \infty)$. Una variable aleatòria X amb valors en els enters no negatius $\mathbb{Z}_+$ és dita de *Poisson* amb paràmetre λ si

$$P\{X = k\} = \mathrm{e}^{-\lambda}\frac{\lambda^k}{k!}. \tag{2.15}$$

Demostrem primer que (2.15) defineix una probabilitat. En efecte, això és una conseqüència immediata del desenvolupament en sèrie de Taylor de la funció $\exp\lambda$,

$$\sum_{k=0}^{\infty} P\{X = k\} = \mathrm{e}^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!} = \mathrm{e}^{-\lambda}\mathrm{e}^{\lambda} = 1.$$

Històricament la llei de Poisson s'ha associat a esdeveniments rars o poc freqüents (suïcidis infantils a Prússia, cavallers prussians morts per una guitza de cavall... són exemples curiosos recollits en un clàssic de Keynes editat el 1921). Altres exemples d'aplicació cal trobar-los en l'estudi de processos de recompte.

Matemàticament, la distribució de Poisson definida en (2.15) pot obtenir-se mitjançant un pas al límit en l'esquema que ens ha portat a definir la llei binomial. En efecte, considerem una successió de lleis binomials $B(n, p_n)$ i suposem que els valors del paràmetre p_n decreixen de manera que $\lim_n np_n = \lambda > 0$. Observeu que això implica $\lim_n p_n = 0$. Aleshores, per a tot nombre natural $k \geq 0$ es té que

$$\begin{aligned}\lim_n B(n, p_n)(\{k\}) &= \lim_n \binom{n}{k} p_n^k (1-p_n)^{n-k} \\ &= \lim_n \frac{n!}{k!(n-k)!} p_n^k (1-p_n)^{n-k} \\ &= \lim_n \frac{(np_n)^k}{k!} \frac{n!}{n^k (n-k)!} [(1-p_n)^{1/p_n}]^{np_n} (1-p_n)^{-k} = \frac{\lambda^k}{k!} \mathrm{e}^{-\lambda},\end{aligned}$$

ja que $\lim_n (1-p_n)^{-k} = 1$ i $\lim_n \frac{(n-k+1)\ldots n}{n^k} = 1$. La llei de Poisson de paràmetre λ serà denotada per $\mathrm{Poiss}(\lambda)$.

2.5 Variables aleatòries absolutament contínues

Les variables aleatòries que tenen una funció de distribució contínua, anomenades *variables aleatòries contínues*, compleixen que

$$P\{X = x\} = 0, \forall x \in \mathbb{R},$$

com es dedueix de la proposició 2.5. En aquesta secció introduïm una subclasse del conjunt de les variables aleatòries contínues a la qual pertanyen molts dels exemples més àmpliament utilitzats.

Definició 2.5 *Una funció $f : \mathbb{R} \to \mathbb{R}$ s'anomena una* densitat *si compleix les condicions següents:*

1. *$f \geq 0$,*

2. *f és integrable (en el sentit de Riemann) en $\mathbb{R}$,*

3. *es té que*

$$\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = 1.$$

Amb la noció de densitat podem introduir el concepte de variable aleatòria absolutament contínua.

Definició 2.6 *Es diu que una variable aleatòria X és* absolutament contínua *(o té llei absolutament contínua) amb* densitat *f si la seva funció de distribució F es pot escriure com*

$$F(x) = \int_{-\infty}^{x} f(y)\,\mathrm{d}y, \tag{2.16}$$

per a tot $x \in \mathbb{R}$, on la funció f satisfà les condicions de la definició 2.5.

La noció introduïda en la definició anterior és una propietat de la funció de distribució. Segons hem comentat en l'observació 2.1, la funció de distribució caracteritza la llei de la variable aleatòria. Per aquest motiu parlem indistintament de variable aleatòria o llei absolutament contínua.

A partir de (2.16) és immediat comprovar que

$$\begin{aligned} P\{X \in (a,b]\} &= P\{X \in (a,b)\} = P\{X \in [a,b)\} \\ &= P\{X \in [a,b]\} = \int_a^b f(x)\,\mathrm{d}x. \end{aligned}$$

Cal fer notar que no tota variable aleatòria contínua és absolutament contínua. Hi ha contraexemples interessants, però complicats, que utilitzen en la seva construcció el conjunt de Cantor. El lector interessat en aquesta qüestió pot consultar la referència [Ch 2].

Tot seguit, donem un bloc de condicions suficients perquè una funció de distribució F sigui absolutament contínua i per determinar-ne la densitat.

Proposició 2.7 *Suposem que la funció de distribució F compleix*

1. *és contínua en tots els punts,*
2. *és derivable, excepte, potser, en un nombre finit de punts,*
3. *la derivada F' és contínua, excepte, potser, en un nombre finit de punts.*

Aleshores, la funció F és absolutament contínua. La seva densitat és la funció f que coincideix amb F' en els punts on aquesta derivada existeix. En els punts on F' no està definida, podem donar a f qualsevol valor.

Aquesta proposició és una conseqüència de l'anomenat *teorema fonamental del càlcul*.

2.6 Exemples de variables aleatòries absolutament contínues

En aquest apartat fem una descripció de les variables aleatòries absolutament contínues que es presenten més freqüentment.

2.6.1 Llei uniforme

Siguin $a, b \in \mathbb{R}$, $a < b$. Definim

$$f(x) = \frac{1}{b-a} \mathbb{1}_{[a,b]}(x). \tag{2.17}$$

És evident que (2.17) defineix una funció de densitat. La funció de distribució associada és

$$F(x) = \int_{-\infty}^{x} f(y)\,\mathrm{d}y = \begin{cases} 0, & \text{si } x \leq a, \\ \frac{x-a}{b-a}, & \text{si } x \in [a,b], \\ 1, & \text{si } x \geq b. \end{cases}$$

Aquest exemple correspon a una distribució de probabilitat amb densitat *constant*, és a dir, *uniforme* en un interval de la recta real. S'anomena la *distribució uniforme* en $[a, b]$ i es designa per $\mathcal{U}[a, b]$. Una imatge física d'aquesta distribució seria la d'una massa igual a una unitat repartida segons una densitat constant. Un exemple aleatori fóra el de l'elecció d'un nombre a l'atzar entre a i b. Pot considerar-se com l'exemple més senzill de distribució amb densitat.

$F(x)$ $f(x)$

1 1

0 1 0 1

Figura 2.5: Funcions de distribució i de densitat d'una llei uniforme $\mathcal{U}[0,1]$

2.6.2 Llei exponencial

Fixem un nombre real $\lambda > 0$. La llei *exponencial* amb paràmetre λ, que denotarem per $\exp(\lambda)$, correspon a la funció de densitat següent:

$$f(x) = \lambda \exp(-\lambda x) \mathbb{1}_{[0,\infty)}(x). \tag{2.18}$$

La funció de distribució associada a aquesta densitat és

$$F(x) = \begin{cases} 0, & \text{si } x \leq 0, \\ 1 - \exp(-\lambda x), & \text{si } x > 0. \end{cases}$$

Un model per a la llei exponencial és el d'una variable aleatòria X positiva que descrigui el temps de vida d'un cert sistema amb les hipòtesis següents:

1. la funció de distribució $F(x)$, $x \geq 0$, és contínua i no nu$\ll a$,

2. no hi ha envelliment. Més precisament, suposem que

$$P\{X \geq x + y / X \geq x\} = P\{X \geq y\}, \tag{2.19}$$

   per a tots $x, y \geq 0$.

Dit altrament, el temps de vida residual a l'instant x segueix la mateixa llei de probabilitat que el temps de vida des de l'instant inicial.

Denotem per $\varphi(x) = 1 - F(x)$. Aleshores, la condició (2.19) s'expressa de la forma equivalent

$$\varphi(x + y) = \varphi(x)\varphi(y).$$

Tenint en compte que $\varphi(0) = 1$, resulta que

$$\varphi(x) = \exp(-\lambda x),$$

amb $\lambda > 0$, car $\lim_{x\to\infty} \varphi(x) = 0$.

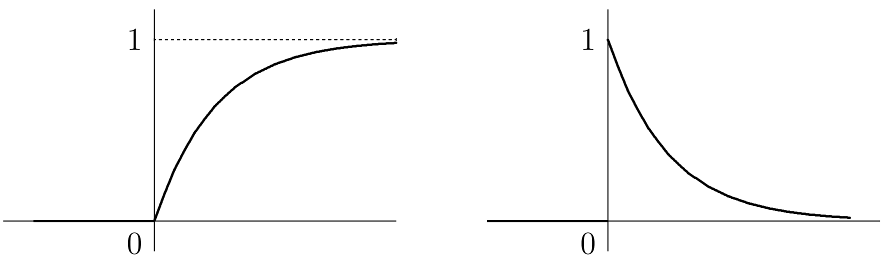

Figura 2.6: Funcions de distribució i de densitat d'una llei exp(1)

2.6.3 Llei normal estàndard

Aquest és, sense cap mena de dubte, l'exemple més important de llei absolutament contínua. El perquè d'aquest paper preponderant es descobrirà al capítol 6 d'aquest llibre. Aquí farem una descripció molt concisa d'aquest exemple, ja que més endavant anirem introduint tècniques que permetran d'analitzar diferents propietats d'aquesta distribució.

Considerem la funció

$$g(x) = \exp\left(-\frac{x^2}{2}\right),$$

$x \in \mathbb{R}$.

Se sap que $\int_{-\infty}^{\infty} g(x) = \sqrt{2\pi}$.
Considerem la funció de densitat $f(x) = \frac{1}{\sqrt{2\pi}} g(x)$, és a dir,

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\Big(-\frac{x^2}{2}\Big). \tag{2.20}$$

Una variable aleatòria amb distribució normal estàndard és la que té per funció de densitat la donada en (2.20).
Denotarem la llei normal estàndard per $\mathcal{N}(0,1)$.

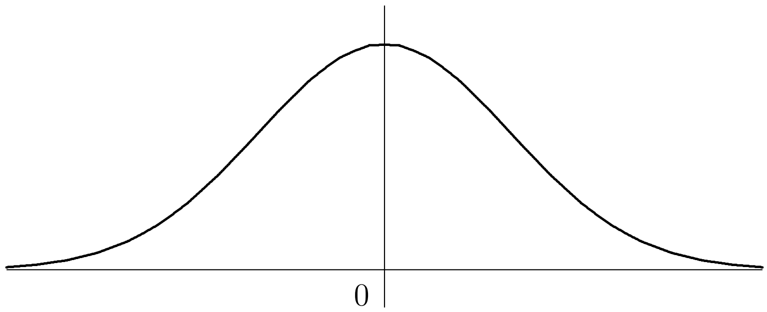

Figura 2.7: Funció de densitat de la llei $\mathcal{N}(0,1)$

2.6.4 Llei Gamma

Aquest exemple cobreix un ampli ventall de distribucions molt utilitzades en l'estadística.
Fixem $p \in (0,\infty)$. Es defineix la *funció gamma* mitjançant la integral

$$\Gamma(p) = \int_0^{\infty} x^{p-1} \mathrm{e}^{-x}\, \mathrm{d}x. \tag{2.21}$$

En particular, $\Gamma(1) = 1$ i $\Gamma(\frac{1}{2}) = \sqrt{\pi}$. Fixem ara $p \in (0,\infty)$ i $\lambda \in (0,\infty)$, i definim

$$f(x) = \begin{cases} \frac{\lambda^p}{\Gamma(p)} x^{p-1} \mathrm{e}^{-\lambda x}, & \text{si } x \geq 0, \\ 0, & \text{en cas contrari.} \end{cases}$$

Fent un canvi de variables, i utilitzant (2.21) és fàcil veure que $f(x)$ defineix una funció de densitat. La llei corresponent s'anomena la *llei gamma* amb paràmetres (p, λ) i es denota per $\mathrm{Gamma}(p, \lambda)$. Observeu que la llei exponencial amb paràmetre λ es pot també descriure com una llei $\mathrm{Gamma}(1, \lambda)$.

2.7 Densitat de transformacions de variables aleatòries

Sigui X una variable aleatòria que pren valors en un interval obert I finit o no. Sigui $g : I \to J$ una transformació bijectiva i de classe C^1 de l'interval I en un altre interval obert J.

Com que g és prou regular, pot demostrar-se que $Y = g(X)$ és una variable aleatòria. Suposem que la variable X té densitat $f_X(x)$. Ens podem preguntar si la variable Y és també absolutament contínua i, en cas afirmatiu, quina és la seva densitat. L'objectiu d'aquest apartat és resoldre aquesta qüestió.

La funció g introduïda anteriorment és estrictament monòtona. Per tant, pel teorema de la funció inversa en $\mathbb{R}$, la funció inversa g^{-1}, que també és estrictament monòtona, és de classe C^1 en J i, per a tot $y \in J$,

$$(g^{-1})'(y) = \frac{1}{g'(g^{-1}(y))}.$$

Demostrem tot seguit que Y té una densitat donada per

$$f_Y(y) = f_X(g^{-1}(y)) \left|(g^{-1})'(y)\right| = f_X(g^{-1}(y)) \frac{1}{|g'(g^{-1}(y))|}, \tag{2.22}$$

$y \in J$.

En efecte, suposem, per exemple, que g és creixent i que $I = (a, b)$, $J = (c, d)$, $-\infty \le a < b \le +\infty, -\infty \le c < d \le +\infty$, tindrem que

$$\begin{aligned} F_Y(y) = P\{Y \le y\} = P\{X \le g^{-1}(y)\} &= \int_a^{g^{-1}(y)} f_X(x)\,\mathrm{d}x \\ &= \int_c^y f_X\big(g^{-1}(z)\big)(g^{-1})'(z)\,\mathrm{d}z, \end{aligned}$$

on la darrera igualtat ha estat obtinguda fent el canvi de variable $x = g^{-1}(z)$. Per a g decreixent la demostració és anàloga i la deixem per al lector.

Exemple 2.3 Sigui X una variable aleatòria amb distribució uniforme en un interval $[a, b]$ on $a \geq 0$. Volem determinar la llei de probabilitat de la variable X^2. D'acord amb la fórmula (2.22) aquesta serà absolutament contínua i tindrà una densitat en l'interval $[a^2, b^2]$ donada per

$$f_{X^2}(y) = \frac{1}{2(b-a)\sqrt{y}},$$

$y \in [a^2, b^2]$.

Exemple 2.4 Considerem una variable aleatòria $\mathcal{N}(0, 1)$ (normal estàndard). Fixem dos paràmetres $m \in \mathbb{R}$ i $\sigma \in (0, \infty)$ i definim la variable aleatòria $Y = \sigma X + m$. Observeu que Y s'obté transformant la variable aleatòria X mitjançant la funció $g : \mathbb{R} \to \mathbb{R}$, $g(x) = \sigma x + m$. Podem, doncs, aplicar la fórmula (2.22) i obtenim

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y-m)^2}{2\sigma^2}\right). \tag{2.23}$$

Utilitzant el canvi de variable $z = \frac{y-m}{\sigma}$ és immediat comprovar que la funció f_Y defineix una densitat. Es diu que la variable aleatòria Y té llei normal amb paràmetres m, σ i aquesta llei es denota per $\mathcal{N}(m, \sigma^2)$.

La fórmula (2.22) admet una generalització que donem tot seguit.

Si el domini de g es pot escriure com una reunió finita d'intervals $I_1 \cup \cdots \cup I_n$ disjunts i la restricció de g a cada $\operatorname{Int} I_j, j = 1, \ldots, n$ és estrictament monòtona i de classe C^1, la densitat de la variable $Y = g(X)$ és

$$f_Y(y) = \sum_{i: y \in g_i(I_i)} f_X(g_i^{-1}(y)) \left|(g_i^{-1})'(y)\right|, \tag{2.24}$$

on g_i indica la restricció de la funció g a l'interval I_i, és a dir,

$$g(x) = \sum_{i=1}^{n} g_i(x) \mathbb{1}_{I_i}(x).$$

En efecte, posem $\operatorname{Int} I_i = (a_i, b_i)$, $i = 1, \ldots, n$, i fixem $y \in J$. Aleshores

$$\begin{aligned} F_Y(y) &= P\{Y \le y\} = P\{g(X) \le y, X \in \cup_{i=1}^n I_i\} \\ &= \sum_{i=1}^n P\{g(X) \le y, X \in I_i\} = \sum_{i=1}^n P\{g_i(X) \le y, X \in I_i\}. \end{aligned}$$

Sigui $i = 1, \ldots, n$, que compleix $y \in g_i(I_i)$. Considerem el canvi de variable donat per $z = g_i(x)$, $g_i : I_i \to g(I_i)$. Si g_i és creixent i $g(a_i) = c_i$, tenim que

$$\begin{aligned} &P\{g_i(X) \le y, X \in I_i\} = P\{a_i \le X \le g_i^{-1}(y)\} \\ &= \int_{a_i}^{g_i^{-1}(y)} f_X(x)\, \mathrm{d}x = \int_{g_i(a_i)}^{y} f_X\big(g_i^{-1}(z)\big)(g_i^{-1})'(z)\, \mathrm{d}z \\ &= \int_{g_i(a_i)}^{y} f_X\big(g_i^{-1}(z)\big)(g_i^{-1})'(z)\, \mathrm{d}z = \int_{c_i}^{y} f_X\big(g_i^{-1}(z)\big) \left|(g_i^{-1})'(z)\right| \mathrm{d}z. \end{aligned}$$

Si g_i és decreixent i $g(b_i) = d_i$,

$$\begin{aligned} &P\{g_i(X) \le y, X \in I_i\} = P\{g_i^{-1}(y) \le X \le b_i\} \\ &= \int_{g_i^{-1}(y)}^{b_i} f_X(x)\, \mathrm{d}x = \int_{y}^{g_i(b_i)} f_X\big(g_i^{-1}(z)\big)(g_i^{-1})'(z)\, \mathrm{d}z \\ &= \int_{g_i(b_i)}^{y} f_X\big(g_i^{-1}(z)\big) \left|(g_i^{-1})'(z)\right| \mathrm{d}z = \int_{d_i}^{y} f_X\big(g_i^{-1}(z)\big) \left|(g_i^{-1})'(z)\right| \mathrm{d}z. \end{aligned}$$

Si $y \notin g_i(I_i)$, ja que $g_i(I_i)$ és un interval, és fàcil veure que el terme $P\{g_i(X) \le y, X \in I_i\}$ és una constant que no depèn de y, per tant, la seva contribució a la funció de densitat serà zero, amb la qual cosa la fórmula (2.24) queda demostrada.

Exemple 2.5 Sigui X una variable aleatòria amb distribució $\mathcal{N}(0, 1)$. Considerem la funció $g : \mathbb{R} \to \mathbb{R}^+$ donada per $g(x) = x^2$. Volem determinar la llei de la variable aleatòria X^2.

Podem descompondre $\mathbb{R}$ en la reunió de dos intervals, $I_1 = (-\infty, 0]$ i $I_2 = (0, \infty)$. La restricció de g a cadascun d'aquests intervals estableix una bijecció de classe $\mathcal{C}^1$. Observeu que $g_1^{-1}(y) = -\sqrt{y}$ i $g_2^{-1}(y) = \sqrt{y}$. Aleshores, aplicant (2.24) amb $n = 2$ s'obté que

$$f_{X^2}(y) = \frac{1}{\sqrt{2\pi}} y^{-\frac{1}{2}} \exp\left(\frac{-y}{2}\right),$$

per a tot $y \in [0, \infty)$, que correspon a la funció de densitat d'una llei Gamma $(1/2, 1/2)$.

2.8 Vectors aleatoris

En l'apartat 2.1 d'aquest capítol hem introduït el concepte de variable aleatòria com una eina per a mesurar una característica numèrica d'un fenomen aleatori. Molt sovint pot interessar mesurar simultàniament diverses propietats, o bé pot ser que una propietat hagi de ser descrita amb diverses variables numèriques. Amb aquesta finalitat en aquest apartat donem la noció de *vector aleatori* i n'estudiem les propietats principals.

Definició 2.7 *Un* vector aleatori m-dimensional *és una aplicació* $X : \Omega \to \mathbb{R}^m$*,* $X = (X_1, \ldots, X_m)$*, tal que cadascun dels components* $X_i : \Omega \to \mathbb{R}$*,* $i = 1, \ldots, m$*, és una variable aleatòria (vegeu la definició 2.1).*

Per a definir la llei d'un vector aleatori cal introduir una σ-àlgebra associada a $\mathbb{R}^n$. No entrem aquí en el detall d'aquesta qüestió. En canvi, estudiarem la *funció de distribució* del vector aleatori X. Amb aquest objectiu considerem en $\mathbb{R}^m$ l'ordre parcial definit de la manera següent: si $x = (x_1, \ldots, x_m)$ i $y = (y_1, \ldots, y_m)$, aleshores

$$x \leq y \text{ si, i només si, } x_1 \leq y_1, \ldots, x_m \leq y_m,$$

Definició 2.8 *La* funció de distribució *d'un vector aleatori* X *és una aplicació* $F : \mathbb{R}^m \to [0, 1]$ *definida per*

$$F(x) = P\{X \leq x\} = P\{X_1 \leq x_1, \ldots X_m \leq x_m\}, \tag{2.25}$$

per a tot $x = (x_1, \ldots, x_m)$.

Cal fer notar que el conjunt $\{X \leq x\}$, $x \in \mathbb{R}$, és un esdeveniment, car és la intersecció dels conjunts $\{X_i \leq x_i\}$, $i = 1, \ldots, m$. Per tant, la definició donada en (2.25) té sentit.

La funció de distribució d'un vector aleatori té propietats similars a les presentades en la proposició 2.4 per a $m = 1$. Més precisament,

1. F és creixent,

2. F és contínua per la dreta en el sentit següent:

$$\lim_{y \downarrow x} F(y) = F(x),$$

   on $y \downarrow x$ significa $y_i \downarrow x_i$ per a tot $i = 1, \ldots, m$.

3. F té les propietats asimptòtiques

$$\lim_{x \to \infty} F(x) = 1, \qquad \lim_{x_i \to -\infty} F(x) = 0,$$

   per a tot $i = 1, \ldots m$.

   Les demostracions d'aquestes propietats són similars a les corresponents per a $m = 1$ i es deixen al lector.

   La propietat següent és específica del cas multidimensional.

4. Siguin $a, b \in \mathbb{R}^m$, $a \leq b$. Definim

$$\begin{aligned} \Delta_{a,b} F = & \sum_{\substack{\varepsilon \in \{0,1\}^m \\ \varepsilon = (\varepsilon_1, \ldots, \varepsilon_m)}} (-1)^{\varepsilon_1 + \cdots + \varepsilon_m} \\ & \times F(b_1 + \varepsilon_1(a_1 - b_1), \ldots, b_m + \varepsilon_1(a_m - b_m)), \end{aligned} \tag{2.26}$$

   $a = (a_1, \ldots, a_m)$, $b = (b_1, \ldots, b_m)$. Aleshores,

$$\Delta_{a,b} F = P(\{X \in (a, b]\}), \tag{2.27}$$

   on $(a, b]$ denota el rectangle de $\mathbb{R}^m$ definit per

$$(a, b] = \{x \in \mathbb{R}^m, a_i < x_i \leq b_i, i = 1, \dots, m\},$$

i, en conseqüència, $\Delta_{a,b}F \geq 0$.

Demostrem (2.27) per a $m = 2$. Per la fórmula (2.26),

$$\begin{aligned}\Delta_{a,b}F &= F(b_1, b_2) - F(b_1, a_2) - F(a_1, b_2) + F(a_1, a_2)\\ &= P\{-\infty < X_1 \leq b_1, a_2 < X_2 \leq b_2\} - P\{-\infty < X_1 \leq a_1, a_2 < X_2 \leq b_2\}\\ &= P(\{X \in (a, b]\}).\end{aligned}$$

La proposició que donem tot seguit estableix una relació entre la funció de distribució d'un vector aleatori X i les funcions de distribució de les variables aleatòries components. Designarem per F_i la funció de distribució de X_i, $i = 1, \cdots, m$, que s'anomena *funció de distribució marginal i-èsima.*

Proposició 2.8 *Per a tot $i = 1, \dots, m$,*

$$F_i(x_i) = \lim_{x_j \to \infty, j \neq i} F(x_1, \dots, x_m).$$

Demostració: Quan $x_j \to \infty, j \neq i$, els conjunts $\{X_1 \leq x_1, \dots, X_m \leq x_m\}$ creixen cap al conjunt $\{X_i \leq x_i\}$. La propietat de continuïtat seqüèncial de la probabilitat permet de concloure la demostració. □

Com en el cas de les variables aleatòries, les classes més importants de vectors aleatoris són els discrets i els absolutament continus.

Definició 2.9 *Un vector aleatori X s'anomena discret si les variables aleatòries components són discretes, ço és, tenen la seva llei concentrada en un conjunt numerable $B_0 \in \mathcal{B}$.*

L'exemple següent generalitza la llei binomial introduïda en l'apartat 2.4.2.

Exemple 2.6 (Llei multinomial) Considerem una experiència aleatòria amb m resultats possibles $A_1, \dots, A_m$, de probabilitats $p_1, \dots, p_m$, respectivament, $\sum_{j=1}^m p_j = 1$. Repetim l'experiència n vegades de manera independent i denotem per X_i, $i = 1, \dots, m$, la variable aleatòria que dóna el nombre de vegades que s'ha produït el resultat A_i.

Clarament, cadascuna de les variables aleatòries X_i té llei $B(n, p_i)$. Considerem el vector aleatori discret $X = (X_1, \dots, X_m)$, que pren valors en el conjunt

$$\mathcal{C} = \left\{ (n_1, \dots, n_m) : n_i \in \mathbb{N}, i = 1, \dots, m, \sum_{i=1}^{m} n_i = n \right\}.$$

Un càlcul de combinatòria senzill demostra que, si $(n_1, \dots, n_m) \in \mathcal{C}$,

$$P\{X_i = n_1, \dots, X_m = n_m\} = \frac{n!}{n_1! \cdots n_m!} p_1^{n_1} \cdots p_m^{n_m}.$$

Es diu que el vector aleatori X, que hem descrit anteriorment, té llei *multinomial* amb paràmetres $n, p_1, \dots, p_m$ i escriurem $M(n; p_1, \dots, p_m)$. Observieu que per a $m = 2$ obtenim la llei $B(n, p_1)$.

Definició 2.10 *Una funció $f : \mathbb{R}^m \to \mathbb{R}$ és una* densitat *en $\mathbb{R}^m$ si es compleixen les condicions següents:*

1. $f \geq 0$,
2. *existeix la integral*

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

en el sentit de Riemann, i

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x)\, \mathrm{d}x = 1.$$

Definició 2.11 *Un vector aleatori X m-dimensional és* absolutament continu *(o té llei* absolutament contínua*) amb densitat f, si la seva funció de distribució F es pot escriure com*

$$F(x) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_m} f(x)\, \mathrm{d}x,$$

on f és una funció de densitat en $\mathbb{R}^m$, és a dir, compleix les condicions de la definició 2.10, i $x = (x_1, \dots, x_m)$.

En els punts $x \in \mathbb{R}^m$ on la funció de densitat f és contínua es té

$$f(x) = \frac{\partial^m F(x)}{\partial x_1 \cdots \partial x_m}.$$

Proposició 2.9 *Sigui X un vector aleatori amb llei absolutament contínua. Aleshores, cadascuna de les variables aleatòries components, X_i, $i = 1, \dots,$ m, són també absolutament contínues i les seves densitats respectives, que denotarem per f_i, s'expressen com*

$$f_i(y) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1 \cdots x_{i-1}, y, x_{i+1}, \cdots x_m) \, \mathrm{d}x_1 \cdots \mathrm{d}x_{i-1} \, \mathrm{d}x_{i+1} \cdots \mathrm{d}x_m, \tag{2.28}$$

on, per a $i = 1$, $x_1 = y$.

Demostració: Fixem $i = 1, \dots, m$ i $x \in \mathbb{R}$. Aleshores, pel teorema de Fubini,

$$\begin{aligned} P\{X_i \le x\} &= P\{X_i \le x, X_j \in \mathbb{R}, j \ne i\} \\ &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{x} \cdots \int_{-\infty}^{\infty} f(y) \, \mathrm{d}y \\ &= \int_{-\infty}^{x} \left(\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(y_1, \dots, y_m) \, \mathrm{d}y_1 \cdots \mathrm{d}y_{i-1} \, \mathrm{d}y_{i+1} \cdots \mathrm{d}y_m \right) \mathrm{d}y_i. \end{aligned}$$

D'on resulta que la variable aleatòria X_i té una densitat donada per (2.28). □

Les densitats f_i s'anomenen *densitats marginals*.

Cal fer notar que les densitats marginals no determinen la densitat del vector aleatori. Per exemple, densitats diferents poden donar densitats marginals idèntiques. En efecte, considerem les densitats en el pla definides per

$$f_1(x, y) = \begin{cases} \frac{1}{4}(1 + xy), & \text{si } -1 \le x, y \le 1, \\ 0, & \text{en cas contrari,} \end{cases}$$

$$f_2(x, y) = \begin{cases} \frac{1}{4}, & \text{si } (x, y) \in [-1, 1]^2, \\ 0, & \text{en cas contrari.} \end{cases}$$

En ambdós casos les marginals són

$$f_1(x) = f_2(x) = \begin{cases} \frac{1}{2}, & \text{si } -1 \leq x \leq 1, \\ 0, & \text{en cas contrari.} \end{cases}$$

Exemple 2.7 (Llei uniforme en un recinte de $\mathbb{R}^2$) Considerem el quadrat unitat en el pla, $T = [0,1]^2$. Definim

$$f(x,y) = \mathbb{1}_T(x,y). \tag{2.29}$$

És obvi que (2.29) defineix una funció de densitat en $\mathbb{R}^2$. Considerem un vector aleatori (X,Y) bidimensional amb aquesta funció de densitat. Aleshores, podem calcular la densitat de cadascun dels components. Per exemple,

$$f_1(x) = \begin{cases} \int_0^1 f(x,y)\,\mathrm{d}y = 1, & \text{si } x \in [0,1], \\ 0, & \text{en cas contrari.} \end{cases}$$

Trobaríem un resultat idèntic per a l'altra densitat marginal, f_2. És a dir, les variables aleatòries components tenen una distribució uniforme en $[0,1]$.

En general, donat un recinte fitat del pla, R, anomenarem *llei uniforme en R* la llei absolutament contínua que té una densitat donada per

$$f(x,y) = \frac{1}{|R|}\mathbb{1}_R(x,y). \tag{2.30}$$

$|R|$ denota l'àrea del recinte R. La constant $\frac{1}{|R|}$ s'hi ha de posar per assegurar que la integral de la funció f sobre $\mathbb{R}^2$ sigui 1.

Per exemple, considerem una llei uniforme en el triangle del pla que té per vèrtexs els punts $(0,0)$, $(1,0)$ i $(0,1)$. Denotem per T aquesta regió. La densitat és donada per $f(x,y) = 2\mathbb{1}_T(x,y)$. Les densitats marginals són $f_X(x) = 2(1-x)\mathbb{1}_{[0,1]}(x)$ i $f_Y(y) = 2(1-y)\mathbb{1}_{[0,1]}(y)$, respectivament. Observeu que, en aquest cas, a diferència del que passa amb la distribució uniforme en el quadrat unitat, $f(x,y) \neq f_X(x)f_Y(y)$.

Exemple 2.8 (Llei normal bidimensional) Introduïm aquí un cas particular d'aquesta llei. Considerem la funció $f : \mathbb{R}^2 \to \mathbb{R}$ definida per

$$f(x,y) = \frac{1}{2\pi} \exp\left(-\frac{x^2+y^2}{2}\right). \tag{2.31}$$

Aquesta funció pot escriure's com a producte de dues densitats de lleis $\mathcal{N}(0,1)$ (vegeu (2.20)). Per aquest motiu és senzill comprovar que (2.31) defineix una densitat. En efecte,

$$\begin{aligned}&\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,\mathrm{d}x\,\mathrm{d}y \\ &\quad = \left(\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)\,\mathrm{d}x\right)\left(\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right)\,\mathrm{d}y\right) = 1.\end{aligned}$$

Observeu que les densitats marginals f_1, f_2 corresponen a lleis $\mathcal{N}(0,1)$. Un vector aleatori bidimensional que té una funció de densitat donada per (2.31) s'anomena *normal bidimensional amb paràmetres* $(0, Id)$ i es denota per $\mathcal{N}_2(0, Id)$. La notació Id significa la matriu identitat en $\mathbb{R}^2$.

2.9 Densitat de transformacions de vectors aleatoris

En aquest apartat fem una extensió al cas multidimensional dels resultats de la secció 2.7.

Sigui X un vector aleatori m-dimensional i $g : \mathbb{R}^m \to \mathbb{R}^d$ una funció *contínua*. Aleshores $Y = g(X)$ és un vector aleatori d-dimensional. Suposem que X té llei absolutament contínua (densitat f_X) i considerem el cas particular següent: $m = d$, U i V són dos oberts de $\mathbb{R}^m$ i $g : U \to V$ és un *difeomorfisme*, ço és, una aplicació bijectiva de classe $\mathcal{C}^1$ amb funció inversa també de classe $\mathcal{C}^1$. Aleshores, el vector aleatori m-dimensional $Y = g(X)$ té llei absolutament contínua i la seva densitat f_Y és

$$f_Y(y) = f_X(g^{-1}(y))\,|J_{g^{-1}}(z)|\,. \tag{2.32}$$

En aquesta fórmula, $J_{g^{-1}}(z)$ designa el jacobià de la transformació g^{-1}, és a dir, el determinant de la matriu de les derivades parcials $\left(\frac{\partial g_i^{-1}}{\partial x_j}\right)_{i,j=1,\dots,m}$.

Exemple 2.9 Sigui X un vector aleatori m-dimensional absolutament continu. Considerem l'aplicació $g : \mathbb{R}^m \to \mathbb{R}^m$ definida per $g(x) = Ax + b$, on A és una matriu quadrada d'ordre m, invertible, i $b \in \mathbb{R}^m$. La funció inversa de g està definida per

$$g^{-1}(y) = A^{-1}(y - b).$$

La funció g és, doncs, un difeomorfisme. Sigui Y el vector aleatori $g(X)$. Aleshores, la fórmula (2.32) permet d'escriure la densitat de Y

$$f_Y(y) = f_X(A^{-1}(y - b)) \left|\det A\right|^{-1}. \tag{2.33}$$

En particular podem considerar un vector aleatori X bidimensional amb llei $\mathcal{N}_2(0, Id)$ (densitat donada per (2.31)). Una altra manera d'escriure la densitat de X és

$$f_X(x) = \frac{1}{2\pi} \exp\left(-\frac{\|x\|^2}{2}\right).$$

Considerem el vector aleatori $Y = AX + b$. Aleshores, per (2.33)

$$f_Y(y) = \frac{1}{2\pi|\det A|} \exp\left(-\frac{\|A^{-1}(y - b)\|^2}{2}\right).$$

Direm que Y té una llei normal bidimensional amb paràmetres (b, A) i ho denotarem per $\mathcal{N}(b, AA^t)$, on A^t designa la matriu transposada de A.

2.10 Variables aleatòries independents

En aquest apartat introduïm el concepte d'independència per a variables aleatòries. Intuïtivament, la independència de les variables aleatòries $X_1, \ldots,$ X_m voldrà dir que els valors presos per una variable qualsevol no afecta els de les altres. Demostrarem que, per a vectors aleatoris amb components independents, les lleis marginals determinen la llei conjunta. L'estudi d'aquesta noció es continuarà en el capítol següent.

Definició 2.12 *Siguin $X_1, \ldots, X_m$ variables aleatòries definides en un espai de probabilitat $(\Omega, \mathcal{A}, P)$. Direm que són* independents *si, per a conjunts qualssevol $B_1, \ldots, B_m \in \mathcal{B}$, es compleix que*

$$P\{X_1 \in B_1, \ldots, X_m \in B_m\} = \prod_{i=1}^{m} P\{X_i \in B_i\}. \tag{2.34}$$

en altres paraules, si els esdeveniments $\{X_1 \in B_1\}, \ldots, \{X_m \in B_m\}$ *són independents.*

En el cas de variables aleatòries discretes la condició d'independència, és a dir (2.34), s'escriu de la manera equivalent següent:
Per a qualsevol $x_1, \ldots, x_m \in \mathbb{R}$

$$P\{X_1 = x_1, \ldots, X_m = x_m\} = \prod_{i=1}^{m} P\{X_i = x_i\} \tag{2.35}$$

En efecte, és evident que (2.34) implica (2.35). Recíprocament, siguin B_1, $\ldots, B_m \in \mathcal{B}$. Aleshores, per (2.35),

$$\begin{aligned} P\{X_1 \in B_1, \ldots, X_m \in B_m\} &= \sum_{x_1 \in B_1, \ldots, x_m \in B_m} P\{X_1 = x_1, \ldots, X_m = x_m\} \\ &= \sum_{x_1 \in B_1, \ldots, x_m \in B_m} P\{X_1 = x_1\} \cdots P\{X_m = x_m\} \\ &= \sum_{x_1 \in B_1} P\{X_1 = x_1\} \cdots \sum_{x_m \in B_m} P\{X_m = x_m\} \\ &= P\{X_1 \in B_1\} \cdots P\{X_m \in B_m\}. \end{aligned}$$

La igualtat (2.34) expressa la propietat que hem avançat abans:
Si $X = (X_1, \ldots, X_m)$ és un vector aleatori discret i les variables aleatòries $X_1, \ldots, X_m$ són independents, la llei de X queda determinada per les seves lleis marginals mitjançant la fórmula (2.35).
Sembla natural que les funcions de variables aleatòries independents produeixin vectors o variables aleatòries independents. En la proposició següent donem una prova d'aquest fet en un cas particular de variables aleatòries discretes.

Proposició 2.10 *Siguin* $X_1, \ldots, X_{m_1}, Y_1, \ldots, Y_{m_2}$ *variables aleatòries discretes, independents, i* $f : \mathbb{R}^{m_1} \to \mathbb{R}$, $g : \mathbb{R}^{m_2} \to \mathbb{R}$ *dues funcions. Aleshores, les variables aleatòries* $f(X_1, \ldots X_{m_1})$ *i* $g(Y_1, \ldots Y_{m_2})$ *són també independents.*

Demostració: Provarem un cas particular d'aquest resultat, per a simplificar l'exposició. Suposem que $m_1 = m_2 = 1$. Considerem dues variables aleatòries, X i Y, i utilitzem les notacions següents: $p(x,y) = P\{(X,Y) = (x,y)\}$, $p_1(x) = P\{X = x\}$ i $p_2(y) = P\{Y = y\}$. Aleshores, per a tot u i v de $\mathbb{R}$ es compleix que

$$P\{f(X) = u, g(Y) = v\} = P\{X \in f^{-1}(u), Y \in g^{-1}(v)\} = \sum_{\substack{x \in f^{-1}(u) \\ y \in g^{-1}(v)}} p(x,y)$$

$$= \sum_{\substack{x \in f^{-1}(u) \\ y \in g^{-1}(v)}} p_1(x)p_2(y) = \sum_{x \in f^{-1}(u)} p_1(x) \sum_{y \in g^{-1}(v)} p_2(y)$$

$$= P\{X \in f^{-1}(u)\}P\{Y \in g^{-1}(v)\} = P\{f(X) = u\}P\{g(Y) = v\}.$$

Això prova la propietat d'independència de les variables aleatòries $f(X)$ i $g(Y)$. □

Cal fer observar que, en l'enunciat de la proposició anterior, cal que f i g siguin prou regulars perquè $f(X_1, \dots X_{m_1})$ i $g(Y_1, \dots Y_{m_2})$ siguin variables aleatòries. Per exemple, això és així si són funcions contínues.

La independència dels components d'un vector aleatori dóna lloc a una forma especial de la seva funció de distribució.

Proposició 2.11 *Considerem un vector aleatori m-dimensional,* $X = (X_1, X_2, \dots, X_m)$*. Denotem per* F *la seva funció de distribució i per* F_i*,* $i = 1, \dots, m$ *les funcions de distribució respectives de les variables aleatòries components. Suposem que les variables* $X_1, X_2, \dots, X_m$ *són independents. Aleshores, es compleix que*

$$F(x_1, \dots, x_m) = F_1(x_1) \dots F_m(x_m), \tag{2.36}$$

per a tot $x = (x_1, \dots, x_m) \in \mathbb{R}$.

Demostració: Siguin $B_1, \dots, B_m \in \mathcal{B}$ borelians arbitraris de $\mathbb{R}$. Aleshores, la independència de les variables aleatòries $X_1, X_2, \dots, X_m$ implica que

$$P\{X_1 \in B_1, \dots, X_m \in B_m\} = P\{X_1 \in B_1\} \cdot \dots \cdot P\{X_m \in B_m\}. \tag{2.37}$$

En particular, prenent $B_1 = (-\infty, x_1], \ldots, B_m = (-\infty, x_m]$ obtenim que el primer terme de (2.37) coincideix amb el primer terme de (2.36) (vegeu (2.25)) i, anàlogament, el segon terme de (2.37) coincideix amb el segon de (2.36), per (2.9). $\square$

El recíproc de la proposició 2.11 és també cert. En altres paraules, la identitat (2.36) es pot considerar com una definició equivalent de la independència de les variables aleatòries $X_1, X_2, \ldots, X_m$. No donem la demostració d'aquest fet, encara que serà utilitzat al llarg d'aquesta secció. Això ens portaria a parlar de la construcció de probabilitats en espais producte, que cau fora dels objectius d'aquest llibre.

Exemple 2.10 Una aplicació senzilla de la proposició 2.11 permet de deduir la distribució del màxim i del mínim d'una família de variables aleatòries $\{X_1, X_2, \ldots, X_n\}$ independents. En efecte, sigui $Y = \max(X_1, \ldots, X_n)$ i $Z = \min(X_1, \ldots, X_n)$. Es compleix que

$$P\{Y \leq x\} = P\{X_1 \leq x, \ldots, X_n \leq x\} = F_1(x) \cdots F_n(x).$$

En particular, si les variables aleatòries $\{X_1, X_2, \ldots, X_n\}$ tenen la mateixa distribució, F,

$$F_Y(y) = \big(F(y)\big)^n.$$

De manera anàloga,

$$P\{Z \geq z\} = P\{X_1 \geq z, \ldots, X_n \geq z\} = \big(1 - F_1(z)\big) \cdots \big(1 - F_n(z)\big).$$

En el cas que el vector aleatori X sigui absolutament continu, la independència dels seus components es pot caracteritzar en termes de funcions de densitat.

Proposició 2.12 *Siguin $X_1, X_2, \ldots, X_m$ variables aleatòries. Suposem que el vector aleatori $X = (X_1, X_2, \ldots, X_m)$ té llei absolutament contínua amb densitat f i densitats marginals $f_1, \ldots, f_m$. Suposem que es compleix que*

$$f(x_1, x_2, \ldots, x_m) = f_1(x_1) f_2(x_2) \cdots f_m(x_m). \tag{2.38}$$

Aleshores, les variables $X_1, X_2, \ldots, X_m$ són independents.

Demostració: Sigui $y = (y_1, \dots, y_m) \in \mathbb{R}^m$. Denotem per F la funció de distribució del vector aleatori X. Aleshores, utilitzant (2.38) s'obté que

$$\begin{aligned} F(y) &= \int_{-\infty}^{y_1} \cdots \int_{-\infty}^{y_m} f(x_1, \dots, x_m)\, \mathrm{d}x_1 \cdots \mathrm{d}x_m \\ &= \int_{-\infty}^{y_1} f_1(x_1)\, \mathrm{d}x_1 \cdots \int_{-\infty}^{y_m} f_m(x_m)\, \mathrm{d}x_m \\ &= F_1(y_1) \cdots F_m(y_m). \end{aligned}$$

Per tant, es compleix (2.36) i, com hem comentat abans, aquesta condició és equivalent a la independència de les variables aleatòries $X_1, X_2, \dots, X_m$. □

Proposició 2.13 *Considerem un vector aleatori* $X = (X_1, X_2, \dots, X_m)$ *tal que els seus components són independents i tenen llei absolutament contínua amb densitats respectives* $f_1, \dots, f_m$. *Aleshores, la llei del vector* X *és també absolutament contínua i la seva densitat* f *és la següent:*

$$f(x_1, \dots, x_m) = f_1(x_1) \cdots f_m(x_m). \tag{2.39}$$

Demostració: Sigui F la funció de distribució del vector aleatori X. Com que estem suposant la independència dels components, podem aplicar la igualtat (2.36). Resulta, doncs,

$$\begin{aligned} F(x_1, \dots, x_m) &= \int_{-\infty}^{x_1} f_1(y)\, \mathrm{d}y \cdots \int_{-\infty}^{x_m} f_m(y)\, \mathrm{d}y \\ &= \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_m} f_1(y_1) \cdots f_m(y_m)\, \mathrm{d}y_1 \cdots \mathrm{d}y_m. \end{aligned}$$

Aquesta igualtat acaba la demostració de la proposició. □

La proposició 2.13 ens diu que, en el cas absolutament continu, si els components són independents, les lleis marginals determinen la llei conjunta.

Sense la condició d'independència, el resultat de la proposició 2.13 és fals. En efecte, considerem una variable aleatòria X_1 absolutament contínua i sigui $X_2 = X_1$. El vector aleatori (X_1, X_2) no pot ser absolutament continu, car si $D = \{(x, y) \in \mathbb{R}^2, x = y\}$, es té $P\{(X_1, X_2) \in D\} = 1$. En canvi, com que

D és un conjunt que té àrea zero, per a qualsevol possible funció de densitat $f(x,y)$ es compleix $\int_D f(x,y)\,\mathrm{d}x\,\mathrm{d}y = 0$.

Observació 2.2 *Considerem un vector aleatori amb llei absolutament contínua. Per simplificar les notacions suposarem que és bidimensional, (X,Y). Suposem que la seva densitat f factoritza, és a dir, existeixen dues funcions f_1 i f_2 tals que*

$$f(x,y) = f_1(x)f_2(y).$$

Aleshores, pot assegurar-se que els components del vector aleatori són variables aleatòries independents.
En efecte, la condició $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y) = 1$ implica

$$\left(\int_{-\infty}^{\infty} f_1(x)\,\mathrm{d}x\right)\left(\int_{-\infty}^{\infty} f_2(x)\,\mathrm{d}x\right) = 1.$$

Sigui $k = \int_{-\infty}^{\infty} f_2(x)\,\mathrm{d}x$. Les densitats marginals són

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f(x,y)\,\mathrm{d}y = f_1(x)\int_{-\infty}^{\infty} f_2(y)\,\mathrm{d}y = kf_1(x), \\ f_Y(y) &= \int_{-\infty}^{\infty} f(x,y)\,\mathrm{d}x = f_2(y)\int_{-\infty}^{\infty} f_1(x)\,\mathrm{d}x = \frac{1}{k}f_2(y). \end{aligned}$$

Per tant,

$$f(x,y) = f_X(x)f_Y(y).$$

Aquesta relació implica la independència, com hem demostrat en la proposició 2.12.

2.11 Distribucions condicionades

En la secció 1.5 d'aquest llibre hem introduït la noció de probabilitat condicionada al coneixement d'un esdeveniment. Aquí presentarem un concepte similar, però referit a les variables aleatòries dels dos tipus estudiats de manera exhaustiva en aquest capítol: les discretes i les absolutament contínues.

Considerem un vector aleatori (X, Y) discret, Definim $p(x, y) = P\{(X, Y) = (x, y)\}$. Recordem que $p(x, y)$, $x, y \in \mathbb{R}$, determina la llei de (X, Y). Anàlogament, sigui $p_Y(y)$ la funció de probabilitat de Y, que determina la llei de Y.

Volem definir la *llei de* X *condicionada a* Y, és a dir, la llei de X coneixent els valors que ha pres la variable aleatòria Y.

Definició 2.13 *La* llei de X condicionada a $\{Y = y\}$ *és la probabilitat sobre* $\mathbb{R}$ *definida per*

$$\bar{p}_X(x/Y = y) = \begin{cases} \frac{p(x,y)}{p_Y(y)}, & si\ p_Y(y) > 0, \\ 0, & en\ cas\ contrari. \end{cases}$$

De manera anàloga, podria definir-se la *llei d'Y condicionada a* $\{X = x\}$.

Si les variables aleatòries X i Y són independents, les lleis condicionades coincideixen amb les respectives marginals, a causa de la propietat (2.35).

La llei de X condicionada a Y i la llei de Y determinen la llei de X. Això és una conseqüència senzilla de la fórmula (1.9), car

$$P\{X = x\} = \sum_{y: p_Y(y) > 0} p(x, y) = \sum_y \bar{p}_X(x/Y = y) p_Y(y).$$

Anàlogament, la llei de Y condicionada a X i la llei de X determinen la llei de Y.

Exemple 2.11 Considerem una variable aleatòria X amb distribució Poiss(λ) i una altra variable aleatòria Y tal que la seva distribució condicionada a $\{X = k\}$ és $B(k, p)$. La llei de Y és una Poiss(λp).

En efecte, fixem un enter positiu l. Aleshores,

$$
\begin{aligned}
P\{Y=l\} &= \sum_{k=0}^{\infty} P\{Y=l/X=k\}P\{X=k\} \\
&= \sum_{k\geq l}^{\infty} \binom{k}{l} p^l (1-p)^{k-l} \mathrm{e}^{-\lambda} \frac{\lambda^k}{k!} \\
&= \mathrm{e}^{-\lambda} \frac{p^l}{l!} \lambda^l \sum_{k=l}^{\infty} \frac{1}{(k-l)!} (1-p)^{k-l} \lambda^{k-l} = \frac{(\lambda p)^l}{l!} \mathrm{e}^{-\lambda p}.
\end{aligned}
$$

Exemple 2.12 Siguin X, Y variables aleatòries independents amb distribucions $B(n_1, p)$, $B(n_2, p)$, respectivament. Trobem la llei de X condicionada per $\{X+Y=n\}$.

La variable aleatòria $X+Y$ té distribució $B(n_1+n_2, p)$. Fixem $n \in \{0, 1, 2, \ldots, n_1+n_2\}$ i $0 \leq k \leq n$. Aleshores,

$$
\begin{aligned}
P\{X=k/X+Y=n\} &= \frac{P\{X=k, X+Y=n\}}{P\{X+Y=n\}} = \frac{P\{X=k, Y=n-k\}}{P\{X+Y=n\}} \\
&= \frac{P\{X=k\}P\{Y=n-k\}}{P\{X+Y=n\}} = \frac{\binom{n_1}{k}\binom{n_2}{n-k}}{\binom{n_1+n_2}{n}}.
\end{aligned}
$$

Resulta, doncs que la llei de X condicionada per $\{X+Y=n\}$ és una *hipergeomètrica* amb paràmetres n, $D=n_1$ i $N=n_1+n_2$, seguint les notacions de la subsecció 2.4.5.

Considerem ara un vector aleatori amb llei absolutament contínua. Per simplificar les notacions, suposarem que és bidimensional, $Z=(X,Y)$. Com que Y té també llei absolutament contínua, la probabilitat d'esdeveniments de la forma $\{Y=y\}$ és zero. Heus aquí un exemple.

Considerem un vector aleatori $Z=(X,Y)$ amb distribució uniforme sobre el cercle de radi 1. La funció de densitat és donada per (2.30) on hem de canviar R pel conjunt $C=\{(x,y)\in\mathbb{R}^2 : x^2+y^2\leq 1\}$. Podem suposar que Z descriu l'elecció a l'atzar d'un punt del cercle, X descriu els valors de les abscisses i Y els de les ordenades. Suposem ara que fixem el valor de l'ordenada d'aquest punt. Quina és la distribució de la variable aleatòria que dóna l'abscissa del punt? Ens estem preguntant per la llei de X sabent que $Y=y$.

Una resposta a aquest problema pot donar-se introduint la noció de *densitat condicionada de X respecte d'Y* que definim de la manera següent:

$$f(x/y) = \frac{f(x,y)}{f_Y(y)}, \tag{2.40}$$

per a tot y tal que $f_Y(y) > 0$, on f designa la densitat conjunta del vector aleatori (X,Y), i f_Y la densitat marginal de X. Si $f_Y(y) = 0$, posem $f(x/y) = 0$.

Anàlogament, podríem definir la *densitat d'Y condicionada per X*.

Observeu que (2.40) determina la densitat conjunta a partir d'una condicionada i la corresponent marginal.

En l'exemple que hem plantejat més amunt, tenim que

$$f(x,y) = \frac{1}{\pi}\mathbb{1}_C(x,y), \qquad f_Y(y) = \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} \frac{1}{\pi}\,\mathrm{d}x = \frac{2}{\pi}\sqrt{1-y^2},$$

si $|y| \leq 1$ i $f_Y(y) = 0$ en cas contrari.

Per tant, per (2.40),

$$f(x/y) = \begin{cases} \frac{1}{2\sqrt{1-y^2}}, & \text{si } -\sqrt{1-y^2} \leq x \leq \sqrt{1-y^2}, \\ 0, & \text{en cas contrari.} \end{cases}$$

Obtenim, doncs, una distribució uniforme, aquesta vegada en l'interval $[-\sqrt{1-y^2}, \sqrt{1-y^2}]$.

Podem també considerar situacions mixtes. Per exemple, suposem que X és una variable aleatòria absolutament contínua i A un esdeveniment de probabilitat estrictament positiva. La *llei de X condicionada per A* pot definir-se a partir de la funció de distribució de la manera següent:

$$F(x/A) = P\{X \leq x/A\} = \frac{P(A \cap \{X \leq x\})}{P(A)}.$$

Si $A = \{X \in B\}$, on B és un conjunt de Borel de $\mathbb{R}$, l'expressió anterior dóna

$$\begin{aligned} F(x/A) &= \frac{P\{X \in B, X \leq x\}}{P(A)} = \frac{\int_{-\infty}^{x} \mathbb{1}_B(y) f_X(y)\,\mathrm{d}y}{P(A)} \\ &= \int_{-\infty}^{x} \mathbb{1}_B(x) \frac{f_X(x)}{P(A)}\,\mathrm{d}x. \end{aligned}$$

Trobem, doncs, que la llei de X condicionada per A és absolutament contínua amb densitat

$$f(x/A) = \begin{cases} \frac{f_X(x)}{P(A)}, & \text{si } x \in B, \\ 0, & \text{en cas contrari.} \end{cases}$$

Exercicis

2.1 *Una urna conté boles numerades des del* 1 *al* N*. Se'n fan* n *extraccions amb reposició i es designa per* X *la variable aleatòria que indica el nombre més gran obtingut. Trobeu la llei de* X*.*

2.2 *Es llancen repetidament i per separat dos daus. Sigui* X *el nombre de llançaments necessaris fins a obtenir un* 1 *en el primer dau i* Y *el nombre de llançaments necessaris per obtenir un* 5 *o un* 6 *amb el segon dau. Es demana:*

1. *les lleis de* X *i de* Y*,*
2. *la llei de* $Z = \max(X, Y)$*,*
3. *el valor de* $P\{X = Z\}$*.*

2.3 *Considerem l'espai de probabilitat* $(\Omega, \mathcal{A}, P)$ *definit per* $\Omega = \mathbb{N}$*,* $\mathcal{A} = \mathcal{P}(\Omega)$*,* $P(\{n\}) = 2^{-n}$*,* $n \in \mathbb{N}$*. Sigui* $k \in \mathbb{N}$ *i definim* $X : \Omega \longrightarrow \mathbb{R}$*,* $X(n) = n \pmod{k}$*. Trobeu la llei de la variable aleatòria* X*.*

2.4 *Considerem un conjunt de* n *punts en el pla tals que no n'hi ha* 3 *d'alineats. Per un procediment aleatori es dibuixen segments que uneixen dos punts qualssevol d'aquest conjunt. Aquests segments tenen probabilitat* p *d'aparèixer, independentment els uns dels altres. Es demana:*

1. *Probabilitat que apareguin exactament* k *segments.*

2. *Probabilitat que apareguin exactament* 3 *segments i formin un triangle.*

2.5 *Considerem k variables aleatòries independents, $X_1, \ldots, X_k$, totes amb distribució* $\text{Poiss}(\lambda)$. *Trobeu la llei condicionada del vector aleatori $(X_1, \ldots, X_k)$ sabent $X_1 + \ldots + X_k = n$.*

2.6 *Llancem una moneda tres vegades. Siguin X el nombre de cares que han sortit, Y el valor absolut de la diferència entre el nombre de cares i de creus. Trobeu:*

1. *La llei del vector aleatori (X, Y).*

2. *Les lleis marginals de X i de Y.*

2.7 *Considerem la funció real definida per*

$$F(x) = \begin{cases} 0, & si\ 0 \leq x < 1, \\ a(x-1)^2, & si\ 1 \leq x < 3, \\ b(2 - \mathrm{e}^{-(x-3)}), & si\ 3 \leq x. \end{cases}$$

1. *Quin ha de ser el valor de b per tal que F sigui una funció de distribució?*

2. *Fixat b, quins són els possibles valors de a?*

3. *Suposem $a = \frac{1}{16}$. És F una funció de distribució absolutament contínua?*

2.8 *Sigui X una variable aleatòria amb distribució* $\exp(\lambda)$. *Definim $Y_1 = \max(X, a)$, $Y_2 = \min(X, b)$, $a, b > 0$, $Y_3 = \max(a, Y_2)$, $b > a$. Trobeu les lleis de les variables aleatòries Y_i, $i = 1, 2, 3$.*

2.9 *Sigui (X, Y) un vector aleatori amb densitat uniforme sobre el quadrat amb vèrtexs $(0,1)$, $(1,0)$, $(0,-1)$, $(-1,0)$. Determineu:*

1. *Les densitats respectives de les variables aleatòries X i Y.*

2. *El valor de $P\left\{Y > \frac{1}{2} / X < \frac{1}{2}\right\}$.*

3. *La densitat de $X + Y$.*

Demostreu finalment que les variables aleatòries $X + Y$ i $X - Y$ són independents.

2.10 *Es tria un punt X sobre l'eix d'abscisses amb distribució $\mathcal{U}[0,1]$ i un punt Y sobre l'eix d'ordenades amb la mateixa distribució i independentment de X. Sigui Z l'àrea del triangle determinat per X, Y i l'origen de coordenades. Trobeu la densitat de la variable aleatòria Z.*

2.11 *Siguin X i Y dues variables aleatòries independents amb llei $\mathcal{U}[0,1]$. Siguin*

$$Z = (-\log X)\cos(2\pi Y), \quad T = (-\log X)\sin(2\pi Y).$$

Trobeu la llei del vector aleatori (Z, T) i demostreu que els components Z i T són independents.

Capítol 3

Esperança matemàtica

En aquest capítol introduïm el concepte de valor mitjà, valor esperat o esperança matemàtica d'una variable aleatòria i desenvolupem un conjunt d'eines de càlcul amb distribucions de probabilitat. El seu contingut, juntament amb els dels dos capítols precedents, formen el nucli bàsic del *càlcul de probabilitats.*

L'esperança matemàtica d'una variable aleatòria és la integral, en el sentit de Lebesgue, de la variable aleatòria respecte de la probabilitat. El nostre interès serà d'expressar l'esperança en termes d'una integral respecte de la llei o distribució de la variable aleatòria. La justificació matemàtica de tots els ingredients que necessitarem aquí requereix el desenvolupament d'uns quants preliminars de teoria de la mesura. Per evitar una presentació que s'allunyi dels objectius del llibre, hem optat per insistir en els càlculs dels dos exemples fonamentals que han estat estudiats en el capítol 2, les lleis discretes i les absolutament contínues, donant referències al lector interessat a aprofundir en el tema.

3.1 Esperança matemàtica de variables aleatòries simples

Com hem dit abans, volem definir una noció de mitjana dels valors presos per una variable aleatòria. La idea de com fer-ho sembla bastant clara en el cas que aquesta sigui *simple*, és a dir, si té una representació

$$X(\omega) = \sum_{i=1}^{r} a_i \mathbb{1}_{A_i}(\omega), \tag{3.1}$$

on $a_i \in \mathbb{R}$, i $A_i \in \mathcal{A}$, $i = 1, \ldots, r$. En efecte, d'una banda haurem de tenir en compte els valors a_i, $i = 1, \ldots, r$, presos per la variable aleatòria i, d'altra banda, la *intensitat* amb què els pren. Això justifica la definició següent:

Definició 3.1 *L'esperança matemàtica de la variable aleatòria definida en (3.1), que denotarem per $E(X)$, es defineix per*

$$E(X) = \sum_{i=1}^{r} a_i P(A_i). \tag{3.2}$$

És fàcil veure, i ho deixem com a exercici per al lector, que la definició (3.2) és independent de la representació particular que tinguem de la variable aleatòria simple X. D'altra banda, és també evident que si X és la funció indicatriu d'un conjunt $A \in \mathcal{A}$ (vegeus l'exemple 2.1), la seva esperança coincideix amb $P(A)$.

Abans d'entrar en el detall de les propietats principals de l'esperança matemàtica de les variables aleatòries simples, en donem alguns exemples.

Exemple 3.1 Sigui X una variable aleatòria amb llei $B(n, p)$. La seva esperança matemàtica val $E(X) = np$. En efecte, d'acord amb (3.2) i tenint en compte (2.11),

$$\begin{aligned} E(X) &= \sum_{k=1}^{n} \binom{n}{k} k p^k (1-p)^{n-k} \\ &= np \sum_{k=1}^{n} \binom{n-1}{k-1} p^{k-1} (1-p)^{n-k} = np. \end{aligned}$$

Exemple 3.2 Considerem una variable aleatòria X amb distribució hipergeomètrica (vegeu l'apartat 2.4.5 per a la definició i el significat de cadascun dels paràmetres). Es té $E(X) = \frac{Dn}{N}$, segons es desprèn dels càlculs següents:

$$\begin{aligned}
E(X) &= \binom{N}{n}^{-1} \sum_{k=(n-N+D)^+}^{n\wedge D} k\binom{D}{k}\binom{N-D}{n-k} \\
&= \binom{N}{n}^{-1} \sum_{k=(n-N+D)\vee 1}^{n\wedge D} D\binom{D-1}{k-1}\binom{N-D}{n-k} \\
&= \binom{N}{n}^{-1} D \sum_{j=(n-1-N+D)^+}^{(n-1)\wedge D} \binom{D-1}{j}\binom{N-D}{n-1-j} \\
&= \binom{N}{n}^{-1} D\binom{N-1}{n-1} = \frac{Dn}{N}.
\end{aligned}$$

Proposició 3.1 *Siguin X i Y variables aleatòries simples i $a, b \in \mathbb{R}$. Aleshores,*

1. $E(aX + bY) = aE(X) + bE(Y)$,
2. *si* $X, Y \geq 0$ *i* $X \leq Y$, *es té* $E(X) \leq E(Y)$.

Demostració: Sigui X com en (3.1) i $Y = \sum_{j=1}^{s} b_j \mathbb{1}_{B_j}$. La variable aleatòria $aX + bY$ té la representació $\sum_{i=1}^{r}\sum_{j=1}^{s}(aa_i + bb_j)\mathbb{1}_{A_i \cap B_j}$ i és, doncs, simple. Per tant,

$$\begin{aligned}
E(aX + bY) &= \sum_{i=1}^{r}\sum_{j=1}^{s}(aa_i + bb_j)P(A_i \cap B_j) \\
&= \sum_{i=1}^{r} aa_i \sum_{j=1}^{s} P(A_i \cap B_j) + \sum_{j=1}^{s} bb_j \sum_{i=1}^{r} P(A_i \cap B_j) \\
&= aE(X) + bE(Y).
\end{aligned}$$

Això acaba la demostració de l'apartat (1).

Pel que fa a (2), com que $Y - X$ és una variable aleatòria simple i positiva, és obvi, a partir de la definició (3.2), que la seva esperança també ha de ser positiva. $\square$

3.2 Esperança matemàtica de variables aleatòries no negatives

Sigui X una variable aleatòria que pren els seus valors en el conjunt $[0, \infty)$. En la proposició 2.3 es va demostrar l'existència d'una successió creixent de variables aleatòries simples i positives, $\{X_n, n \geq 1\}$, que convergeix puntualment cap a X. Basant-nos en aquest resultat i en l'apartat (b) de la proposició 3.1 donem la definició següent.

Definició 3.2 *L'esperança matemàtica d'una variable aleatòria X no negativa és*

$$E(X) = \lim_{n\to\infty} E(X_n),$$

on $\{X_n, n \geq 1\}$ és una successió de variables aleatòries simples i positives que s'aproximen a X en el sentit que acabem de precisar.

Observem que $E(X) \in [0, \infty]$.

Per a poder considerar acceptable aquesta definició cal, almenys, justificar dues qüestions. En primer lloc, que la definició és independent de la successió que s'aproxima a X i, en segon lloc, que és consistent amb la definició 3.1.

Proposició 3.2 *Siguin $\{X_n, n \geq 1\}$, $\{Y_n, n \geq 1\}$ dues successions de variables aleatòries simples i positives que s'aproximen a la variable aleatòria X en el sentit de la proposició 2.3. Aleshores,*

$$\lim_{n\to\infty} E(X_n) = \lim_{n\to\infty} E(Y_n).$$

Demostració: És suficient demostrar la propietat següent: Sigui Z una variable aleatòria simple tal que $Z \leq X$. Es compleix que

$$E(Z) \leq \lim_{n\to\infty} E(X_n). \tag{3.3}$$

En efecte, aplicant (3.3) a $Z = Y_n$ obtindrem la desigualtat $\lim_{n\to\infty} E(X_n) \geq \lim_{n\to\infty} E(Y_n)$ i, intercanviant els papers de les successions, $\lim_{n\to\infty} E(X_n) \leq \lim_{n\to\infty} E(Y_n)$.

Sigui $\varepsilon > 0$. Definim $A_n = \{X_n > Z - \varepsilon\}$, $n \geq 1$. Els conjunts A_n creixen, quan n tendeix a infinit, cap a Ω i $X_n \geq (Z - \varepsilon)\mathbb{1}_{A_n}$. En conseqüència,

$$\begin{aligned} E(X_n) &\geq E[(Z-\varepsilon)\mathbb{1}_{A_n}] = E(Z\mathbb{1}_{A_n}) - \varepsilon P(A_n) \\ &\geq E(Z) - E(Z\mathbb{1}_{A_n^c}) - \varepsilon \\ &\geq E(Z) - P(A_n^c)\max_{\omega} Z - \varepsilon. \end{aligned}$$

Com que $\lim_{n\to\infty} P(A_n^c) = 0$, el darrer terme d'aquesta desigualtat tendeix a $E(Z) - \varepsilon$ quan $n \to \infty$ i, com que ε és arbitrari, això acaba la demostració de (3.3). □

Si X és una variable aleatòria simple, podem considerar com a successió aproximadora, la definida per $X_n = X$, per a tot n. Per tant, a la vista de la proposició 3.2, les definicions 3.1 i 3.2 són coherents.

Els resultats demostrats en la proposició 3.1 de linealitat i monotonia de l'esperança, s'estenen sense cap dificultat a l'esperança de les variables aleatòries no negatives.

3.3 Variables aleatòries amb esperança matemàtica finita

Sigui X una variable aleatòria arbitrària. Considerem la descomposició $X = X^+ - X^-$ en termes de les seves parts positiva i negativa respectivament (vegeu la secció 2.1). Com que X^+, X^- són variables aleatòries positives, podem definir $E(X^+)$, $E(X^-)$ a partir de la definició 3.2. Volem definir l'esperança matemàtica de X de manera que es compleixi la propietat de linealitat que hem expressat en la proposició 3.1. Sembla, doncs, natural proposar $E(X) = E(X^+) - E(X^-)$. Ara bé, aquesta proposta pot no tenir sentit si ambdues esperances, $E(X^+)$ i $E(X^-)$, prenen el valor infinit. Aquestes consideracions porten a la definició següent.

Definició 3.3 *Donada una variable aleatòria X qualsevol, direm que té esperança matemàtica* finita *si, i només si, $E(|X|) < \infty$ i, en aquest cas, definim*

$$E(X) = E(X^+) - E(X^-).$$

Per simplificar el llenguatge, a partir d'ara denotarem per $L^1(\Omega)$ el conjunt de les variables aleatòries definides en un espai de probabilitat $(\Omega, \mathcal{A}, P)$ que tenen esperança matemàtica finita.

De la descomposició $|X| = X^+ + X^-$ es dedueix que $X \in L^1(\Omega)$ si, i només si, $E(X^+), E(X^-)$ són ambdues finites.

En les dues proposicions que segueixen s'enuncien algunes de les propietats més importants de l'esperança matemàtica.

Proposició 3.3 *Siguin $X, Y \in L^1(\Omega)$, $a, b \in \mathbb{R}$.*

1. *La variable aleatòria $aX + bY \in L^1(\Omega)$ i*

$$E(aX + bY) = aE(X) + bE(Y).$$

2. *Propietat de convexitat: Es compleix*

$$|E(X)| \leq E(|X|).$$

Demostració: Comencem provant que $aX \in L^1(\Omega)$ i $E(aX) = aE(X)$. Si $a > 0$, el resultat és obvi. Si $a < 0$, hem de tenir en compte la descomposició següent:

$$aX = (aX)^+ - (aX)^- = (-a)X^- - (-a)X^+.$$

Ara, per acabar la demostració de l'apartat (1), podem suposar $a = b = 1$. Per la desigualtat triangular, $X + Y \in L^1(\Omega)$. Aleshores, la fórmula $E(X + Y) = E(X) + E(Y)$ és una conseqüència de la identitat

$$X + Y = (X^+ + Y^+) - (X^- + Y^-).$$

Per a la demostració de (2) utilitzem la definició d'esperança, la desigualtat triangular i la linealitat de l'esperança de variables aleatòries positives:

$$|E(X)| = |E(X^+) - E(X^-)| \leq E(X^+) + E(X^-) = E(|X|).$$

□

Proposició 3.4 (Resultats de convergència)

1. *Convergència monòtona. Sigui* $\{X_n, n \geq 1\}$ *una successió creixent de variables aleatòries positives que convergeix puntualment cap a una variable aleatòria* X*. Aleshores,*

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

2. *Convergència dominada. Sigui* $\{X_n, n \geq 1\}$ *una successió de variables aleatòries convergent puntualment cap a una variable aleatòria* X*. Suposem que* $|X_n| \leq Y$*, on* $Y \in L^1(\Omega)$*. Aleshores,*

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

La demostració pot trobar-se, per exemple, en [C] o [N-S].

3.4 Càlcul d'esperances matemàtiques

En les seccions anteriors hem donat les definicions precises d'esperança matemàtica per a diferents tipus de variables aleatòries. En el cas de les variables aleatòries simples hem pogut utilitzar la definició per al càlcul efectiu en algun exemple. Això resulta que és molt més difícil per a les variables aleatòries positives o les generals. L'objectiu d'aquesta secció és de donar resultats precisos per al càlcul d'esperances en els dos tipus de variables aleatòries que hem introduït en el capítol 2, les discretes i les absolutament contínues.

3.4.1 Variables aleatòries discretes

Sigui X una variable aleatòria discreta *no negativa* amb una representació com en (2.10), és a dir,

$$X = \sum_{i\in\mathbb{N}} a_i \mathbb{1}_{A_i}, \tag{3.4}$$

on els conjunts A_i, $i \in \mathbb{N}$, són de $\mathcal{A}$ i formen una partició de Ω. Es compleix que

$$E(X) = \sum_{i\in\mathbb{N}} a_i P(A_i). \tag{3.5}$$

Aquest resultat és una conseqüència immediata de la definició 3.1 i del teorema de la convergència monòtona.

Exemple 3.3 Per a una variable aleatòria amb distribució de Poisson amb paràmetre $\lambda > 0$ (vegeu l'apartat 2.4.6) es té $E(X) = \lambda$. En efecte, segons (3.5),

$$\begin{aligned} E(X) &= \sum_{k=0}^{\infty} k \exp(-\lambda) \frac{\lambda^k}{k!} \\ &= \exp(-\lambda)\lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \\ &= \lambda. \end{aligned}$$

Exemple 3.4 Considerem una variable aleatòria amb distribució geomètrica de paràmetre $p \in (0,1)$ (vegeu l'apartat 2.4.3). La seva esperança matemàtica val $\frac{1}{p}$. En efecte, d'acord amb (3.5),

$$E(X) = \sum_{k=1}^{\infty} kp(1-p)^{k-1} = \frac{1}{p}.$$

En efecte, la sèrie $\sum_{k=1}^{\infty} k(1-p)^{k-1}$ suma $\frac{1}{p^2}$. Aquest resultat s'obté per derivació de la funció $\phi(p) = \sum_{k=0}^{\infty}(1-p)^k$. Observeu que el radi de convergència d'aquesta sèrie és menor que 1.

Suposem ara que X és una variable aleatòria discreta qualsevol, és a dir, no necessàriament *no negativa*. Aleshores $X \in L^1(\Omega)$ si, i només si,

$$\sum_{i \in \mathbb{N}} |a_i| P(A_i) < \infty, \tag{3.6}$$

i, en aquest cas,

$$E(X) = \sum_{i \in \mathbb{N}} a_i P(A_i). \tag{3.7}$$

Aquest resultat s'obté particularitzant la definició 3.3 a les variables aleatòries que tenen una representació (3.4). En efecte, es té que

$$X^+ = \sum_{i\in\mathbb{N}} a_i^+ \mathbb{1}_{A_i},$$
$$X^- = \sum_{i\in\mathbb{N}} a_i^- \mathbb{1}_{A_i}.$$

Exemple 3.5 Sigui X una variable aleatòria que pren els valors $\frac{2^n}{n}(-1)^{n+1}$ amb probabilitat $\frac{1}{2^n}$, $n \geq 1$. X no té esperança finita. En efecte, la sèrie $\sum_{i\in\mathbb{N}} |a_i|P(A_i)$ suma infinit, ja que

$$\sum_{n=1}^{\infty} |(-1)^{n+1}| \frac{2^n}{n} \frac{1}{2^n} = \sum_{n=1}^{\infty} \frac{1}{n}.$$

Tanmateix, $\sum_{i\in\mathbb{N}} a_i P(A_i) < \infty$, car es tracta d'una sèrie harmònica alternada.

3.4.2 Variables aleatòries absolutament contínues

Una variable aleatòria X amb densitat f pertany a $L^1(\Omega)$ si, i només si,

$$\int_{-\infty}^{\infty} |x| f(x)\,\mathrm{d}x < \infty \tag{3.8}$$

i, en aquest cas,

$$E(X) = \int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x. \tag{3.9}$$

Les integrals en (3.8) i (3.9) són en el sentit de Lebesgue.

No donem aquí la justificació de l'equivalència entre la definició 3.3 i les condicions (3.8), (3.9) ja que per a això calen certs ingredients de la teoria de la mesura que s'escapen de l'abast d'aquest llibre (teorema de la mesura imatge i teorema de Radon-Nikodym).

Al llarg d'aquest text entendrem que el càlcul d'esperances de variables aleatòries amb densitat el fem mitjançant *integrals de Riemann*. En aquest sentit ens pot ser útil el resultat següent:

Sigui g una funció fitada en $[a, b]$.

1. Si g és contínua, llevat potser en un nombre finit de punts, és integrable en el sentit de Riemann.

2. Si g és integrable en el sentit de Riemann, també ho és en el sentit de Lebesgue i ambdues integrals coincideixen.

Exemple 3.6 Considerem una variable aleatòria amb distribució $\mathcal{U}[a,b]$ (vegeu l'apartat 2.6.1). És evident que en aquest cas es compleix la condició (3.8). A més,

$$E(X) = \frac{1}{b-a}\int_a^b x\,\mathrm{d}x = \frac{a+b}{2}.$$

Exemple 3.7 Sigui X una distribució $\mathcal{N}(0,1)$. Demostrem tot seguit que X té esperança finita i

$$E(X) = 0.$$

Per comprovar la condició (3.8) observem que la funció $\psi(x) = |x|\mathrm{e}^{-\frac{x^2}{2}}$ és parella. Aleshores,

$$\int_{-\infty}^{\infty} |x|\mathrm{e}^{-\frac{x^2}{2}}\,\mathrm{d}x = 2\int_0^{\infty} x\mathrm{e}^{-\frac{x^2}{2}}\,\mathrm{d}x = 2.$$

Una vegada comprovada la condició d'integrabilitat, és clar que

$$\int_{-\infty}^{\infty} x\mathrm{e}^{-\frac{x^2}{2}}\,\mathrm{d}x = 0,$$

ja que la funció $\varphi(x) = x\mathrm{e}^{-\frac{x^2}{2}}$ és imparella.

Exemple 3.8 Considerem una variable aleatòria X amb distribució Gamma(p,λ) (vegeu l'apartat 2.6.4). En aquest cas, analitzar l'existència de $E(X)$ i calcular el seu valor comporta el càlcul de la mateixa integral:

$$J = \int_0^{\infty} x\frac{\lambda^p}{\Gamma(p)}x^{p-1}\mathrm{e}^{-\lambda x}\,\mathrm{d}x.$$

Es té, fent servir (2.21),

$$J = \frac{1}{\Gamma(p)} \int_0^\infty \lambda^p x^p \mathrm{e}^{-\lambda x}\, \mathrm{d}x = \frac{\Gamma(p+1)}{\lambda \Gamma(p)}.$$

La distribució *exponencial* amb paràmetre $\lambda > 0$ és una Gamma$(1, \lambda)$. Com que per a tot $p \geq 1$ es compleix que $\Gamma(p+1) = p\Gamma(p)$, dels càlculs anteriors es dedueix que la mitjana d'una variable aleatòria amb llei $\exp(\lambda)$ és $\frac{1}{\lambda}$.

Exemple 3.9 Considerem una variable aleatòria amb una funció de distribució com la donada en la figura 3.1. És clar que X no és ni discreta ni absolutament contínua. Ara bé, a causa del resultat de descomposició de la proposició 2.6 i a la forma particular de la part contínua en aquest exemple, es pot calcular l'esperança de X utilitzant els resultats d'aquest apartat i el precedent.

Més concretament, suposem que en la descomposició de la proposició 2.6,

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

on f és una funció positiva, integrable en el sentit de Riemann en $\mathbb{R}$. Per exemple, sigui

$$F(x) = \begin{cases} 0, & \text{si } x < 0, \\ \frac{1}{2}(x+1), & \text{si } x \in [0,1), \\ 1, & \text{si } x \geq 1. \end{cases}$$

1

0.5

1

Figura 3.1: Gràfica de la funció de distribució F

En aquest cas, $F_c(x) = \int_{-\infty}^{x} \frac{1}{2}\mathbb{1}_{[0,1]}\,\mathrm{d}y$, $F_d(x) = \frac{1}{2}\mathbb{1}_{[0,\infty)}$. Aleshores, si X és una variable aleatòria amb funció de distribució F, X té esperança finita i

$$E(X) = 0 \cdot \frac{1}{2} + \int_0^1 \frac{y}{2}\,\mathrm{d}y = \frac{1}{4}.$$

La justificació heurística d'aquest càlcul és la següent. L'esperança matemàtica de X és la integral de Lebesgue de X respecte de la probabilitat P. Mitjançant el teorema de la mesura imatge, que juga el paper anàleg al del canvi de variable per a la integral de Riemann, l'esperança pot calcular-se com una integral respecte de la mesura sobre $\mathbb{R}$ definida mitjançant la funció de distribució. Aquesta és una integral de Lebesgue-Stieltjes. En el cas discret, això ens porta a l'expressió (3.7) i, en el cas absolutament continu, a la (3.9). En un cas mixt, com en l'exemple anterior, ja que la integral és additiva respecte a la mesura, la integral es calcularà fent-ne la suma respecte de les parts discontínua i contínua de la mesura.

3.5 Esperança matemàtica de funcions de variables aleatòries

Sigui X una variable aleatòria i $g : \mathbb{R} \longrightarrow \mathbb{R}$ tal que $g(X)$ és també una variable aleatòria. Per exemple, això és així si g és mesurable, és a dir, si l'antiimatge d'un conjunt de la σ-àlgebra de Borel és també un conjunt de la mateixa σ-àlgebra. En aquest apartat estem interessats en l'existència i el càlcul de l'esperança matemàtica de $g(X)$.

Si coneixem la llei de $g(X)$, podríem, en principi, utilitzar els resultats dels apartats precedents. Recordem que, si $g(x)$ és prou regular, la densitat de $g(X)$ es pot trobar explícitament. Podem, però, fer càlculs més directes.

Proposició 3.5 *Sigui X una variable aleatòria discreta amb una representació com en (2.10), $g : \mathbb{R} \to \mathbb{R}$ una funció tal que $Y := g(X)$ és una variable aleatòria. Aleshores, Y té esperança finita si, i només si,*

$$\sum_i |g(a_i)| P\{X = a_i\} < \infty \tag{3.10}$$

i, en aquest cas,

$$E(Y) = \sum_i g(a_i) P\{X = a_i\}. \tag{3.11}$$

Demostració: Designem per $\{y_1, \dots, y_n, \dots\}$ els valors que pren la variable aleatòria Y. Es té la identitat següent:

$$\{Y = y_j\} = \cup_{g(a_i)=y_j} \{X = a_i\},$$

que és una reunió disjunta. Per tant,

$$\begin{aligned} \sum_j |y_j| P\{Y = y_j\} &= \sum_j |y_j| \sum_{i:g(a_i)=y_j} P\{X = a_i\} \\ &= \sum_j \sum_{i:g(a_i)=y_j} |g(a_i)| P\{X = a_i\} = \sum_i |g(a_i)| P\{X = a_i\}. \end{aligned}$$

Això dóna la condició necessària i suficient per a l'existència de $E(Y)$. L'expressió de $E(Y)$ prové dels càlculs que acabem de fer, però sense escriure els valors absoluts. □

A continuació donem la versió absolutament contínua de la proposició anterior.

Proposició 3.6 *Sigui X una variable aleatòria amb llei absolutament contínua. Denotem per f la seva densitat. Sigui g com en la proposició 3.5. Definim $Y = g(X)$. Aleshores, Y té esperança finita si, i només si,*

$$\int_{-\infty}^{\infty} |g(x)| f(x)\, \mathrm{d}x < \infty, \tag{3.12}$$

i, en aquest cas,

$$E(Y) = \int_{-\infty}^{\infty} g(x) f(x)\, \mathrm{d}x. \tag{3.13}$$

Demostració: Farem la prova en un cas particular. Considerem una situació de validesa de la fórmula (2.22). Més precisament, suposarem que X pren els seus valors en un interval obert I i g és una transformació bijectiva, contínuament diferenciable de I en un altre interval obert J. Arguments anàlegs s'apliquen a la situació menys restrictiva sota la qual val (2.24).

D'acord amb (3.8), la variable aleatòria Y té esperança finita si, i només si,

$$K = \int_J |y| f\bigl(g^{-1}(y)\bigr) \frac{1}{|g'\bigl(g^{-1}(y)\bigr)|} \,\mathrm{d}y < \infty.$$

Fent el canvi de variable $y = g(x)$, resulta que

$$K = \int_I |g(x)| f(x) \,\mathrm{d}x,$$

que és la condició (3.12) per a aquest cas particular. Endemés, per les mateixes raons que abans,

$$E(Y) = \int_I g(x) f(x) \,\mathrm{d}x.$$

Això acaba la demostració en el cas particular que hem considerat. □

3.6 Moments de variables aleatòries

En aquest apartat introduïm paràmetres relacionats amb la llei d'una variable aleatòria que proporcionen informació sobre aquesta. Un cas particular és l'esperança matemàtica, que ha estat estudiada a bastament en les seccions precedents.

3.6.1 Moments d'ordre k

Definició 3.4 *Donada una variable aleatòria X i $k \in \mathbb{N}$, direm que X té* moment d'ordre k *finit si X^k té esperança finita.*

Denotarem per $\mu_k = E(X^k)$ el moment d'ordre k. El moment d'ordre $k = 1$ és l'esperança matemàtica.

Si X té esperança finita, direm que té *moment centrat d'ordre k* finit, si $X - E(X)$ té moment d'ordre k finit. Denotarem per $\tilde{\mu}_k = E(X - E(X))^k$ el moment centrat d'ordre k.

Des del punt de vista del coneixement de la distribució de X, un element d'interès és el *moment centrat d'ordre* 2, anomenat també *variància*, i que designarem per $\operatorname{Var} X$ o $\sigma^2(X)$. Representa la mitjana de la desviació de X en un entorn del seu valor mitjà. En termes físics, representa el moment d'inèrcia d'una distribució de masses. L'arrel quadrada de la variància de X s'anomena *desviació típica* de X i es denota per $\sigma(X)$.

És important saber quina relació hi ha entre els moments de diversos ordres. Part de la proposició següent dóna un resultat en aquesta direcció.

Proposició 3.7

1. *Sigui X una variable aleatòria amb moment d'ordre k finit. Aleshores, X té també moment d'ordre r finit, per a qualsevol* $1 \leq r \leq k$.
2. *Si X i Y són dues variables aleatòries amb moment d'ordre k finit, la variable aleatòria $X+Y$ té també moment d'ordre k finit.*

Demostració: L'apartat (1) de la proposició és una conseqüència de la desigualtat de Hölder; l'apartat (2) s'obté per la desigualtat de Minkowski. No donem aquí la demostració d'aquestes desigualtats. El lector interessat pot consultar, per exemple, [C].

A tall d'il·lustració provem la proposició en el cas particular de variables aleatòries discretes.

Fixem $1 \leq r \leq k$. Per a tot $x \in \mathbb{R}$ es compleix que

$$|x|^r \leq 1 + |x|^k.$$

En efecte, si $|x| < 1$, aleshores $|x|^r \leq 1$. Mentre que si $|x| \geq 1$, aleshores $|x|^r \leq |x|^k$.

En conseqüència, suposant que $\{a_1, a_2, \ldots\}$ és el conjunt de valors presos per la variable X i denotant per $p(a_i)$ les probabilitats $P\{X = a_i\}$, tenim que

$$\sum_i |a_i|^r\, p(a_i) \leq \sum_i (1 + |a_i|^k)\, p(a_i) \leq 1 + \sum_i |a_i|^k\, p(a_i).$$

Per hipòtesi, el darrer membre d'aquesta expressió és finit. Això demostra la primera part de l'enunciat.

La demostració de la segona part utilitza la desigualtat següent:

Per a tot nombre natural k i per a tots $x, y \in \mathbb{R}$ es compleix

$$|x+y|^k \leq 2^{k-1}(|x|^k + |y|^k).$$

Denotem per $\{a_i, i \geq 1\}$, $\{b_j, j \geq 1\}$, els valors presos per les variables aleatòries X i Y, respectivament i per $p(a_i, b_j)$ les probabilitats $P\{X = a_i, Y = b_j\}$. Aleshores, tenim les desigualtats següents:

$$\begin{aligned}
\sum_{i,j} |a_i + b_j|^k p(a_i, b_j) &\leq 2^{k-1} \sum_{i,j} (|a_i|^k + |b_j|^k) p(a_i, b_j) \\
&= 2^{k-1} \sum_i |a_i|^k \sum_j p(a_i, b_j) + 2^{k-1} \sum_j |b_j|^k \sum_i p(a_i, b_j) \\
&= 2^{k-1} \sum_i |a_i|^k p(a_i) + 2^{k-1} \sum_j |b_j|^k p(b_j).
\end{aligned}$$

Això demostra la segona part de la proposició. □

3.6.2 Desigualtats

Dediquem aquest apartat a presentar dues desigualtats utilitzades molt sovint.

Proposició 3.8 (Desigualtat de Txebixev) *Sigui X una variable aleatòria no negativa i $f : \mathbb{R}_+ \longrightarrow \mathbb{R}_+$ una funció creixent tal que $E\big(f(X)\big) < \infty$. Per a tot nombre real a es compleix que*

$$f(a) P\{X \geq a\} \leq E\big(f(X)\big). \tag{3.14}$$

Demostració: Considerem la desigualtat

$$f(a) \mathbb{1}_{\{X \geq a\}} \leq f(X).$$

Prenent esperances s'obté que

$$f(a) P\{X \geq a\} \leq E\big(f(X)\big),$$

que és la desigualtat buscada. $\square$

En particular, si X té moment d'ordre k finit i $a > 0$, tindrem que

$$P\{|X| \geq a\} \leq \frac{E(|X|^k)}{a^k}. \tag{3.15}$$

En efecte, (3.15) s'obté aplicant la desigualtat (3.14) a la funció $f(x) = x^k$ i a la variable $|X|$.

Sigui X una variable aleatòria amb moment de segon ordre finit. La desigualtat (3.15) aplicada a la variable aleatòria $|X - E(X)|$ i a $k = 2$ ens dóna

$$P\{|X - E(X)| \geq a\} \leq \frac{\sigma^2(X)}{a^2}. \tag{3.16}$$

En particular, per a un valor fix $\alpha \in (0,1)$, podrem assegurar que, amb probabilitat d'error inferior a α, la variable X pertany a l'interval $\left[E(X) - \frac{\sigma(X)}{\sqrt{\alpha}}, E(X) + \frac{\sigma(X)}{\sqrt{\alpha}}\right]$. Observeu que la longitud d'aquest interval és proporcional a la desviació típica $\sigma(X)$.

Proposició 3.9 (Desigualtat de Jensen) *Sigui g una funció real* convexa *i X una variable aleatòria de $L^1(\Omega)$ tal que $g(X) \in L^1(\Omega)$. Aleshores,*

$$g(E(X)) \leq E(g(X)). \tag{3.17}$$

Demostració: Com que g és convexa, existeix $a \in \mathbb{R}$ tal que per a tot $x \in \mathbb{R}$

$$g(x) \geq g(E(X)) + a(x - E(X)).$$

Substituint x per $X(\omega)$ tenim, per a tot $\omega \in \Omega$,

$$g(X) \geq g(E(X)) + a(X - E(X)).$$

La propietat de monotonia de l'esperança matemàtica permet de concloure la demostració. $\square$

La funció $g(x) = |x|$ és convexa; per tant, la desigualtat de Jensen generalitza la propietat establerta en l'apartat (2) de la proposició 3.3.

3.7 Independència de variables aleatòries i moments

En l'apartat 3.5 vàrem donar fórmules per al càlcul d'esperances de funcions de variables aleatòries en el cas discret i absolutament continu, respectivament. Les fórmules (3.11) i (3.13) admeten una extensió a vectors aleatoris i proporcionen una eina per a l'estudi de propietats dels moments relacionades amb la independència.

En la proposició següent recollim les extensions a vectors aleatoris de les proposicions 3.5 i 3.6. Aquestes són les notacions que utilitzem: X denota un vector aleatori n-dimensional, $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ és una funció mesurable. Recordem que aquesta propietat implica que $g(X)$ és una variable aleatòria.

Proposició 3.10

1. *Si X és un vector aleatori discret que pren els seus valors en un conjunt numerable $\mathcal{B}_0$, la variable aleatòria $g(X)$ té esperança finita si, i només si,*

$$\sum_{(x_1,\ldots,x_n)\in\mathcal{B}_0} |g(x_1,\ldots,x_n)| P\{X_1 = x_1,\ldots,X_n = x_n\} < \infty \tag{3.18}$$

   i, en aquest cas,

$$E(g(X)) = \sum_{(x_1,\ldots,x_n)\in\mathcal{B}_0} g(x_1,\ldots,x_n) P\{X_1 = x_1,\ldots,X_n = x_n\}. \tag{3.19}$$

2. *Si X és un vector aleatori absolutament continu, amb densitat f, la variable aleatòria $g(X)$ té esperança finita si, i només si,*

$$\int_{\mathbb{R}^n} |g(x_1,\ldots,x_n)| f(x_1,\ldots,x_n)\, \mathrm{d}x_1 \cdots \mathrm{d}x_n < \infty \tag{3.20}$$

   i, en aquest cas,

$$E(g(X)) = \int_{\mathbb{R}^n} g(x_1,\ldots,x_n) f(x_1,\ldots,x_n)\, \mathrm{d}x_1 \cdots \mathrm{d}x_n. \tag{3.21}$$

La proposició següent és una conseqüència de (2.35) i (2.39).

Proposició 3.11 *Siguin X i Y dues variables aleatòries independents amb esperança finita. Aleshores, la variable aleatòria producte XY té també esperança finita i*

$$E(XY) = E(X)E(Y).$$

Demostració: Considerem primer el cas discret. Siguin $\{x_1, x_2, \ldots\}$ i $\{y_1, y_2, \ldots\}$ els conjunts on prenen els seus valors les variables aleatòries X i Y, respectivament. Per provar que XY té esperança finita, cal comprovar que la sèrie

$$\sum_{i,j} |x_i y_j| P\{X = x_i, Y = y_j\}$$

convergeix. Ara bé, per la hipòtesi d'independència (vegeu (2.35)), es té que

$$\begin{aligned}\sum_{i,j} |x_i y_j| P\{X = x_i, Y = y_j\} &= \sum_i |x_i| P\{X = x_i\} \sum_j |y_j| P\{Y = y_j\} \\ &= E(|X|)E(|Y|) < \infty.\end{aligned}$$

D'altra banda, els càlculs fets anteriorment, sense prendre els valors absoluts, demostren la igualtat $E(XY) = E(X)E(Y)$.
Suposem ara que X i Y tenen llei absolutament contínua amb densitats f_X, f_Y, respectivament. Per la proposició 2.13, el vector (X, Y) té també llei absolutament contínua i la seva densitat és

$$f(x, y) = f_X(x) f_Y(y).$$

En conseqüència,

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |xy| f(x, y)\,\mathrm{d}x\,\mathrm{d}y = \left(\int_{-\infty}^{\infty} |x| f_X(x)\,\mathrm{d}x \right) \left(\int_{-\infty}^{\infty} |y| f_Y(y)\,\mathrm{d}y \right) < \infty,$$

i, segons (3.20), tenim l'existència d'esperança per a XY. D'altra banda,

$$\begin{aligned}
&\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)\,\mathrm{d}x\,\mathrm{d}y \\
&\quad = \left(\int_{-\infty}^{\infty} xf_X(x)\,\mathrm{d}x\right)\left(\int_{-\infty}^{\infty} yf_Y(y)\,\mathrm{d}y\right) \\
&\quad = E(X)E(Y).
\end{aligned}$$

□

Heus aquí una conseqüència important de la proposició anterior.

Proposició 3.12 *Siguin X i Y dues variables aleatòries independents amb moments de segon ordre finits. Aleshores,*

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y).$$

Demostració: Per la definició de la variància,

$$\begin{aligned}
\mathrm{Var}(X+Y) &= E\left[\big((X+Y) - E(X+Y)\big)^2\right] \\
&= E\left[\big((X-E(X)) + (Y-E(Y))\big)^2\right] \\
&= E\left[(X-E(X))^2\right] + E\big[(Y-E(Y))^2\big] + 2E\left[(X-E(X))(Y-E(Y))\right].
\end{aligned}$$

La propietat d'independència i la linealitat de l'esperança implica

$$E\left[\big(X-E(X)\big)\big(Y-E(Y)\big)\right] = 0,$$

per tant, el resultat queda demostrat.

□

La proposició anterior s'estén sense dificultat a un nombre finit de variables aleatòries independents.

Per mesurar el grau de dependència entre dues variables aleatòries X i Y s'introdueix la noció de *covariància*. Suposem que les variables tenen moments de segon ordre finits. La *covariància* entre X i Y es defineix per

$$\mathrm{Cov}(X,Y) = E\left[\big(X-E(X)\big)\big(Y-E(Y)\big)\right].$$

Observeu que si $X = Y$, aleshores $\operatorname{Cov}(X, Y) = \operatorname{Var}(X)$.

Si les variables X i Y són independents, aleshores $\operatorname{Cov}(X, Y) = 0$. Les variables aleatòries que tenen covariància nul·la s'anomenen *incorrelacionades*. Ara bé, dues variables aleatòries poden tenir covariància nul·la i no ser independents. Per exemple, considerem un vector aleatori (X, Y) amb distribució uniforme en el conjunt $\{(-1, 0), (0, 1), (0, -1), (1, 0)\}$. La llei marginal de X es concentra en $\{-1, 0, 1\}$ amb $P\{X = -1\} = \frac{1}{4}$, $P\{X = 0\} = \frac{1}{2}$, $P\{X = 1\} = \frac{1}{4}$. Anàlogament, Y es concentra també en $\{-1, 0, 1\}$ i $P\{Y = -1\} = \frac{1}{4}$, $P\{Y = 0\} = \frac{1}{2}$, $P\{Y = 1\} = \frac{1}{4}$. Aleshores, $E(X) = E(Y) = 0$ i $\operatorname{Cov}(X, Y) = 0$. Però X i Y no són independents, car $P\{(X, Y) = (-1, 0)\} = \frac{1}{4}$ i $P\{X = -1\}P\{Y = 0\} = \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{8}$.

A la covariància, convenientment normalitzada, se l'anomena *coeficient de correlació*. Més precisament, es defineix el *coeficient de correlació* entre dues variables aleatòries X i Y amb moments de segon ordre finits per

$$\rho_{X,Y} = \frac{\operatorname{Cov}(X, Y)}{\sqrt{\operatorname{Var}(X) \operatorname{Var}(Y)}} = \frac{\operatorname{Cov}(X, Y)}{\sigma_X \sigma_Y}.$$

Ja hem vist que si les variables X i Y són independents, aleshores $\rho_{X,Y} = 0$. La proposició següent ens precisa els valors que pot prendre el coeficient de correlació: $\rho_{X,Y} \in [-1, 1]$.

Proposició 3.13 *Siguin X i Y variables aleatòries amb moments de segon ordre finits. Aleshores es compleix que*

$$\big(\operatorname{Cov}(X, Y)\big)^2 \leq \operatorname{Var}(X) \operatorname{Var}(Y).$$

Demostració: Sigui $\lambda \in \mathbb{R}$. Es satisfà que

$$0 \leq \operatorname{Var}(\lambda X - Y) = \lambda^2 \operatorname{Var}(X) - 2\lambda \operatorname{Cov}(X, Y) + \operatorname{Var}(Y).$$

Per tant, la funció de λ (trinomi de segon grau) donada per

$$\lambda^2 \operatorname{Var}(X) - 2\lambda \operatorname{Cov}(X, Y) + \operatorname{Var}(Y)$$

no pot prendre valors negatius. Això vol dir que el seu discriminant és menor o igual a zero. És a dir,

$$\big(\operatorname{Cov}(X,Y)\big)^2 - \operatorname{Var}(X)\operatorname{Var}(Y) \le 0,$$

que és el que es volia provar. □

Exemple 3.10 Una variable aleatòria X binomial $B(n,p)$ (vegeu 2.4.2), pot obtenir-se fent la suma de n variables aleatòries *independents* de Bernoulli $B(1,p)$ Resulta, doncs, immediat que $E(X) = np$. D'altra banda, la variància d'una variable aleatòria $B(1,p)$ és pq. En conseqüència, $\operatorname{Var}(X) = npq$.

Exemple 3.11 Sigui (X,Y) un vector aleatori absolutament continu amb densitat constant en el quadrat de vèrtexs $(-1,0),(0,-1),(1,0),(0,1)$. Definim

$$Z = \begin{cases} X^2Y, & \text{si } X \le 0,\\ XY^2, & \text{si } X > 0. \end{cases}$$

Trobem el valor mitjà de Z.

En primer lloc, escriurem la densitat f_{XY} del vector aleatori (X,Y). Aquesta és donada per (2.30), on R indica el quadrat de l'enunciat. Tenim que

$$f_{XY}(x,y) = \frac{1}{2}\mathbb{1}_{\{(x,y)\in\mathbb{R}^2:-1\le x\le 0,-1-x\le y\le 1+x\}\cup\{(x,y)\in\mathbb{R}^2:0\le x\le 1,x-1\le y\le 1-x\}}.$$

Per al càlcul de l'esperança matemàtica de Z farem servir la proposició 3.10 per a la funció $g:\mathbb{R}^2 \longrightarrow \mathbb{R}$ donada per

$$g(x,y) = \begin{cases} x^2y, & \text{si } x \le 0,\\ xy^2, & \text{si } x > 0. \end{cases}$$

La condició (3.20) es compleix clarament. Per tant, segons (3.21),

$$\begin{aligned} E(Z) &= \frac{1}{2}\left(\int_{-1}^{0} \mathrm{d}x \int_{-1-x}^{1+x} \mathrm{d}y (x^2y) + \int_0^1 \mathrm{d}x \int_{x-1}^{1-x} \mathrm{d}y (xy^2)\right)\\ &= \frac{1}{2}\int_0^1 \mathrm{d}x \left(\frac{(1-x)^3}{3} - \frac{(x-1)^3}{3}\right) x = \frac{1}{60}. \end{aligned}$$

mn$(0, 1)$

mn$y - x = 1$ mn$x + y = 1$

mn$(-1, 0)$ mn$(1, 0)$

mn$x + y = -1$ mn$x - y = 1$

mn$(0, -1)$

Figura 3.2: Exemple 3.1

3.8 Regressió lineal

En aquest apartat considerarem dues variables aleatòries amb *moments de segon ordre finits* i variàncies estrictament positives. El problema que plantegem és el d'aproximar la variable aleatòria Y per una funció afí de X, $aX+b$. Hem de concretar el sentit d'aquesta aproximació: imposarem que

$$E\big((Y - aX - b)^2\big) \tag{3.22}$$

sigui mínim. És a dir, hem de trobar dos nombres reals a i b tals que la funció

$$\varphi(a, b) = E\big((Y - aX - b)^2\big),$$

sigui mínima.

Minimitzar (3.22) significa garantir un valor petit per al valor mitjà (en sentit aleatori) dels errors comesos en *linearitzar* Y. Els errors es mesuren mitjançant la funció $\psi(y) = |y|^2$.

Definim $\bar{X} = X - E(X)$ i $\bar{Y} = Y - E(Y)$. Aleshores,

$$\varphi(a, b) = E\left[(\bar{Y} - a\bar{X})^2\right] + E[\big(b - E(Y) + aE(X)\big)^2\Big] .$$

Per tant, haurem de prendre $b = E(Y) - aE(X)$. Un càlcul senzill prova que

$$E\left[(\bar{Y}-a\bar{X})^2\right]=\sigma^2(Y)+a^2\sigma^2(X)-2a\rho_{X,Y}\sigma(X)\sigma(Y).$$

Es demostra fàcilment que aquesta funció de a assoleix el seu valor mínim per a $a=\rho_{X,Y}\frac{\sigma(Y)}{\sigma(X)}$.

Resumint, el valor mínim de la funció $\varphi(a,b)$ es pren en el punt (a, b) donat per

$$\begin{aligned} a &= \rho_{X,Y}\frac{\sigma(Y)}{\sigma(X)} \\ b &= E(Y)-\rho_{X,Y}\frac{\sigma(Y)}{\sigma(X)}E(X). \end{aligned}$$

A més, el valor pres per la funció en aquest punt és

$$E\big((Y-aX-b)\big)^2=\sigma^2(Y)(1-\rho^2_{X,Y}).$$

Aquesta darrera quantitat, $\varepsilon := \sigma^2(Y)(1-\rho^2_{X,Y})$, mesura l'error comès en l'aproximació. Observeu que $\varepsilon = 0$ si, i només si, $|\rho_{X,Y}| = 1$.

La recta $y = ax + b$ s'anomena *recta de regressió de la variable Y sobre la variable X*. El mètode que hem utilitzat per determinar-la és el de *mínims quadrats.*

Els papers de les variables aleatòries X i Y es poden intercanviar i així determinar la *recta de regressió de la variable X sobre la variable Y*.

3.9 Funcions generatrius

En aquest apartat introduïm una tècnica molt útil en el càlcul de probabilitats i, en particular, en el tractament de les variables aleatòries independents. La idea és fer una transformació de la llei d'una variable aleatòria de manera que l'objecte obtingut conservi tota la informació que teníem al començament i que sigui un procés reversible. Amb la transformació volem obtenir més operativitat en el tractament de certs problemes. Aquesta tècnica és bastant usual en l'anàlisi matemàtica.

En tot aquest apartat, X designarà una variable aleatòria amb valors en els enters no negatius, i denotarem per $p(n)$ les probabilitats $P\{X=n\}$. Observeu que

$$\sum_{n\in\mathbb{Z}^+} p(n) = 1.$$

3.9.1 Generalitats

Definició 3.5 *S'anomena* funció generatriu *de la variable aleatòria X la funció $\varphi_X : \mathbb{R} \to \mathbb{R}$ definida per*

$$\varphi_X(z) = E[z^X]. \tag{3.23}$$

Veiem, doncs, que l'únic ingredient que necessitem per definir la funció generatriu d'una variable aleatòria és la seva llei. És a dir, *si dues variables aleatòries tenen la mateixa llei, també han de tenir la mateixa funció generatriu.* En efecte, a partir de (3.11) resulta que, si $\varphi_X(z)$ existeix, es compleix que

$$\varphi_X(z) = \sum_{n=0}^{\infty} z^n p(n). \tag{3.24}$$

Perquè la funció generatriu d'una variable aleatòria estigui ben definida, cal que la sèrie que apareix en el segon membre de (3.24) sigui absolutament convergent (en altres paraules, que la variable aleatòria z^X tingui esperança finita). En general, això serà així si $|z| \le 1$. En efecte, en aquest cas,

$$\sum_{n=0}^{\infty} |z|^n p(n) \le \sum_{n\in\mathbb{Z}^+} p(n) = 1.$$

A partir de la funció generatriu d'una variable aleatòria podem retrobar la seva llei, ço és, hem establert un procés de canvi *reversible.* Aquesta propietat és una conseqüència immediata de la definició (3.24). En efecte, hem donat la definició de la funció φ_X mitjançant el seu desenvolupament en sèrie de potències i, per resultats de l'anàlisi real, se sap que aquest desenvolupament és únic. Endemés, hi ha una fórmula que dóna els coeficients del desenvolupament. Aplicant aquests resultats al nostre cas obtenim

$$p(n) = \frac{1}{n!} \left.\frac{\mathrm{d}^n \varphi_X(z)}{\mathrm{d}z^n}\right|_{z=0}.$$

Donem tot seguit una de les propietats més importants de les funcions generatrius.

Proposició 3.14 *Siguin X i Y dues variables aleatòries a valors en $\mathbb{Z}^+$, independents. Aleshores,*

$$\varphi_{X+Y}(z) = \varphi_X(z)\varphi_Y(z),$$

per a tot $|z| \leq 1$.

Demostració: Fixem $|z| \leq 1$. Les variables aleatòries z^X i z^Y són independents i tenen moments de primer ordre finit. Aleshores, per la proposició 3.11 tindrem

$$\varphi_{X+Y}(z) = E\big[z^{X+Y}\big] = E\big[z^X z^Y\big] = E\big[z^X\big]E\big[z^Y\big] = \varphi_X(z)\varphi_Y(z).$$

Això acaba la demostració de la proposició. □

3.9.2 Càlcul d'algunes funcions generatrius

Llei binomial

Considerem una variable aleatòria X amb distribució $B(n,p)$. La seva funció generatriu es pot calcular de la manera següent:

$$\begin{aligned}\varphi_X(z) &= \sum_{k=0}^{n} z^k \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=0}^{n} \binom{n}{k} (zp)^k (1-p)^{n-k} \\ &= (1-p+zp)^n.\end{aligned} \tag{3.25}$$

El mateix resultat pot obtenir-se calculant primer la funció generatriu d'una variable aleatòria amb distribució de Bernoulli de paràmetre p i utilitzant després la proposició 3.14.

Llei de Poisson

Per a una variable aleatòria X amb distribució de Poisson de paràmetre λ es té que

$$\begin{aligned}\varphi_X(z) &= \mathrm{e}^{-\lambda}\sum_{n=0}^{\infty} z^n \frac{\lambda^n}{n!} \\ &= \mathrm{e}^{-\lambda}\sum_{n=0}^{\infty}\frac{(z\lambda)^n}{n!} = \mathrm{e}^{-\lambda}\mathrm{e}^{\lambda z} \\ &= \mathrm{e}^{\lambda(z-1)}.\end{aligned} \tag{3.26}$$

Llei geomètrica

La funció generatriu de la llei geomètrica (vegeu apartat 2.4.3) s'obté sumant una sèrie geomètrica:

$$\begin{aligned}\varphi_X(z) &= \sum_{k=0}^{\infty} z^{k+1}p(1-p)^k = z\sum_{k=0}^{\infty} p\big(z(1-p)\big)^k \\ &= \frac{zp}{1-z(1-p)}.\end{aligned} \tag{3.27}$$

Observeu que cal imposar $|z(1-p)| < 1$. Aquesta condició dóna els valors de z per als quals la funció generatriu està definida.

La proposició 3.14 i les expressions de les funcions generatrius de les lleis binomial i de Poisson (vegeu (3.25) i (3.26), respectivament) permeten d'obtenir un caràcter d'autoreproducció d'aquestes lleis.

Més precisament, la suma de dues variables aleatòries independents, X i Y, amb lleis $B(n_1, p)$ i $B(n_2, p)$, respectivament, és una variable aleatòria amb distribució $B(n_1 + n_2, p)$. El mateix tipus de propietat val també per a la llei de Poisson: la suma de dues variables aleatòries independents amb lleis Poiss(λ_1) i Poiss(λ_2), respectivament, és una variable aleatòria amb distribució Poiss($\lambda_1 + \lambda_2$).

Pel que fa a la llei binomial, el resultat que acabem de provar ja el coneixíem. En efecte, ja hem comentat abans que una variable aleatòria amb distribució $B(n, p)$ pot obtenir-se com la suma de n variables aleatòries independents amb llei $B(1, p)$.

3.9.3 Funcions generatrius i càlcul de moments

A partir de la funció generatriu d'una variable aleatòria es poden trobar els seus moments, sempre que es compleixin determinades condicions. En efecte, sigui ρ el radi de convergència de la sèrie del segon terme de (3.24); sabem que $\rho \geq 1$. Fixem $z \in \mathbb{R}$, $|z| < \rho$. Derivant (3.24) hom obté que

$$(\varphi_X)'(z) = \sum_{n=1}^{\infty} n z^{n-1} p(n).$$

Aleshores, si el radi de convergència de la sèrie en (3.24) és més gran que 1, podrem escriure

$$(\varphi_X)'(1) = \sum_{n=1}^{\infty} n p(n) = E(X).$$

Anàlogament, derivant dues vegades s'obté que

$$(\varphi_X)''(z) = \sum_{n=2}^{\infty} n(n-1) z^{n-2} p(n),$$

i, sota les mateixes condicions que abans,

$$\begin{aligned}(\varphi_X)''(1) &= \sum_{n=2}^{\infty} (n^2 - n) p(n) = \sum_{n=0}^{\infty} (n^2 - n) p(n) \\ &= E(X^2) - E(X).\end{aligned}$$

En particular,

$$\mathrm{Var}(X) = E(X^2) - E(X)^2 = (\varphi_X)''(1) + (\varphi_X)'(1) - (\varphi_X)'(1)^2.$$

Aquest procés pot iterar-se. En conclusió, si el radi de convergència de la sèrie que defineix la funció generatriu és més gran que 1, la llei té moments de tots els ordres, i aquests s'obtenen a partir de les derivades successives de la funció generatriu en el punt $z = 1$.

Per exemple, per la *llei geomètrica*, a partir de (3.27) s'obté que

$$(\varphi_X)'(z) = \frac{p}{(1-z(1-p))^2}.$$

Com que $p \in (0,1)$, la condició $|z(1-p)| < 1$ implica que el radi de convergència de la sèrie que defineix $\varphi_X(z)$ és més gran que 1. En conseqüència,

$$(\varphi_X)'(1) = \frac{1}{p} = E(X),$$

on X és una variable aleatòria amb llei geomètrica de paràmetre p.

D'altra banda, com que

$$(\varphi_X)''(z) = \frac{2p(1-p)}{(1-z(1-p))^3},$$

resulta

$$(\varphi_X)''(1) = \frac{2(1-p)}{p^2}.$$

Per tant, la variància de la llei geomètrica de paràmetre p és

$$\operatorname{Var} X = \frac{1-p}{p^2}.$$

Observeu que aquests càlculs també justifiquen que l'esperança d'una variable aleatòria *binomial negativa* és $\frac{n}{p}$, ja que tal variable aleatòria pot obtenir-se com a suma de n variables aleatòries independents amb llei geomètrica de paràmetre p.

3.10 Funcions generatrius de moments

En l'apartat 3.9.3 hem vist com les funcions generatrius permeten de trobar els moments de variables aleatòries discretes. Hi ha altres transformacions de la llei d'una variable aleatòria que aconsegueixen el mateix objectiu. Una és la que definim tot seguit.

Definició 3.6 *La* funció generatriu de moments *d'una variable aleatòria X és la funció $\psi_X(t) = E(\exp(tX))$, definida en un entorn de $t = 0$.*

Observem que, en la definició anterior, estem suposant que, per a tot t en un entorn de l'origen, la variable aleatòria $\exp(tX)$ té esperança matemàtica finita.

Clarament, si X i Y són variables aleatòries independents, per la proposició 3.11 obtenim que

$$\psi_{X+Y}(t) = \psi_X(t)\psi_Y(t),$$

per a tot t per al qual ambdós termes de la igualtat anterior tinguin sentit. Aquest resultat és l'anàleg a la proposició 3.14 per a funcions generatrius de moments.

El fet que la funció generatriu de moments caracteritza la llei de la variable aleatòria és una qüestió delicada que queda fora de l'abast d'aquest llibre.

Als efectes d'utilitat per als càlculs donem finalment un resultat sobre determinació de moments de qualsevol ordre.

Proposició 3.15 *Sigui X una variable aleatòria i ψ_X la seva funció generatriu de moments. Per a tot $k \in \mathbb{N}$, X té moment d'ordre k finit i*

$$E(X^k) = \psi_X^{(k)}(0). \tag{3.28}$$

La demostració d'aquesta proposició es basa en resultats de derivació d'integrals que depenen d'un paràmetre t. Podem però donar un argument sense rigor per a justificar la validesa de (3.28).

Si poguéssim invertir les operacions de derivació respecte de t i d'esperança matemàtica, obtindríem

$$\psi_X^{(k)}(t) = E(X^k \exp(tX)).$$

Per a $t = 0$ aquesta igualtat dóna (3.28).

Exercicis

3.1 *Sigui X una variable aleatòria absolutament contínua, amb funció de densitat donada per*

$$f(x) = \begin{cases} \frac{1}{2x^2}, & \textit{si } |x| \geq 1, \\ 0, & \textit{si } |x| > 1. \end{cases}$$

Justifiqueu quina de les afirmacions següents és correcta:

1. *Com que $f(x)$ és simètrica respecte de l'origen, el seu valor mitjà és zero.*

2. *Per consideracions de simetria,*

$$E(X) = \int_{\mathbb{R}} x f(x)\,\mathrm{d}x = 2\int_1^{\infty} \frac{1}{2x}\,\mathrm{d}x = \infty.$$

3. *L'esperança matemàtica de la variable aleatòria X no existeix.*

3.2 *Per quin valor del paràmetre b la funció*

$$F(x) = \begin{cases} 0, & \textit{si } 0 \leq x < 1, \\ a(x-1)^2, & \textit{si } 1 \leq x < 3, \\ b(2 - \exp(-(x-3))), & \textit{si } 3 \leq x. \end{cases}$$

és una funció de distribució? Donat aquest b, quins sóns els valors possibles de a? Sigui X una variable aleatòria amb funció de distribució F, definida anteriorment, amb $a = \frac{1}{8}$. Determineu si X és absolutament contínua i calculeu $E(X)$.

3.3 *Sigui $g : [0, b] \longrightarrow \mathbb{R}_+$ una funció contínua i estrictament decreixent. Considerem un vector aleatori (X, Y) amb densitat*

$$f(x, y) = \begin{cases} \alpha, & \textit{si } x \in [0, b], g(b) \leq y \leq g(x), \\ 0, & \textit{en cas contrari.} \end{cases}$$

1. *Trobeu α.*

2. *Trobeu les lleis marginals.*

3. *Són X i Y independents? Per què?*

4. *Proveu que $X + Y$ i $X - Y$ són incorrelacionades si, i només si,* $\mathrm{Var}(X) = \mathrm{Var}(Y)$.

3.4 *Sigui X una variable aleatòria amb funció de densitat*

$$f(t) = \left(\frac{1}{2} + at^3\right) \mathbb{1}_{(-1,1)}(t), \quad t \in \mathbb{R},$$

on a és un paràmetre real.

1. *Determineu els possibles valors de a per als quals f és una densitat.*

2. *Calculeu, en el cas que existeixin, els moments de primer i segon ordre de la variable aleatòria*

$$Y = \frac{1}{\sqrt{|X|}}.$$

3. *Calculeu la densitat de la variable aleatòria $Z = X^2 - X + 1$.*

3.5 *Siguin X, Y variables aleatòries independents, amb*

$$P(X = k) = \frac{1}{n}, k \in \{1, \dots, n\},$$
$$P(Y = k) = \frac{1}{2}, k = 0 \text{ o } k = n.$$

Calculeu la llei de la variable $X + Y$ i la seva esperança.

3.6 *Sigui X una variable aleatòria amb llei donada per*

$$P\{X = x\} = \begin{cases} \frac{1}{2k^2}, & \text{si } x = \mu + k\sigma, \text{ o } x = \mu - k\sigma, \\ 1 - \frac{1}{k^2}, & \text{si } x = \mu, \\ 0, & \text{en cas contrari.} \end{cases}$$

1. *Determineu si X té moments de qualsevol ordre.*

2. *Demostreu que $P\{|X - \mu| \geq k\sigma\} = \frac{1}{k^2}$. Escriviu una fita per a $P\{|X - \mu| \geq k\sigma\}$ a partir de la desigualtat de Txebixev.*

3. *Demostreu que la llei de X condicionada per $\{|X - \mu| \geq k\sigma\}$ és uniforme discreta, amb la mateixa mitjana que X.*

3.7 *Sigui (X,Y) un vector aleatori amb llei*

$$P\{(X,Y)=(i,j)\}=\begin{cases} c(i+j), & \text{per a } i\in\{1,\dots,m\}, j\in\{1,\dots,n\},\\ 0, & \text{en cas contrari.}\end{cases}$$

1. *Determineu la constant c perquè l'expressió anterior defineixi una distribució de probabilitat.*
2. *Trobeu les lleis marginals de X i Y. Comproveu que X i Y són independents.*
3. *Calculeu $E\left(\frac{1}{X+Y}\right)$.*

3.8 *Siguin X, Y dues variables aleatòries independents, amb lleis binomial $B(n,p)$ i $U(0,1)$, respectivament.*

1. *Calculeu la funció de distribució de $Z=X-Y$. És contínua? És Z una variable aleatòria absolutament contínua?*
2. *Calculeu, si existeixen, $E(Z)$ i $\mathrm{Var}(Z)$.*

3.9 *Per a $r,s>0$ es defineix*

$$f(x)=\begin{cases} cx^{-(s+1)}, & \text{si } x>r,\\ 0, & \text{si } x\le r.\end{cases}$$

1. *Determineu c perquè f sigui una densitat de probabilitat.*
2. *Sigui X una variable aleatòria amb densitat f. Per a quins valors de s la variable aleatòria X té esperança finita?*

3.10 *Sigui N una variable aleatòria amb distribució uniforme en $\{1,2,\dots,10\}$.*

1. *Considerem les variables aleatòries $X=\mathbb{1}_{\{N\le 5\}}$, $Y=\mathbb{1}_{\{N \text{ parell}\}}$. Sense fer cap càlcul, argumenteu el resultat següent:*

$$E(X)=E(Y)=\frac{1}{2}.$$

Són X i Y independents? Calculeu $\mathrm{Cov}(X,Y)$ i $E((X+Y)^2)$. Calculeu la funció generatriu de N.

2. *Sigui* $\{X_n, n = 1, \dots, 10\}$ *una família de variables aleatòries independents amb llei* $\text{Poiss}(\lambda)$, *independents també de la variable aleatòria* N. *Definim*

$$S_N(\omega) = X_1(\omega) + \cdots + X_{N(\omega)}(\omega).$$

Calculeu la funció generatriu de S_N.

Capítol 4

Successions de variables aleatòries

L'objecte d'estudi d'aquest capítol són les successions de variables aleatòries i els diferents tipus de convergència. Ja hem trobat anteriorment situacions en les quals es fa necessari introduir successions de variables aleatòries. Per exemple, per a l'estudi de la llei geomètrica o de la llei binomial negativa. Altres motivacions importants són l'aproximació de la probabilitat d'un esdeveniment per la successió de les seves freqüències relatives (llei feble dels grans nombres de Bernoulli) i l'aproximació de la llei binomial per la llei de Poisson o la llei normal.

Comencem el capítol amb una secció tècnica sobre successions d'esdeveniments que utilitzarem posteriorment. Introduïm després diversos tipus de convergència que corresponen a diferents maneres de mesurar la *distància* i els relacionem. Ultra el seu interès matemàtic, el contingut d'aquest capítol és bàsic per poder formular i estudiar els teoremes límit fonamentals del càlcul de probabilitats: les lleis dels grans nombres i el teorema del límit central.

Cal fer observar que la construcció de l'espai de probabilitat $(\Omega, \mathcal{A}, P)$ on podem considerar definida la successió de variables aleatòries $\{X_n, n \geq 1\}$ no és un problema trivial. Però aquí no entrarem en la discussió d'aquest punt.

4.1 Lemes de Borel-Cantelli

Aquesta secció és dedicada a l'estudi d'alguns resultats notables sobre successions d'esdeveniments. El seu contingut ens serà d'utilitat per a l'estudi de la convergència *quasi segura*. Comencem donant les definicions de *límit superior* i *límit inferior* d'una successió de conjunts. Per establir-les necessitem únicament una σ-àlgebra $\mathcal{A}$ de parts d'un conjunt Ω.

Definició 4.1 *Considerem una successió $\{A_n, n \geq 1\}$ de conjunts de $\mathcal{A}$. Definim*

$$\limsup_n A_n = \bigcap_{n\geq 1} \bigcup_{k\geq n} A_k,$$
$$\liminf_n A_n = \bigcup_{n\geq 1} \bigcap_{k\geq n} A_k.$$

A causa de l'estructura de σ-àlgebra de $\mathcal{A}$, clarament $\limsup_n A_n$ i $\liminf_n A_n$ són elements de $\mathcal{A}$.

D'aquestes definicions es dedueix fàcilment que

1. Un element ω pertany a $\limsup_n A_n$ si, i únicament si, ω pertany a infinits A_n.
2. Un element ω pertany a $\liminf_n A_n$ si, i únicament si, existeix un nombre natural n_0, que depèn de ω, tal que ω pertany a tots els A_n per $n \geq n_0$.

En efecte, si ω pertany a $\limsup_n A_n$, es tindrà que, per a tot $n \geq 1$, existeix $k \geq n$ tal que $\omega \in A_k$, i, en conseqüència, ω pertany a infinits conjunts A_k. Recíprocament, suposem que ω pertany a infinits conjunts A_k, que denotarem per $A_{k_1}, \ldots, A_{k_n}, \ldots$ Aleshores,

$$\omega \in \cap_{j\in\mathbb{N}} A_{k_j} \subset \cap_{n\geq 1} \cup_{k\geq n} A_k.$$

Suposem ara que ω pertany a $\liminf_n A_n$. Aleshores, existeix $n_0(\omega) \geq 1$ i per a tot $k \geq n_0(\omega)$, $\omega \in A_k$. En altres paraules, ω pertany a tots els conjunts A_k a partir d'un cert ordre $k = n_0(\omega)$. Recíprocament, si es compleix aquesta propietat, ω pertany a $\liminf_n A_n$.

A partir de les caracteritzacions expressades a (1) i (2) es dedueix fàcilment que

$$\liminf_n A_n \subset \limsup_n A_n.$$

D'altra banda, de les regles d'operacions amb conjunts s'obté que

$$(\limsup_n A_n)^c = \liminf_n A_n^c,$$

$$(\liminf_n A_n)^c = \limsup_n A_n^c.$$

Exemple 4.1 Considerem la σ-àlgebra $\mathcal{B}(\mathbb{R}^2)$, generada pels conjunts oberts de $\mathbb{R}^2$, i la successió de conjunts $A_n \in \mathcal{B}(\mathbb{R}^2)$ definida per

$$A_n = \{(x,y) \in \mathbb{R}^2 : n^2x^2 + n^4y^2 \geq 1\}, n \geq 1.$$

Es compleix $\limsup_n A_n = \liminf_n A_n = \mathbb{R}^2 - \{(0,0)\}$. Observeu que A_n correspon als punts del pla exteriors a l'el·lipse amb diàmetres $\frac{1}{n}$ i $\frac{1}{n^2}$ respectivament (vegeu la figura 4.1).

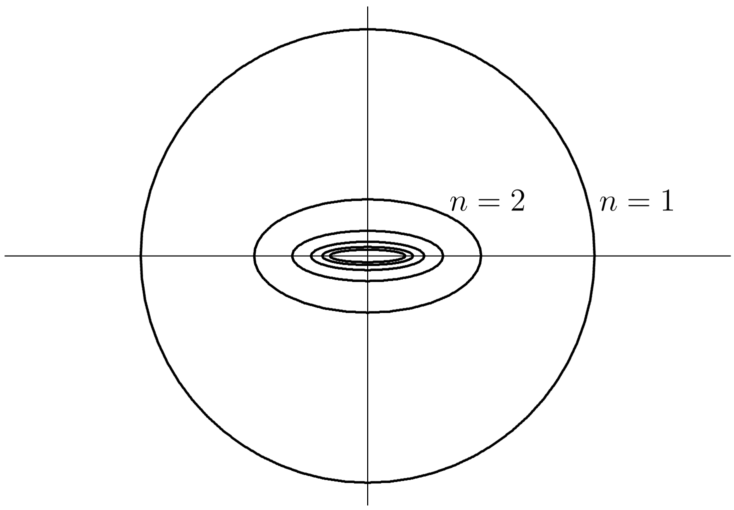

Figura 4.1: Exemple 4.1

Els conjunts A_n formen una successió *creixent*. Aquesta propietat implica $\limsup_n A_n = \liminf_n A_n = \cup_{n=1}^{\infty} A_n$. Vegem finalment que

$$\cup_{n=1}^{\infty} A_n = \mathbb{R}^2 - \{(0,0)\}.$$

En efecte, si $(x,y) \in \cup_{n=1}^{\infty} A_n$, existeix $n_0 \in \mathbb{N}$ tal que $n^2x^2 + n^4y^2 \geq 1$, per tant, $(x,y) \neq (0,0)$. Recíprocament, donat $(x,y) \in \mathbb{R}^2 - \{(0,0)\}$, existeix $n_1 \in \mathbb{N}$ tal que $(x,y) \in A_{n_1}$.

El lema següent serà útil a l'hora de relacionar diferents tipus de convergència. A partir d'ara suposem donat un espai de probabilitat de referència $(\Omega, \mathcal{A}, P)$.

Lema 4.1 *Donada una successió $\{A_n, n \geq 1\}$ d'esdeveniments de $\mathcal{A}$, es compleix que*

$$P\{\limsup_n A_n\} \geq \limsup_n P(A_n),$$

i

$$P\{\liminf_n A_n\} \leq \liminf_n P(A_n).$$

Demostració: Podem escriure

$$\begin{aligned} P\{\limsup_n A_n\} &= \lim_n P\Big\{\bigcup_{m \geq n} A_m\Big\} \geq \lim_n \Big(\sup_{m \geq n} P(A_m)\Big) \\ &= \limsup_n P(A_n). \end{aligned}$$

L'altre resultat es dedueix d'aquest per pas al complementari. En efecte,

$$\begin{aligned} P\{\liminf_n A_n\} &= 1 - P\Big((\liminf_n A_n)^c\Big) = 1 - P(\limsup_n A_n^c) \\ &\leq 1 - \limsup_n P(A_n^c) = \liminf_n P(A_n). \end{aligned}$$

Això acaba la demostració. $\square$

Presentem, tot seguit, els lemes de Borel-Cantelli, que donen condicions suficients (i en un cert sentit, també necessàries) perquè $P\{\limsup_n A_n\} = 0$.

Lema 4.2 (Primer lema de Borel-Cantelli) *Sigui $\{A_n, n \geq 1\}$ una successió de conjunts de $\mathcal{A}$. Si $\sum_{n=1}^{\infty} P(A_n) < \infty$, es compleix que*

$$P\{\limsup_n A_n\} = 0.$$

Demostració: Es té que

$$P\{\limsup_n A_n\} = P\left\{\bigcap_{n=1}^{\infty} \bigcup_{m \geq n} A_m\right\} = \lim_n P\left\{\bigcup_{m \geq n} A_m\right\}$$
$$\leq \lim_n \sum_{m \geq n} P(A_m) = 0$$

i, en conseqüència, el resultat. □

Lema 4.3 (Segon lema de Borel-Cantelli) *Sigui $\{A_n, n \geq 1\}$ una successió d'esdeveniments independents. Si $\sum_{n=1}^{\infty} P(A_n) = \infty$, aleshores*

$$P\{\limsup_n A_n\} = 1.$$

Demostració: Veurem que $P\left\{(\limsup_n A_n)^c\right\} = P\left\{\bigcup_{n=1}^{\infty} \bigcap_{m \geq n} A_m^c\right\} = 0$. I per a això demostrarem que $P\{\bigcap_{m \geq n} A_m^c\} = 0$ per a tot $n \geq 1$.
Utilitzant la desigualtat $1 - x \leq e^{-x}, x \geq 0$, i a causa de la independència, tenim que

$$P\left\{\bigcap_{m=n}^{n+j} A_m^c\right\} = \prod_{m=n}^{n+j} (1 - P(A_m)) \leq \prod_{m=n}^{n+j} e^{-P(A_m)}$$
$$= \exp\left(-\sum_{m=n}^{n+j} P(A_m)\right),$$

expressió que tendeix a zero quan j tendeix a infinit, a causa de la hipòtesi. □

Exemple 4.2 ([Bo]) Juguem a *cara o creu* amb una moneda que té probabilitat p que surti cara en llançar-la. Denotem per Y_n el resultat de la n-èsima tirada (1 si ha sortit cara, -1 si ha sortit creu) i definim $S_n = Y_1 + \cdots + Y_n$. L'esdeveniment $\{S_n = 0\}$ té probabilitat zero si n és senar. Si $n = 2k$,

$$P\{S_n = 0\} = \binom{2k}{k} p^k (1-p)^k.$$

Utilitzant l'equivalent de Stirling ($n! \sim \sqrt{2\pi n}\, n^n e^{-n}$) resulta que

$$P\{S_{2k} = 0\} \sim (\pi k)^{-\frac{1}{2}} (2p)^k \left(2(1-p)\right)^k.$$

Si $p \neq \frac{1}{2}$, la sèrie $\sum_n P\{S_n = 0\}$ convergeix i, pel primer lema de Borel-Cantelli, $P\{\limsup(S_n = 0)\} = 0$. És a dir, *per a una moneda trucada, amb probabilitat igual a* 1*, la variable aleatòria* S_n *només s'anul·la un nombre finit de vegades.*

Exemple 4.3 Sigui $\{X_n, n \geq 1\}$ una successió de variables aleatòries independents que prenen els valors 0 i 1 amb probabilitat $\frac{1}{2}$. Sigui $\alpha = (\alpha_1, \ldots, \alpha_j)$ una col·lecció ordenada de zeros i uns que considerem fixada. Tenim la propietat següent: Amb probabilitat 1, la col·lecció ordenada α apareix infinites vegades en la successió aleatòria $\{X_n, n \geq 1\}$.

En efecte, els esdeveniments

$$A_n = \{X_{(n-1)j+1} = \alpha_1, \ldots, X_{nj} = \alpha_j\}, n \geq 1,$$

són independents i tots tenen probabilitat 2^{-j}. Per tant, la sèrie $\sum_{n=1}^{\infty} P(A_n)$ divergeix i, pel segon lema de Borel-Cantelli, $P\{\limsup_n A_n\} = 1$, que és el que volíem provar.

Donada una successió d'esdeveniments de $\mathcal{A}$, $\{A_n, n \geq 1\}$, és clar que, o bé $0 \leq \sum_{n=1}^{\infty} P(A_n) < \infty$, o bé $\sum_{n=1}^{\infty} P(A_n) = \infty$. Per tant, com a conseqüència dels lemes 4.2 i 4.3 demostrats abans, tenim el resultat següent.

Lema 4.4 (Llei del zero-u) *Sigui* $\{A_n, n \geq 1\}$ *una successió d'esdeveniments independents d'un espai de probabilitat* $(\Omega, \mathcal{A}, P)$*. Aleshores, o bé* $P\{\limsup_n A_n\} = 0$*, o bé* $P\{\limsup_n A_n\} = 1$*, segons que* $\sum_{n=1}^{\infty} P(A_n) < \infty$ *o* $\sum_{n=1}^{\infty} P(A_n) = \infty$.

4.2 Convergència quasi segura

En aquest apartat introduïm un primer tipus de convergència de successions de variables aleatòries, que és el més semblant a la convergència puntual de successions de funcions. Donem també criteris per obtenir aquesta convergència.

Definició 4.2 *Direm que una successió de variables aleatòries* $\{X_n,\ n \geq 1\}$ *convergeix* quasi segurament *cap a una variable aleatòria* X *si existeix un conjunt* $N \in \mathcal{A}$ *de probabilitat zero tal que*

$$\lim_{n\to\infty} X_n(\omega) = X(\omega),$$

per a tot $\omega \notin N$.

Si X_n convergeix quasi segurament cap a X, quan n tendeix cap a ∞, escriurem $X_n \longrightarrow X$, q.s. o bé $\lim_{n\to\infty} X_n = X$, q.s.
Es tracta, doncs, d'una convergència puntual de la successió numèrica $X_n(\omega)$ en tots els punts $\omega \in \Omega - N$.
La variable límit és única llevat en conjunts de probabilitat zero. És clar, també, que la successió $\{X_n, n \geq 1\}$ convergeix quasi segurament si, i només si, és de Cauchy amb probabilitat 1.

Exemple 4.4 Considerem l'espai de probabilitat format per $\Omega = [0,1]$, $\mathcal{A}$, la σ-àlgebra de Borel de $[0,1]$, és a dir, la σ-àlgebra generada pels conjunts oberts de $[0,1]$; com a probabilitat P considerem la mesura de Lebesgue, és a dir, la distribució uniforme en $[0,1]$ (la *longitud*). Aquest és un espai de probabilitat molt útil per proposar exemples sobre convergències.
Definim

$$X_n(x) = \begin{cases} 2^n, & \text{si } 0 \leq x \leq \frac{1}{n}, \\ 0, & \text{en cas contrari.} \end{cases}$$

És fàcil comprovar que $X_n \longrightarrow 0$, q.s.
En efecte, si $x \in (0,1]$, podem trobar n_0 tal que per a tot $n \geq n_0$, $x \in \left(\frac{1}{n}, 1\right]$ i, per tant, $X_n(x) = 0$ per a tot $n \geq n_0$. És a dir,

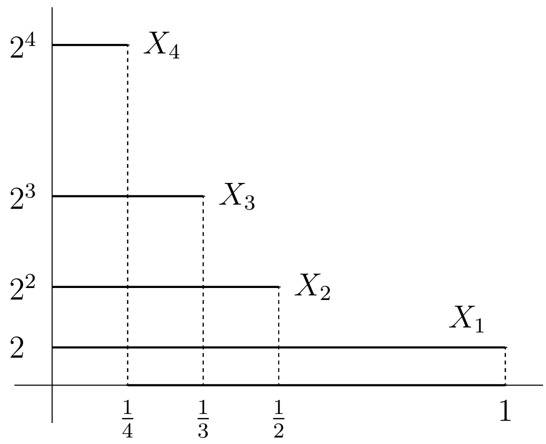

Figura 4.2: Exemple 4.4

$$\lim_{n\to\infty} X_n(x) = 0,$$

si $x \in (0, 1]$. Si $x = 0$, $\lim_{n\to\infty} X_n(x) = \infty$. Però la distribució uniforme en $[0, 1]$ assigna probabilitat zero a cada punt. En altres paraules, un punt de $\mathbb{R}$ té longitud (probabilitat) zero. Així, doncs, $\lim_{n\to\infty} X_n(x) = 0$, q.s.

El tractament de l'exemple precedent depèn molt del cas particular de l'espai de probabilitat. En espais de probabilitat generals, l'estudi de la convergència quasi segura no és un problema fàcil. El resultat següent ens dóna una condició equivalent per a la convergència quasi segura d'una successió de variables aleatòries.

Proposició 4.1 *La successió $\{X_n, n \geq 1\}$ convergeix quasi segurament cap a la variable aleatòria X si, i únicament si, per a tot $\varepsilon > 0$*

$$P\left\{\liminf_n \{|X_n - X| \leq \varepsilon\}\right\} = 1.$$

Demostració: Suposem primer que $\lim_{n\to\infty} X_n = X$, q.s. i sigui N el subconjunt de $\mathcal{A}$ de probabilitat zero on falla la convergència puntual. Designem per Ω_0 el conjunt $\Omega - N$. Es compleix que

$$\{\sup_{n\geq m}|X_n - X| \leq \varepsilon\} = \bigcap_{n\geq m}\{|X_n - X| \leq \varepsilon\} := A_m^\varepsilon.$$

Els conjunts A_m^ε formen una successió creixent en m. D'altra banda, es té la inclusió

$$\Omega_0 \subset \bigcup_{m=1}^\infty A_m^\varepsilon.$$

Resulta, doncs,

$$1 = P(\Omega_0) \leq P\left(\bigcup_{m=1}^\infty A_m^\varepsilon\right) = \lim_{m\to\infty} P(A_m^\varepsilon),$$

i, en conseqüència,

$$P\left\{\liminf_n\{|X_n - X| \leq \varepsilon\}\right\} = \lim_{m\to\infty} P(A_m^\varepsilon) = 1.$$

Recíprocament, fixem un nombre racional $\varepsilon > 0$. Podem determinar un conjunt N_ε de probabilitat zero tal que per a tot $\omega \notin N_\varepsilon$, existeix $m_0(\varepsilon)$ i per a tot $m \geq m_0, |X_m(\omega) - X(\omega)| \leq \varepsilon$. Sigui $N = \bigcup_{\varepsilon\in\mathbb{Q}^+} N_\varepsilon$. Observeu que $P(N) = 0$, car N és una reunió numerable de conjunts de probabilitat zero. Aleshores, per a tot $\omega \notin N$ tenim $\lim_{n\to\infty} X_n(\omega) = X(\omega)$, és a dir, la convergència q.s. de la successió $\{X_n, n \geq 1\}$ cap a X. Això acaba la demostració de la proposició. □

El resultat següent és un criteri de convergència quasi segura que resulta d'aplicar el primer lema de Borel-Cantelli.

Proposició 4.2 *Sigui $\{\varepsilon_n, n \geq 1\}$ una successió de nombres reals positius tal que $\sum_{n=1}^\infty \varepsilon_n < \infty$. Sigui $\{X_n, n \geq 1\}$ una successió de variables aleatòries i suposem que*

$$\sum_{n=1}^\infty P\{|X_{n+1} - X_n| \geq \varepsilon_n\} < \infty.$$

Aleshores, $\{X_n, n \geq 1\}$ convergeix quasi segurament.

Demostració: Sigui $A_n = \{|X_{n+1} - X_n| > \varepsilon_n\}$. Com que per hipòtesi $\sum_{n=1}^{\infty} P(A_n) < \infty$, aplicant el primer lema de Borel-Cantelli s'obté que

$$P(\limsup_n A_n) = 0$$

o, equivalentment, $P(\liminf_n A_n^c) = 1$. Per a tot ω de $\liminf_n A_n^c$ existeix un $n_0(\omega)$ tal que $\omega \in A_n^c$ per a tot $n \geq n_0(\omega)$. Per tant, si $n \geq n_0(\omega)$, tindrem que $|X_{n+1}(\omega) - X_n(\omega)| \leq \varepsilon_n$. És a dir, la sèrie $\sum_{n=1}^{\infty}[X_{n+1}(\omega) - X_n(\omega)]$ convergeix absolutament amb probabilitat 1 (per a un conjunt de ω de probabilitat 1). D'aquí es dedueix que la successió

$$X_n(\omega) = X_1(\omega) + \sum_{k=1}^{n-1}[X_{k+1}(\omega) - X_k(\omega)]$$

és convergent quasi segurament. □

A nivell pràctic, l'esforç necessari per a aplicar les proposicions 4.1 i 4.2 és similar. En efecte, la condició que apareix en la proposició 4.1 és equivalent a

$$P\{\limsup_n \{|X_n - X| \geq \varepsilon\} = 0.$$

Si $\sum_{n=1}^{\infty} P\{|X_n - X| \geq \varepsilon\} < \infty$, el primer lema de Borel-Cantelli assegura la seva validesa. Observem també que, en la proposició 4.1 apareix un candidat a variable aleatòria límit, mentre que en la proposició 4.2 es demostra únicament l'existència del límit.

Exemple 4.5 Considerem una successió de variables aleatòries independents $\{X_n, n \geq 1\}$, amb $P\{X_n = 0\} = 1 - \frac{1}{n}$, $P\{X_n = 1\} = \frac{1}{n}$. Fixem $\varepsilon > 0$. La sèrie $\sum_{n=1}^{\infty} P\{|X_{n^2}| > \varepsilon\}$ convergeix. En efecte,

$$\sum_{n=1}^{\infty} P\{|X_{n^2}| > \varepsilon\} = \sum_{n=1}^{\infty} P\{X_{n^2} = 1\} = \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty.$$

Pel primer lema de Borel-Cantelli,

$$P\left\{\limsup_n \{|X_{n^2}| \geq \varepsilon\}\right\} = 0$$

i, per la proposició 4.1, deduïm que

$$\lim_{n\to\infty} X_{n^2} = 0, \text{ q.s.}$$

En canvi, la successió $\{X_n, n \geq 1\}$ no convergeix quasi segurament cap a 0. En efecte, si això fos cert, per la proposició 4.1 es tindria que

$$P\left\{\limsup_n \{|X_n| > \varepsilon\}\right\} = 0,$$

per a tot $\varepsilon > 0$. Ara bé,

$$\sum_{n=1}^{\infty} P\{|X_n| > \varepsilon\} = \sum_{n=1}^{\infty} P\{|X_n| = 1\} = \sum_{n=1}^{\infty} \frac{1}{n} = \infty.$$

Com que els esdeveniments $\{|X_n| > \varepsilon\}$ són independents, ja que les variables aleatòries X_n ho són, el segon lema de Borel-Cantelli implica $P\left\{\limsup_n \{|X_n| > \varepsilon\}\right\} = 1$.

4.3 Convergència en probabilitat

En la convergència quasi-segura, la probabilitat s'utilitza per "mesurar" el conjunt de punts $\omega \in \Omega$ per als quals falla la convergència puntual. En aquesta secció introduïm una nova noció de convergència, la convergència en probabilitat, en la qual la probabilitat juga un paper més important: serveix per mesurar el grau d'apropament (o distanciament) dels termes de la successió al límit hipotètic.

Comencem donant la definició, estudiant les propietats més importants i, finalment, demostrant que és una convergència metritzable. En la teoria de la mesura, el concepte anàleg al de convergència en probabilitat s'anomena "convergència en mesura".

Definició 4.3 *Direm que una successió de variables aleatòries* $\{X_n, n \geq 1\}$ convergeix en probabilitat *cap a una variable aleatòria* X *si, per a tot* $\varepsilon > 0$,

$$\lim_{n\to\infty} P\{|X_n - X| > \varepsilon\} = 0. \tag{4.1}$$

Escriurem $P - \lim X_n = X$.

La definició anterior expressa que, en augmentar n, és cada vegada menys probable que X_n i X difereixin en més de ε, per a tot $\varepsilon > 0$ fixat.

Més endavant demostrarem que el límit en probabilitat d'una successió de variables aleatòries, si existeix, és únic.

La definició 4.3 s'estén de manera natural a successions de vectors aleatoris. Únicament cal substituir en (4.1) el valor absolut $|\cdot|$ per la norma $\|\cdot\|$ de l'espai $\mathbb{R}^m$ on prenen els seus valors els vectors aleatoris.

La successió de variables aleatòries definides en l'exemple 4.4 convergeix en probabilitat cap a 0.

Exemple 4.6 Sigui $\Omega = [0,1]$, $\mathcal{A}$ la σ-àlgebra de Borel de $[0,1]$ i P la llei uniforme en $[0,1]$ Per a tot $n \geq 1, 1 \leq m \leq n$ definim

$$X_{n,m}(x) = \begin{cases} 1, & \text{si } \frac{m-1}{n} < x \leq \frac{m}{n}, \\ 0, & \text{en cas contrari.} \end{cases}$$

La successió de variables aleatòries $\{X_{n,m}, n \geq 1, 1 \leq m \leq n\}$ convergeix en probabilitat cap a 0. En efecte,

$$P\{|X_{n,m}| > \varepsilon\} = \begin{cases} 0, & \text{si } \varepsilon > 1, \\ P\{x : X_{n,m}(x) = 1\}, & \text{si } \varepsilon \leq 1. \end{cases}$$

La convergència cap a zero es dedueix del fet que $P\{x : X_{n,m}(x) = 1\} = \frac{1}{n}$.

En l'exemple anterior hem pogut fer un càlcul exacte de la successió $P\{|X_n - X| > \varepsilon\}$ que apareix en la definició de convergència en probabilitat. D'altres vegades, només és possible fer-ne una fitació, per exemple, mitjançant la desigualtat de Txebixev. Aquest és el mètode emprat en el teorema 5.2 per a demostrar la convergència en probabilitat de les freqüències relatives d'un esdeveniment a la seva probabilitat.

El resultat següent permet d'operar amb successions pel que fa a la convergència en probabilitat.

Proposició 4.3 *Siguin $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ una funció contínua, i $\{X_n, n \geq 1\}, \{Y_n, n \geq 1\}$ sengles successions de variables aleatòries que convergeixen en probabilitat cap a les variables aleatòries X i Y, respectivament. Aleshores, la successió $\{f(X_n, Y_n), n \geq 1\}$ convergeix en probabilitat cap a $f(X, Y)$.*

Demostració: Fixem $\varepsilon > 0, \eta > 0$. És fàcil comprovar que existeix $k > 0$ tal que $P\{|X| > \frac{k}{2}\} \leq \eta$ i $P\{|Y| > \frac{k}{2}\} \leq \eta$. En efecte, per a una variable aleatòria Z qualsevol, es compleix que

$$\lim_{n\to\infty} P\{|Z| > n\} = 0,$$

com a conseqüència de les propietats asimptòtiques de les funcions de distribució. Com que f és uniformement contínua en el compacte $[-k,k]^2$, existirà $\delta > 0$ tal que, si $|x'|$, $|x|$, $|y'|$, $|y| \leq k$, $|x-x'| \leq \delta$, $|y-y'| \leq \delta$, aleshores $|f(x,y) - f(x',y')| \leq \varepsilon$. Considerem els conjunts

$$\begin{aligned}
A &= \{|f(X_n,Y_n) - f(X,Y)| > \varepsilon\},\\
B &= \{|X| \leq \frac{k}{2}, |Y| \leq \frac{k}{2}, |X_n - X| \leq \frac{k}{2}, |Y_n - Y| \leq \frac{k}{2}\}.
\end{aligned}$$

Escrivim

$$P(A) = P(A \cap B) + P(A \cap B^c).$$

Per la desigualtat triangular,

$$B \subset \{|X| \leq k, |Y| \leq k, |X_n| \leq k, |Y_n| \leq k\}.$$

Tenim, doncs,

$$\begin{aligned}
P\{|f(X_n,Y_n) - f(X,Y)| > \varepsilon\} &\leq P\big\{|X| > k/2\big\}\\
&+ P\big\{|Y| > k/2\big\} + P\big\{|X - X_n| > k/2\big\}\\
&+ P\big\{|Y - Y_n| > k/2\big\} + P\big\{|X| \leq k,\\
&\quad |Y| \leq k, |X_n| \leq k, |Y_n| \leq k, |f(X_n,Y_n) - f(X,Y)| > \varepsilon\big\}\\
&\leq 2\eta + P\big\{|X - X_n| > k/2\big\} + P\big\{|Y - Y_n| > k/2\big\}\\
&+ P\{|X - X_n| > \delta\} + P\{|Y - Y_n| > \delta\}.
\end{aligned}$$

Fent tendir n cap a infinit, i tenint en compte que $\eta > 0$ és arbitrari, obtenim la convergència desitjada. $\square$

Com a conseqüència immediata de la proposició anterior tenim que el límit d'una suma (producte) de successions convergents en probabilitat és la suma (producte) dels límits respectius.

Acabem aquesta secció amb un resultat que interpreta la convergència en probabilitat com la convergència pròpia d'un espai mètric adequat.

En el conjunt de les variables aleatòries $X : \Omega \longrightarrow \mathbb{R}$ definim la relació d'equivalència següent:

$$X \sim Y \iff P\{X = Y\} = 1.$$

Si $X \sim Y$, direm que $X = Y$ quasi segurament.

Denotem per $L^0(\Omega, \mathcal{A}, P)$ el conjunt de les classes d'equivalència de variables aleatòries mòdul aquesta relació i definim l'aplicació $d : L^0(\Omega, \mathcal{A}, P) \times L^0(\Omega, \mathcal{A}, P) \to \mathbb{R}$ mitjançant

$$d(X, Y) = E\left(\frac{|X - Y|}{1 + |X - Y|}\right). \tag{4.2}$$

Proposició 4.4 *L'aplicació d definida en (4.2) és una distància sobre el conjunt $L^0(\Omega, \mathcal{A}, P)$ que metritza la convergència en probabilitat.*

Demostració: Provem en primer lloc que d és una distància. Clarament d és simètrica i $d(X, Y) = 0$ si, i únicament si, $X = Y$ en $L^0(\Omega, \mathcal{A}, P)$, ja que, si una variable aleatòria positiva té esperança 0, $X \sim 0$. Queda, doncs, per demostrar la desigualtat triangular.

Donats $x, y, z \in \mathbb{R}$ és immediat comprovar que

$$\frac{|x + y|}{1 + |x + y|} \leq \frac{|x|}{1 + |x|} + \frac{|y|}{1 + |y|}.$$

Prenent $x = X - Y$, $y = Y - Z$ i prenent esperances, resulta que

$$d(X, Z) \leq d(X, Y) + d(Y, Z).$$

Demostrem ara que d metritza la convergència en probabilitat. Considerem una successió de variables aleatòries $\{X_n, n \geq 1\}$ convergent en probabilitat cap a X. Podem suposar, prenent la successió $Y_n = X_n - X$, $X = 0$. Donat $\varepsilon > 0$, sigui $n_0 \in \mathbb{N}$ tal que per a tot $n \geq n_0$,

$$P\{|X_n| > \varepsilon\} < \varepsilon.$$

Denotem per A_n^ε el conjunt $\left\{\frac{|X_n|}{1+|X_n|} > \frac{\varepsilon}{1+\varepsilon}\right\}$ i considerem la descomposició

$$E\left(\frac{|X_n|}{1+|X_n|}\right) = E\left(\frac{|X_n|}{1+|X_n|}\mathbb{1}_{A_n^\varepsilon}\right) + E\left(\frac{|X_n|}{1+|X_n|}\mathbb{1}_{(A_n^\varepsilon)^c}\right). \tag{4.3}$$

El primer sumand del segon membre de la igualtat anterior està fitat per ε, ja que la variable aleatòria $\frac{|X_n|}{1+|X_n|}$ està fitada per 1 i el conjunt A_n^ε es pot expressar de manera equivalent com $\{|X_n| > \varepsilon\}$. Per hipòtesi, aquest conjunt té probabilitat inferior a ε. Pel que fa al segon sumand de (4.3), utilitzem que, sobre el conjunt $(A_n^\varepsilon)^c$, la variable aleatòria $\frac{|X_n|}{1+|X_n|}$ està fitada per ε. En conseqüència,

$$E\left(\frac{|X_n|}{1+|X_n|}\right) < 2\varepsilon.$$

Això prova

$$\lim_{n\to\infty} d(X_n, 0) = 0.$$

Recíprocament, suposem ara $\lim_{n\to\infty} d(X_n, 0) = 0$. Es compleix que

$$E\left(\frac{|X_n|}{1+|X_n|}\right) \geq E\left(\frac{|X_n|}{1+|X_n|}\mathbb{1}_{A_n^\varepsilon}\right) > \frac{\varepsilon}{1+\varepsilon}P(A_n^\varepsilon).$$

D'on,

$$P\{|X_n| > \varepsilon\} = P(A_n^\varepsilon) < \frac{1+\varepsilon}{\varepsilon}E\left(\frac{|X_n|}{1+|X_n|}\right).$$

Per tant, la successió $\{X_n, n \geq 1\}$ convergeix en probabilitat cap a 0. $\square$

De la proposició anterior es dedueix que el límit en probabilitat d'una successió de variables aleatòries, si existeix, és únic.

Hi ha altres distàncies en $L^0(\Omega, \mathcal{A}, P)$ que metritzen la convergència en probabilitat. Per exemple, la mètrica de Ky Fan definida per

$$\tilde{d}(X, Y) = \inf\{\varepsilon \geq 0 : P(|X - Y| > \varepsilon) \leq \varepsilon\}.$$

Definició 4.4 *Una successió* $\{X_n, n \geq 1\}$ *de variables aleatòries és de Cauchy en probabilitat si*

$$P - \lim_{n,m\to\infty} (X_m - X_n) = 0.$$

Si una successió de variables aleatòries convergeix en probabilitat, és de Cauchy en probabilitat. En efecte, de la desigualtat triangular es dedueix la inclusió

$$\{|X_m - X_n| > \varepsilon\} \subset \{|X_m - X| > \frac{\varepsilon}{2}\} \cup \{|X_n - X| > \frac{\varepsilon}{2}\}.$$

Més endavant provarem que l'espai $L^0(\Omega, \mathcal{A}, P)$ amb la convergència en probabilitat és un espai mètric *complet*, és a dir, tota successió de Cauchy en probabilitat té límit en probabilitat.

4.4 Convergència en mitjana d'ordre p

Per a $p \in [1, \infty)$, definim $L^p(\Omega, \mathcal{A}, P)$ com el conjunt de classes d'equivalència de variables aleatòries, mòdul la relació

$$X \sim Y \iff P\{X = Y\} = 1,$$

que tenen moment d'ordre p finit, ço és, $E|X|^p < \infty$. Per a successions de variables aleatòries en aquest conjunt hom pot estudiar la convergència en *mitjana*. El cas més utilitzat en aquest llibre és $p = 2$, que s'anomena *convergència en mitjana quadràtica*.

Definició 4.5 *Sigui* $\{X_n, n \geq 1\}$ *una successió de variables aleatòries de* $L^p(\Omega, \mathcal{A}, P)$. *Direm que aquesta successió convergeix en* mitjana d'ordre p *cap a una variable aleatòria* X*, que té també moment d'ordre* p *finit, si*

$$\lim_{n\to\infty} E(|X_n - X|^p) = 0.$$

Escriurem $L^p - \lim X_n = X$.

Exemple 4.7 Considerem el mateix espai de probabilitat de l'exemple 4.4. Donat un nombre natural $n \geq 1$ existeixen dos únics naturals p i q tals que

$n = 2^p + q, p \geq 0, 0 \leq q \leq 2^p$. Considerem la successió de variables aleatòries definida per $X_n = \mathbb{1}_{[q2^{-p},(q+1)2^{-p}]}$. Per a tot $p \geq 1, \{X_n, n \geq 1\}$ convergeix en mitjana d'ordre p cap a 0. En efecte,

$$E(|X_n|^p) = \int_{q2^{-p}}^{(q+1)2^{-p}} \mathrm{d}x = \frac{1}{2^p},$$

que tendeix a zero quan n tendeix a infinit.

Exemple 4.8 Sigui $\{X_n, n \geq 1\}$ una successió de variables aleatòries amb

$$P\{X_n = n\} = \frac{1}{n^2}, \qquad P\{X_n = \frac{1}{n}\} = 1 - \frac{1}{n^2}.$$

Hi ha convergència en probabilitat d'aquesta successió cap a la variable aleatòria $X = 0$. En efecte, fixat $\varepsilon \in (0,1)$ existeix $n_0 \in \mathbb{N}$ tal que per a tot $n \geq n_0$, es té $\frac{1}{n} < \varepsilon$. En conseqüència,

$$P\{|X_n| > \varepsilon\} = P\{X_n = n\} = \frac{1}{n^2} \xrightarrow[n\to\infty]{} 0.$$

En canvi, no hi ha convergència en mitjana quadràtica cap a zero, ja que

$$E\left(|X|^2\right) = 1 + \frac{1}{n^2} - \frac{1}{n^4} \xrightarrow[n\to\infty]{} 1.$$

Utilitzant la desigualtat de Hölder pot demostrar-se la inclusió $L^p(\Omega, \mathcal{A}, P) \subset L^q(\Omega, \mathcal{A}, P)$ i també que la convergència en mitjana d'ordre p implica la d'ordre q, si $1 < q < p$.

4.5 Convergència en llei

El tipus de convergència que estudiem en aquest apartat es formula en termes de les lleis de les variables aleatòries de la successió i del possible límit. És un cas particular de la convergència feble de mesures en espais mètrics i la noció de convergència adequada per a formular el teorema del límit central.

Definició 4.6 *Sigui $\{X_n, n \geq 1\}$ una successió de variables aleatòries i $\{F_n, n \geq 1\}$ la successió de les seves funcions de distribució respectives. Di-*

rem que la successió $\{X_n, n \geq 1\}$ convergeix en llei *(o* en distribució*) cap a una variable aleatòria* X *que té funció de distribució* F*, si es compleix que*

$$\lim_{n\to\infty} F_n(x) = F(x),$$

per a tot $x \in \mathbb{R}$ *on la funció* F *sigui contínua.*

Denotarem aquest tipus de convergència mitjançant el simbolisme

$$\mathcal{L} - \lim_{n\to\infty} X_n = X.$$

El límit en distribució d'una successió de variables aleatòries no és, en general, únic. Només la llei de la variable aleatòria límit queda identificada sense ambigüitat.

En el cas de variables aleatòries a valors en un subconjunt S de $\mathbb{R}$ numerable i sense punts d'acumulació es poden demostrar criteris molt útils de la convergència en llei. Per simplificar l'exposició suposarem $S = \mathbb{Z}_+$. Els dos resultats següents corresponen a aquesta situació.

Proposició 4.5 *Sigui* $\{X_n, n \geq 1\}$ *una successió de variables aleatòries i* X *una variable aleatòria, totes a valors en el conjunt dels nombres enters positius. Una condició necessària i suficient perquè* $\mathcal{L} - \lim_{n\to\infty} X_n = X$ *és que, per a tot* $k \in \mathbb{Z}_+$*, es compleixi que*

$$\lim_{n\to\infty} P\{X_n = k\} = P\{X = k\}. \tag{4.4}$$

Demostració: Suposem primer que hi ha convergència en llei. En tot punt de la forma $k + \frac{1}{2}$ la funció de distribució F de la variable aleatòria X és contínua. En conseqüència,

$$\begin{aligned}
\lim_{n\to\infty} P\{X_n = k\} &= \lim_{n\to\infty} \left(F_n\left(k + \frac{1}{2}\right) - F_n\left(k - \frac{1}{2}\right)\right) \\
&= F\left(k + \frac{1}{2}\right) - F\left(k - \frac{1}{2}\right) = P\{X = k\}.
\end{aligned}$$

Recíprocament, suposem que es compleix (4.4). Sigui $x \in \mathbb{R}$, Aleshores,

$$F_n(x) = P\{X_n \le x\} = \sum_{k=0}^{[x]} P\{X_n = k\}$$

$$\xrightarrow[n\to\infty]{} \sum_{k=0}^{[x]} P\{X = k\} = P\{X \le x\} = F(x).$$

Això acaba la demostració del criteri. $\square$

Proposició 4.6 *Sigui* $\{p_k^{(n)}, k \ge 0\}$, $n \ge 1$, *una successió de distribucions de probabilitat sobre* $\mathbb{Z}_+$. *Hi ha equivalència entre els dos enunciats següents:*

1. *Existeix una successió* $\{p_k^{(0)}, k \ge 0\}$ *tal que*

$$\lim_{n\to\infty} p_k^{(n)} = p_k^{(0)}, \quad k \ge 0.$$

2. *Existeix una funció* $\varphi(s)$, $0 < s < 1$, *tal que*

$$\lim_{n\to\infty} \varphi_n(s) = \varphi(s), \quad s \in (0,1),$$

on $\varphi_n(s)$ *és la funció generatriu de la distribució de probabilitat* $\{p_k^{(n)}$, $k \ge 0\}$, *és a dir,*

$$\varphi_n(s) = \sum_{k=0}^{\infty} p_k^{(n)} s^k.$$

Si es compleix (2) i, per tant, (1), $\varphi(s) = \sum_{k=0}^{\infty} p_k^{(0)} s^k$, *i,*

$$\sum_{k=0}^{\infty} p_k^{(0)} = 1 \iff \lim_{s\uparrow 1} \varphi(s) = 1.$$

Abans de donar la demostració d'aquesta proposició, farem uns quants comentaris per aclarir el seu significat.

Suposem que $\{p_k^{(0)}, k \geq 0\}$ sigui una distribució de probabilitat sobre $\mathbb{Z}_+$, és a dir, es compleixi la condició $\sum_{k=0}^{\infty} p_k^{(0)} = 1$. La proposició 4.5 ens diu que, si es compleix (1), hi ha convergència en distribució (de la successió de variables aleatòries amb llei donada per $\{p_k^{(n)}, k \geq 0\}$ cap a la variable aleatòria amb llei $\{p_k^{(0)}, k \geq 0\}$). Aquesta convergència és equivalent (vegeu (2)) a la convergència de la successió de les funcions generatrius cap a la funció generatriu del límit.

La condició (1) de la proposició 4.6 implica $p_k^{(0)} \in [0, 1]$, però no necessàriament $\sum_{k=0}^{\infty} p_k^{(0)} = 1$. En altres paraules, el límit pot no definir una distribució de probabilitat. L'exemple següent posa en evidència aquest fet. Definim $p_k^{(n)} = \delta_{k,n}$, la delta de Kronecker. En aquest cas, $\lim_{n\to\infty} p_k^{(n)} = 0$, per a tot $k \geq 0$. La darrera part de la proposició dóna una condició necessària i també suficient perquè el límit $\{p_k^{(0)}, k \geq 0\}$ defineixi una distribució de probabilitat sobre $\mathbb{Z}_+$.

Demostració de la proposició 4.5: Fixem $s \in (0, 1)$, $\varepsilon > 0$ i $m \in \mathbb{N}$ tal que

$$\sum_{i=m+1}^{\infty} s^i < \varepsilon.$$

Definim $\varphi(s) = \sum_{k=0}^{\infty} p_k^{(0)} s^k$ i suposem que es compleix (1). Aleshores,

$$|\varphi_n(s) - \varphi(s)| \leq \sum_{k=0}^{\infty} |p_k(n) - p_k^{(0)}| s^k \leq \sum_{k=0}^{m} |p_k(n) - p_k^{(0)}| + 2 \sum_{k=m+1}^{\infty} s^k.$$

Per tant,

$$\lim_{n\to\infty} |\varphi_n(s) - \varphi(s)| \leq 2\varepsilon$$

i, com que ε és arbitrari, això prova (2).

L'element $\{p_k^{(n)}, k \geq 0\}$ pertany a l'espai *compacte* $[0, 1]^{\mathbb{N}}$. Existeix, doncs, una subsuccessió $\{p_{k_j}^{(n)}, j \geq 0\}$ de $\{p_k^{(n)}, k \geq 0\}$ convergent. Com que ja hem demostrat que de (1) se'n segueix (2), podem afirmar

$$\lim_{j\to\infty} \varphi_{k_j}(s) = \varphi(s), \quad s \in (0,1).$$

D'aquí deduïm que tota successió parcial $\{p_{k_j}^{(n)}, j \geq 0\}$ de $\{p_k^{(n)}, k \geq 0\}$ convergent té necessàriament el mateix límit. Això implica que la successió $\{p_k^{(n)}, k \geq 0\}$ convergeix quan n tendeix cap a infinit i, per tant, es compleix (1). □

Exemple 4.9 Considerem una successió de variables aleatòries $\{X_n, n \geq 1\}$ tals que $P\{X_n = a_n\} = 1$. Suposem, a més, que la successió $\{a_n, n \geq 1\}$ convergeix cap a $a \in \mathbb{R}$ quan $n \to \infty$. Aleshores, la successió $\{X_n, n \geq 1\}$ convergeix en llei cap a una variable aleatòria X tal que $P\{X = a\} = 1$.

En efecte, la funció de distribució de X té un únic punt de discontinuïtat en a. Fixem $x \neq a$. Si $x < a$, $F(x) = 0$. D'altra banda, existeix $n_0 \geq 1$ tal que, per a tot $n \geq n_0$, $F_n(x) = 0$. Per tant, hi ha convergència de la successió $F_n(x), n \geq 1$ cap a $F(x)$. Anàlogament, si $x > a$, aleshores $F(x) = 1$, i existeix $n_1 \geq 1$ tal que, per a tot $n \geq n_1$, $F_n(x) = 1$. Es compleix, doncs, $\lim_{n\to\infty} F_n(x) = F(x)$ en tot punt x on la funció F és contínua.

Aquest exemple correspon a una situació *determinista*.

Exemple 4.10 En la subsecció 2.4.6 vàrem introduir la distribució de Poisson i vàrem demostrar el resultat següent:

Considerem una successió de variables aleatòries $\{X_n, n \geq 1\}$ amb distribució $B(n, p_n)$, on $\{p_n, n \geq 1\}$ és una successió en $[0,1]$ tal que $\lim_{n\to\infty} np_n = \lambda$, per a un $\lambda > 0$. Aleshores,

$$\lim_{n\to\infty} P\{X_n = k\} = \frac{\lambda^k}{k!} \exp(-\lambda).$$

D'acord amb la proposició 4.5 aquest resultat significa que una variable aleatòria amb llei Poiss(λ), $\lambda > 0$, s'obté com el límit en llei d'una successió de distribucions binomials com les descrites anteriorment.

Exemple 4.11 Considerem una successió de variables aleatòries $\{X_n, n \geq 1\}$ amb les característiques següents: cadascuna de les X_n pren els seus valors en el conjunt $\left\{0, \frac{1}{n}, \frac{2}{n}, \ldots \frac{n-1}{n}\right\}$ i cadascun d'aquests valors és pres amb probabilitat $\frac{1}{n}$. Aleshores, la successió $\{X_n, n \geq 1\}$ convergeix en llei cap a una variable X que té distribució uniforme en $[0,1]$.

En efecte, fixem $x \in [0,1]$. Hi ha $[nx]+1$ punts de la forma $\frac{k}{n}$ $(k = 0,1,\ldots)$ que compleixen $\frac{k}{n} \leq x$. En conseqüència,

$$F_n(x) = P\{X_n \leq x\} = \frac{1}{n}([nx]+1).$$

Resulta, doncs,

$$\lim_{n\to\infty} F_n(x) = \lim_{n\to\infty} \frac{[nx]+1}{n} = x.$$

Per a $x < 0$, $F_n(x) = 0$ i per a $x \geq 1$, $F_n(x) = 1$. Queda així justificada la convergència cap a la llei uniforme.

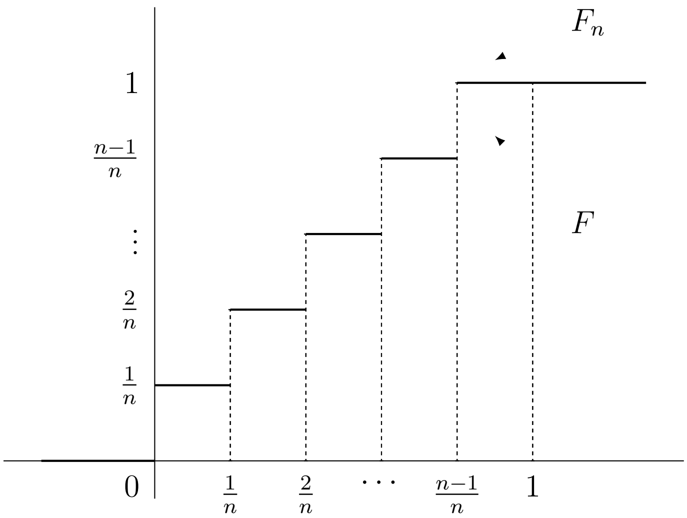

Figura 4.3: Exemple 4.11

Exemple 4.12 Sigui $\{X_n, n \geq 1\}$ una successió de variables aleatòries absolutament contínues amb funcions de densitat, $f_n, n \geq 1$ donades per

$$f_n(x) = n\frac{x^{n-1}}{\theta^n}\mathbb{1}_{(0,\theta)}(x),$$

amb $\theta > 0$. Definim

$$Z_n = n(\theta - X_n), n \geq 1.$$

La successió $Z_n = n(\theta - X_n), n \geq 1$ convergeix en distribució cap a una variable X aleatòria amb distribució exponencial de paràmetre $\frac{1}{\theta}$.

En efecte, en $z \leq 0$, les funcions de distribució de les variables aleatòries Z_n i la de X valen totes 0. Fixem $z > 0$; existeix $n_0 \in \mathbb{N}$ i, per a tot $n \geq n_0$, $\frac{z}{n} \leq \theta$. En conseqüència, si $n \geq n_0$,

$$F_{Z_n}(z) = P\{Z_n \leq z\} = 1 - \int_0^{\theta - \frac{z}{n}} n \frac{x^{n-1}}{\theta^n} \, \mathrm{d}x = 1 - \Big(1 - \frac{z}{n\theta}\Big)^n.$$

Per tant, per a $z > 0$

$$\lim_{n\to\infty} F_{Z_n}(z) = 1 - \exp\left(-\frac{z}{\theta}\right) = F_X(z).$$

Com que $F_X(z)$ és contínua en tot punt, això acaba la prova.

Acabem aquesta secció donant una definició equivalent de la convergència en llei. A més del seu interès intrínsec, aquest resultat serà útil a l'hora d'estudiar la convergència en llei de transformacions, com la suma o el producte, de successions de variables aleatòries, que abordarem en la secció següent.

Denotarem per $\mathcal{C}_b$ el conjunt de les funcions reals contínues i fitades.

Teorema 4.1 *Una successió $\{X_n, n \geq 1\}$ de variables aleatòries convergeix en llei cap a una variable aleatòria X si, i únicament si, per a tota funció $f \in \mathcal{C}_b$*

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

Demostració: Suposem primer que

$$\mathcal{L} - \lim_{n\to\infty} X_n = X,$$

és a dir, la convergència de les funcions de distribució F_n de X_n cap a F, la funció de distribució de X, en els punts de continuïtat del límit. Fixem $\varepsilon > 0$ i k real i positiu tal que $F(k) - F(-k) > 1 - \varepsilon$ i tant k com $-k$ són punts de continuïtat de F.

Sigui $f \in \mathcal{C}_b$. Com que f és uniformement contínua en $[-k, k]$, existeix una funció esglaonada $g = \sum_{j=1}^{r} b_j \mathbb{1}_{(a_j, a_{j+1}]}$ amb $b_j \in \mathbb{R}$, els punts a_j són punts de continuïtat de F, $j = 1, \dots, r+1$, i $\sup_{x \in [-k,k]} |f(x) - g(x)| < \varepsilon$.
Es compleix que

$$|E\left(f(X_n) - f(X)\right)| \le A_n + B_n,$$

amb

$$\begin{aligned} A_n &= \left|E\left(f(X_n)\mathbb{1}_{\{|X_n|>k\}}\right) - E\left(f(X)\mathbb{1}_{\{|X|>k\}}\right)\right|, \\ B_n &= \left|E\left(f(X_n)\mathbb{1}_{\{|X_n|\le k\}}\right) - E\left(f(X)\mathbb{1}_{\{|X|\le k\}}\right)\right|. \end{aligned}$$

Com que f és fitada,

$$A_n \le |f|_\infty \left(P\{|X_n| > k\} + P\{|X| > k\}\right).$$

Per hipòtesi,

$$P\{|X| > k\} \le 1 - P\{-k < X \le k\} < \varepsilon$$

i, en conseqüència,

$$P\{|X_n| > k\} < \varepsilon.$$

Així, doncs,

$$A_n \le 2\varepsilon |f|_\infty.$$

Descomponem el terme B_n de la manera següent:

$$B_n = \sum_{i=1}^{3} B_n^i,$$

amb

$$\begin{aligned} B_n^1 &= \left|E\left(f(X_n) - g(X_n)\right)\mathbb{1}_{\{|X_n|\le k\}}\right|, \\ B_n^2 &= \left|E\left(g(X_n)\mathbb{1}_{\{|X_n|\le k\}}\right) - E\left(g(X)\mathbb{1}_{\{|X|\le k\}}\right)\right|, \\ B_n^3 &= \left|E\left(\left(g(X) - f(X)\right)\mathbb{1}_{\{|X|\le k\}}\right)\right|. \end{aligned}$$

Clarament,

$$B_n^1 + B_n^3 \leq 2\varepsilon.$$

D'altra banda,

$$B_n^2 = \left| \sum_{j=1}^{r} b_j \left((F_n(a_{j+1}) - F_n(a_j)) - (F(a_{j+1}) - F(a_j)) \right) \right|,$$

que tendeix a zero quan n tendeix a infinit, per hipòtesi. Com que ε és arbitrari, resulta $\lim_{n\to\infty} E\left(f(X_n)\right) = E\left(f(X)\right)$.

Recíprocament, fixem un punt x de continuïtat de la funció F i considerem les funcions contínues i fitades definides per

$$\begin{aligned} f_\varepsilon^+(y) &= \mathbb{1}_{(-\infty,x]}(y) + \left(1 - \frac{y-x}{\varepsilon}\right) \mathbb{1}_{(x,x+\varepsilon)}(y), \\ f_\varepsilon^-(y) &= \mathbb{1}_{(-\infty,x-\varepsilon]}(y) + \frac{x-y}{\varepsilon} \mathbb{1}_{(x-\varepsilon,x)}(y), \end{aligned}$$

on ε és un nombre real estrictament positiu que fixem.

Posem $\ell = \liminf_{n\to\infty} F_n(x)$, $L = \limsup_{n\to\infty} F_n(x)$. Tenim que

$$\begin{aligned} F(x-\varepsilon) &\leq E\left(f_\varepsilon^-(X)\right) = \lim_n E\left(f_\varepsilon^-(X_n)\right) \\ &\leq \liminf_n F_n(x) = \ell \leq L = \limsup_n F_n(x) \\ &\leq \lim_n E\left(f_\varepsilon^+(X)\right) \leq F(x+\varepsilon). \end{aligned}$$

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

Observació 4.1 *De la demostració de la segona part del teorema anterior es dedueix que, per a comprovar la convergència* $\mathcal{L} - \lim_{n\to\infty} X_n = X$, *és suficient veure que* $\lim_{n\to\infty} E\left(f(X_n)\right) = E\left(f(X)\right)$ *per a tota funció* f uniformement contínua *i fitada, ja que les funcions* f_ε^+ *i* f_ε^- *són d'aquest tipus. Endemés, regularitzant les funcions* f_ε^+ *i* f_ε^- *en els punts* x, $x+\varepsilon$ *i* $x-\varepsilon$, x, *respectivament, pot demostrar-se, de manera similar, la segona part del teorema anterior suposant* $\lim_{n\to\infty} E(f(X_n)) = E(f(X))$, *per a tota funció*

$f \in \mathcal{C}_b^\infty$ (funcions infinitament diferenciables, fitades i amb derivades fitades).

4.6 Relacions entre els diferents tipus de convergències i aplicacions

L'objectiu d'aquest apartat és d'estudiar les possibles relacions entre les diverses nocions de convergència de successions de variables aleatòries que hem introduït en les seccions 4.2 a 4.5. En primer lloc, estudiarem la connexió entre la convergència quasi segura i en probabilitat; després, la relació entre la convergència en mitjana d'ordre p i en probabilitat. Finalment, veurem que la convergència en llei és la noció menys restrictiva de totes les que hem estudiat. Al llarg de la secció presentarem exemples dels tipus que, en general, no es relacionen.

Teorema 4.2 *Sigui $\{X_n, n \geq 1\}$ una successió de variables aleatòries i X una altra variable aleatòria.*

1. *Si $X_n \longrightarrow X$, q.s. quan n tendeix a infinit, aleshores $P - \lim_{n\to\infty} X_n = X$.*
2. *Si $P - \lim_{n\to\infty} X_n = X$, existeix una subsuccessió $\{X_{n_k}, k \geq 1\}$ que convergeix quasi segurament cap a la variable X, quan k tendeix a infinit.*

Demostració: Comencem per l'apartat (1). Fixem $\varepsilon > 0$. Per la proposició 4.1,

$$P(\liminf_n \{|X_n - X| \leq \varepsilon\}) = 1.$$

Aleshores, aplicant el lema 4.1, obtenim que

$$\limsup_n P\{|X_n - X| > \varepsilon\} \leq P\left(\limsup_n \{|X_n - X| > \varepsilon\}\right) = 0,$$

i, en conseqüència, $P - \lim_{n\to\infty} X_n = X$.

Demostrem ara (2). Suposem que $P - \lim_{n\to\infty} X_n = X$. Es compleix la condició de Cauchy, $\lim_{m,n\to\infty} P\{|X_n - X_m| > \varepsilon\} = 0$, per a tot $\varepsilon > 0$. Considerem dues sèries de termes positius $\sum_{n=1}^{\infty} \varepsilon_n < \infty$ i $\sum_{n=1}^{\infty} \delta_n < \infty$.

Definirem una successió de nombres naturals $\{n_k, k \geq 1\}$ estrictament creixent de nombres naturals de la manera següent. Posem $n_0 = 0$ i per a tot $k \geq 1$ sigui n_k el primer natural més gran que n_{k-1} que satisfà la condició

$$P\{|X_n - X_m| > \varepsilon_k\} < \delta_k,$$

per a tot $n, m \geq n_k$.

Aleshores,

$$\sum_{k=1}^{\infty} P\{|X_{n_{k+1}} - X_{n_k}| > \varepsilon_k\} < \sum_{k=1}^{\infty} \delta_k < \infty.$$

Utilitzant el criteri demostrat en la proposició 4.2, resulta que la successió parcial $\{X_{n_k}, k \geq 1\}$ convergeix quasi segurament. El límit haurà de ser necessàriament la variable X, ja que la convergència quasi segura implica la convergència en probabilitat (tal com hem demostrat en l'apartat (1)) i sabem que $P - \lim_{k\to\infty} X_{n_k} = X$.

La proposició queda totalment demostrada. $\square$

Observeu que en la demostració de l'apartat (2) del teorema anterior hem utilitzat únicament que la successió $\{X_n, n \geq 1\}$ és de Cauchy en probabilitat. En la proposició 4.7 demostrem que aquesta condició és equivalent a la convergència en probabilitat.

Considerem l'exemple 4.6. És fàcil veure que la successió de variables aleatòries que allí proposàvem no convergeix quasi segurament cap a zero. De fet, per a tot x fix de $[0, 1]$ la successió $\{X_{n,m}(x), n \geq 1, 1 \leq m \leq n\}$ conté infinits zeros i infinits uns. Resulta, doncs, que la convergència en probabilitat no implica la convergència quasi segura.

Ara bé, si la successió de variables aleatòries és *monòtona* (creixent o decreixent), aleshores la convergència en probabilitat implica la quasi segura. En efecte, suposem, per exemple, que la successió $\{X_n, n \geq 1\}$ convergeix en probabilitat cap a X i és creixent quasi segurament, és a dir, $X_n(\omega) \leq X_{n+1}(\omega)$, per a tot $\omega \in C$ tal que $P(C) = 1$. Fixat $\varepsilon > 0$, els conjunts $\{|X_n - X| > \varepsilon\}$ decreixen amb n. En conseqüència,

$$\limsup_n\{|X_n - X| > \varepsilon\} = \bigcap_n\{|X_n - X| > \varepsilon\}$$

i, per la propietat de continuïtat de la probabilitat,

$$P\{\limsup_n\{|X_n - X| > \varepsilon\} = \lim_n P\{|X_n - X| > \varepsilon\} = 0.$$

La proposició 4.1 ens assegura que $\{X_n, n \geq 1\}$ convergeix quasi segurament cap a X.
Per a successions decreixents l'argument és similar.

Donem ara una conseqüència del teorema 4.2 que completa l'estudi de la metritzabilitat de la convergència en probabilitat.

Proposició 4.7 *L'espai $L^0(\Omega, \mathcal{A}, P)$, amb la mètrica definida a partir de la convergència en probabilitat, és un espai mètric complet.*

Demostració: Sigui $\{X_n, n \geq 1\}$ una successió de Cauchy en probabilitat. En la demostració del teorema 4.2 hem vist que aquesta condició implica l'existència d'una subsuccessió $\{X_{n_k}, k \geq 1\}$ que convergeix cap a un cert límit X, en el sentit de la convergència quasi segura. Considerem la descomposició

$$P\{|X_n - X| > \varepsilon\} \leq P\{X_n - X_{n_k}| \geq \frac{\varepsilon}{2}\} + P\{|X_{n_k} - X| \geq \frac{\varepsilon}{2}\}.$$

Com que $\{X_n, n \geq 1\}$ és una successió de Cauchy en probabilitat, existeix un nombre natural n_0 tal que, si $n, n_k > n_0$, es té $P\{X_n - X_{n_k}| \geq \frac{\varepsilon}{2}\} \leq \frac{\varepsilon}{2}$. D'altra banda, com que $\{X_{n_k}, k \geq 1\}$ convergeix en probabilitat cap a X, existeix $n_1 \in \mathbb{N}$ tal que per tot $n_k \geq n_1$ es compleix $P\{|X_{n_k} - X| \geq \frac{\varepsilon}{2}\} \leq \frac{\varepsilon}{2}$. Aleshores, prenent $N_0 = \max(n_0, n_1)$ tindrem que

$$P\{|X_n - X| > \varepsilon\} \leq \varepsilon, n \geq N_0.$$

□

Tot seguit estudiem la relació entre convergència en probabilitat i convergència en mitjana d'ordre p.

Proposició 4.8 *Siguin* $\{X_n, n \geq 1\}$ *i* X *variables aleatòries de* $L^p(\Omega, \mathcal{A}, P)$. *Si* $L^p - \lim_{n\to\infty} X_n = X$, *també* $P - \lim_{n\to\infty} X_n = X$.

Demostració: És una conseqüència senzilla de la desigualtat de Txebixev aplicada a la funció $f(x) = |x|^p$ i la variable aleatòria $X_n - X$. En efecte, fixem $\varepsilon > 0$. Aleshores,

$$\lim_{n\to\infty} P\{|X_n - X| > \varepsilon\} \leq \frac{1}{\varepsilon^p} \lim_{n\to\infty} E\left(|X_n - X|^p\right) = 0,$$

que és el que volíem provar. $\square$

El recíproc de la proposició anterior no és cert. En altres paraules, existeixen successions de variables aleatòries que convergeixen en probabilitat i no convergeixen en mitjana d'ordre p. Tampoc les convergències quasi segura i en mitjana d'ordre p estan relacionades si no imposem condicions addicionals. Exemples estudiats anteriorment confirmen el que acabem de dir.

En l'exemple 4.7 hem vist que hi ha convergència en mitjana d'ordre p, per a tot $p \geq 1$; en canvi no hi ha convergència quasi segura. En efecte, fixat un element $x \in [0, 1]$ existeixen infinits intervals del tipus $[q2^{-p}, (q + 1)2^{-p}]$ que el contenen i també infinits que no el contenen. És a dir, $\limsup_n X_n(x) = 1$ i $\liminf_n X_n(x) = 0$. Per tant, la successió X_n no convergeix en cap punt.

Considerem novament el mateix espai de probabilitat que en l'exemple 4.4. La successió $\{X_n,\ n \geq 1\}$ definida per $X_n = 2^n\ \mathbb{1}_{[0,\frac{1}{n}]}$ convergeix quasi segurament cap a zero; en canvi,

$$E(|X_n|^p) = \frac{2^{np}}{n},$$

que tendeix a infinit quan n tendeix a infinit, qualsevol que sigui $p \geq 1$.

Aquest exemple demostra que la convergència quasi segura (o en probabilitat) no implica la convergència en mitjana d'ordre p.

Amb una condició de fitació uniforme de la successió per una variable aleatòria de $L^p(\Omega, \mathcal{A}, P)$ pot provar-se que la convergència quasi segura implica la convergència en mitjana d'ordre p. No és, però, la nostra intenció entrar en el detall d'aquests resultats.

Finalment, ens ocupem de la convergència en llei.

Proposició 4.9 *Sigui $\{X_n, n \geq 1\}$ una successió de variables aleatòries que convergeix en probabilitat cap a una variable aleatòria X. Aleshores, la successió convergeix també en llei cap a X.*

L'afirmació recíproca és també certa si la variable aleatòria X és constant quasi segurament, és a dir, $X(\omega) = k$, per a tot $\omega \notin N$, on N és un conjunt de $\mathcal{A}$ de probabilitat zero, i $k \in \mathbb{R}$.

Demostració: Per a tot $\varepsilon > 0$ i $x \in \mathbb{R}$ es tenen les desigualtats següents:

$$\begin{aligned} P\{X_n \leq x\} &\leq P\{|X_n - X| \geq \varepsilon\} + P\{X \leq x + \varepsilon\}, \\ P\{X \leq x - \varepsilon\} &\leq P\{|X_n - X| \geq \varepsilon\} + P\{X_n \leq x\}. \end{aligned} \tag{4.5}$$

En efecte, per a demostrar la primera considerem les inclusions

$$\begin{aligned} \{X_n \leq x\} &\subset \{X_n \leq x\} \cap \{X > x + \varepsilon\} \cup \{X \leq x + \varepsilon\} \\ &\subset \{|X_n - X| \geq \varepsilon\} \cup \{X \leq x + \varepsilon\}. \end{aligned}$$

La segona es demostra de manera anàloga.

Aleshores, utilitzant primer la segona desigualtat de (4.5) i després la primera,

$$F(x - \varepsilon) - P\{|X_n - X| \geq \varepsilon\} \leq F_n(x) \leq P\{|X_n - X| \geq \varepsilon\} + F(x + \varepsilon).$$

Sigui x un punt on la funció de distribució de X, F, és contínua. Per a tot $\delta > 0$ podem trobar $\varepsilon > 0$ tal que

$$F(x + \varepsilon) \leq F(x) + \frac{\delta}{2},$$

i

$$F(x - \varepsilon) > F(x) - \frac{\delta}{2}.$$

També podem determinar $n_0 \geq 1$ tal que per a tot $n \geq n_0$,

$$P\{|X_n - X| \geq \varepsilon\} \leq \frac{\delta}{2},$$

ja que $\{X_n, n \geq 1\}$ convergeix en probabilitat cap a X.
Tenim, doncs, que

$$\begin{aligned}
F(x) - \delta &\leq F(x-\varepsilon) - \frac{\delta}{2} \\
&\leq F(x-\varepsilon) - P\{|X_n - X| \geq \varepsilon\} \leq F_n(x) \\
&\leq P\{|X_n - X| \geq \varepsilon\} + F(x+\varepsilon) \leq \delta + F(x).
\end{aligned}$$

En conseqüència,

$$\lim_{n\to\infty} F_n(x) = F(x),$$

per a tot x on la funció F és contínua. Per tant, queda establerta la convergència en llei.

Passem ara a demostrar la segona part de la proposició. Per a tot $\varepsilon > 0$ els punts $k+\varepsilon$ i $k-\varepsilon$ són punts de continuïtat de la funció de distribució de X. Tindrem, doncs, que

$$\begin{aligned}
P\{|X_n - k| > \varepsilon\} &= P\{X_n > k+\varepsilon\} + P\{X_n < k-\varepsilon\} \\
&= 1 - F_n(k+\varepsilon) + F_n\left((k-\varepsilon)-\right).
\end{aligned}$$

Fent $n \to \infty$ s'obté que

$$\lim_{n\to\infty} P\{|X_n - k| \geq \varepsilon\} \leq 1 - F(k+\varepsilon) + F(k-\varepsilon) = 0.$$

Això acaba la demostració de la proposició. $\square$

Exemple 4.13 Sigui $\{X_n, n \geq 1\}$ una successió de variables aleatòries, X_n amb distribució uniforme en l'interval $[-\frac{1}{n}, \frac{1}{n}]$, $n \geq 1$. A partir de les gràfiques de les funcions de distribució s'intueix que el límit en distribució d'aquesta successió és la variable aleatòria $X = 0$.

Per a demostrar-ho calculem primer la funció de distribució F_n de cada X_n, $n \geq 1$. Es té que

$$F_n(x) = \begin{cases} 0, & \text{si } x \leq -1/n, \\ \frac{n}{2}(x + \frac{1}{n}), & \text{si } x \in [-1/n, 1/n], \\ 1, & \text{si } x \geq 1/n. \end{cases}$$

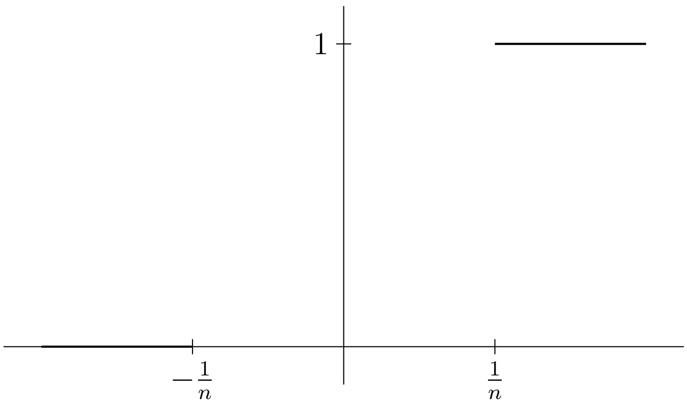

Figura 4.4: Exemple 4.13

La funció de distribució de la variable aleatòria $X = 0$ és $F(x) = \mathbb{1}_{[0,\infty)}$. Fixem $x \neq 0$. Podem trobar $n_0 \in \mathbb{N}$ tal que $x \notin [-\frac{1}{n}, \frac{1}{n}]$, per a tot $n \geq n_0$. Per tant, si $x < 0$, $F_n(x) = F(x) = 0$ i si $x > 0$, $F_n(x) = F(x) = 1$, per a tot $n \geq n_0$. En definitiva,

$$\lim_{n\to\infty} F_n(x) = F(x), x \neq 0.$$

Com que l'únic punt de discontinuïtat de la funció de distribució F és $x = 0$, queda establerta la convergència en distribució. Aquesta convergència també té lloc en probabilitat, ja que el límit en distribució és una variable aleatòria constant.

Considerem ara la successió de variables aleatòries $Y_n = nX_n$, $n \geq 1$. Es tracta d'una successió de variables amb la mateixa distribució, una llei uniforme en l'interval $[-1, 1]$; per tant, òbviament convergirà en llei cap a una variable aleatòria Y amb la mateixa distribució que la de les variables que formen la successió. El límit ara no és, com en el cas anterior, una variable aleatòria constant i, en general, no hi haurà convergència en probabilitat.

A continuació presentem el *teorema de representació de Skorohod.* Cal fer notar que no es tracta d'una relació entre les convergències quasi segura i en distribució, sinó d'una representació de la segona en termes de la primera.

Sigui F una funció de distribució. Definim la *funció quantil* o *pseudoinversa* de F per

$$Q(y) = \inf\{x : F(x) \geq y\},$$

$0 < y < 1$.

Si F és una funció contínua i bijectiva, ço és, estrictament creixent, aleshores $Q = F^{-1}$ i, clarament,

$$y \leq F(x) \quad \text{si, i només si} \quad Q(y) \leq x. \tag{4.6}$$

És fàcil veure que la demostració de (4.6) només requereix de la continuïtat per la dreta de la funció F i, per tant, és certa per a tota funció de distribució.

Sigui U una variable aleatòria amb distribució uniforme en $(0, 1)$. Aleshores, $Q(U)$ és també una variable aleatòria i la seva funció de distribució és F. En efecte, per (4.6),

$$P\{Q(U) \leq x\} = P\{F(x) \geq U\} = F(x).$$

Utilizant aquest resultat podem donar la demostració del teorema de representació de Skorohod en $\mathbb{R}$.

Teorema 4.3 *Sigui $\{Y_n, n \geq 1\}$ una successió de variables aleatòries que convergeixen en llei cap a una variable aleatòria Y. Existeixen una successió de variables aleatòries $\{X_n, n \geq 1\}$ i una variable aleatòria X que compleixen les propietats següents:*

1. *La llei de Y_n és la mateixa que la de X_n, per a tot $n \geq 1$ i la llei de Y la mateixa que la de X.*
2. *La successió $\{X_n, n \geq 1\}$ convergeix quasi segurament cap a la variable aleatòria X.*

Demostració: Denotem per F_n, F, les funcions de distribució de Y_n, Y, respectivament, i per Q_n la funció quantil de F_n, $n \geq 1$. Sigui U una variable aleatòria amb distribució uniforme en $(0,1)$. Definim $X_n = Q_n(U)$. Demostrem tot seguit que $\{Q_n(u), n \geq 1\}$ convergeix per a tot u, llevat en un subconjunt numerable de punts de $(0,1)$. Aleshores, es pot definir X (llevat en un conjunt de probabilitat zero), com el límit de la successió $\{X_n, n \geq 1\}$.

Sigui T_0 el conjunt de punts de continuïtat de la funció de distribució F, que és dens en $[0,1]$. Si per a algun $u \in (0,1)$, la successió $\{Q_n(u), n \geq 1\}$ no convergís, existirien punts $t, s \in T_0$ tals que

$$\liminf_n Q_n(u) < t < s < \limsup Q_n(u).$$

Per (4.6), això implica $F_n(t) \geq u$ i $F_n(s) \leq u$. Aleshores, prenent límits per a alguna subsuccessió, obtenim $F(t) \geq u$ i $F(s) \leq u$, la qual cosa implica que F pren el valor constant igual a u en l'interval $[t,s]$. Per les propietats de les funcions de distribució això només pot passar per a un conjunt de valors de u que, com a màxim, serà numerable. Per tant, $\{Q_n(u), n \geq 1\}$ convergeix per a tot u, llevat en un subconjunt numerable de punts de $(0,1)$, que és el que volíem provar.

Vegem finalment que el límit puntual de la successió $\{X_n, n \geq 1\}$ no pot ser infinit. Donat $u \in (0,1)$, sigui $t \in T_0$ tal que $F(t) > u$. La convergència en els punts de T_0 implica $F_n(t) > u$ per a tot n prou gran, o sigui, $Q_n(u) \leq t$ per a tot n prou gran. En conseqüència, $\limsup Q_n(u) < \infty$. Un argument similar demostraria $\liminf Q_n(u) > -\infty$. □

Proposició 4.10 (Teorema de Slutsky) *Siguin* $\{X_n, n \geq 1\}$ *i* $\{Y_n, n \geq 1\}$ *successions de variables aleatòries tals que* $\mathcal{L} - \lim_{n\to\infty} X_n = X$, $\mathcal{L} - \lim_{n\to\infty} Y_n = c$, *on* X *és una variable aleatòria i* $c \in \mathbb{R}$. *Aleshores,*

1. $\mathcal{L} - \lim_{n\to\infty}(X_n + Y_n) = X + c$,
2. $\mathcal{L} - \lim_{n\to\infty}(X_n Y_n) = cX$.

Demostració: Pel teorema 4.1 i l'observació 4.1 és suficient provar que, per a tota funció f uniformement contínua i fitada, es compleix que

$$\lim_{n\to\infty} E\big(f(X_n + Y_n)\big) = E\big(f(X + c)\big), \tag{4.7}$$

$$\lim_{n\to\infty} E\big(f(X_nY_n)\big) = E\big(f(cX)\big), \tag{4.8}$$

respectivament.

Fixem, doncs, f que compleixi aquestes propietats. Aleshores, novament pel teorema 4.1,

$$\lim_{n\to\infty} E\big(f(X_n + c)\big) = E\big(f(X + c)\big), \tag{4.9}$$

$$\lim_{n\to\infty} E\big(f(cX_n)\big) = E\big(f(cX)\big), \tag{4.10}$$

ja que les funcions $h_1(x) = f(x + c)$, $h_2(x) = f(cx)$ són també contínues i fitades.

Observem que, per la proposició 4.9, la successió $\{Y_n, n \geq 1\}$ convergeix en probabilitat cap a c.

Comencem donant la demostració de (4.7):

Donat $\varepsilon > 0$, sigui $\delta > 0$ tal que per a tots $x, y \in \mathbb{R}$, $|x - y| < \delta$, es té $|f(x) - f(y)| < \varepsilon$. Aleshores,

$$\begin{aligned}
&\big|E\big(f(X_n + Y_n)\big) - E\big(f(X + c)\big)\big| \\
&\quad \leq E\big(\big|f(X_n + Y_n) - f(X_n + c)\big|\mathbb{1}_{\{|Y_n - c| > \delta\}}\big) \\
&\quad + E\big(\big|f(X_n + Y_n) - f(X_n + c)\big|\mathbb{1}_{\{|Y_n - c| \leq \delta\}}\big) \\
&\quad + |E\big(f(X_n + c)\big) - E\big(f(X + c)\big)\big| \\
&\quad \leq 2\|f\|_\infty P\{|Y_n - c| > \delta\} + \varepsilon + |E\big(f(X_n + c)\big) - E\big(f(X + c)\big)\big|.
\end{aligned}$$

Utilitzant ara (4.9) i la convergència en probabilitat de $\{Y_n, n \geq 1\}$ cap a 0,

$$0 \leq \lim_{n\to\infty} |E\big(f(X_n + Y_n)\big) - E\big(f(X + c)\big)\big| \leq \varepsilon$$

i, com que ε és arbitrari, això completa la prova de (4.7).

Demostrem ara (4.8). Sigui F la funció de distribució de X i F_n la de X_n. Fixem $\varepsilon > 0$ i k real i positiu tal que $F(k) - F(-k) > 1 - \varepsilon$ i tant k com $-k$ són punts de continuïtat de F. Aleshores, també $F_n(k) - F_n(-k) > 1 - \varepsilon$. Com per a la demostració de (4.7), fixem f uniformement contínua, fitada i $\delta > 0$ tal que per a tots els $x, y \in \mathbb{R}$, $|x - y| < \delta$, es té $|f(x) - f(y)| < \varepsilon$.

Es compleix

$$
\begin{aligned}
&\big|E\big(f(X_nY_n)-f(cX)\big)\big| \\
&\quad\le E\Big(\big|f(X_nY_n)-f(cX_n)\big|\mathbb{1}_{\{|Y_n-c|>\frac{\delta}{k}\}}\Big) \\
&\quad+E\Big(\big|f(X_nY_n)-f(cX_n)\big|\mathbb{1}_{\{|Y_n-c|\le\frac{\delta}{k},|X_n|>k\}}\Big) \\
&\quad+E\Big(\big|f(X_nY_n)-f(cX_n)\big|\mathbb{1}_{\{|Y_n-c|\le\frac{\delta}{k},|X_n|\le k\}}\Big) \\
&\quad+\big|E\big(f(cX_n)\big)-E\big(f(cX)\big)\big| \\
&\quad\le 2\|f\|_\infty\Big(P\{|Y_n-c|>\delta/k\}+P\{|X_n|>k\}\Big) \\
&\quad+\varepsilon+\big|E\big(f(cX_n)\big)-E\big(f(cX)\big)\big| \\
&\quad\le 2\|f\|_\infty\Big(P\{|Y_n-c|>\delta/k\}+\varepsilon\Big)+\varepsilon+\big|E\big(f(cX_n)\big)-E\big(f(cX)\big)\big|.
\end{aligned}
$$

Aquestes fitacions, juntament amb (4.10) acaben la demostració de (4.8). □

Exercicis

4.1 *Sigui $\{X_n, n\ge 1\}$ una successió de variables aleatòries independents amb distribució uniforme en l'interval $[0,1]$. Demostreu que la successió de variables aleatòries $\{m_n, n\ge 1\}$ definida per*

$$m_n=\min(X_1,\cdots,X_n)$$

convergeix en probabilitat i quasi segurament cap a la variable aleatòria $X=0$.

4.2 *Considerem una successió de variables aleatòries $\{X_n, n\ge 1\}$ tals que*

$$P(X_n=0)=1-\alpha_n,\quad P(X_n=n)=\alpha_n,$$

on $\{\alpha_n, n\ge 1\}$ és una successió en $[0,1]$.
Demostreu que $\{X_n, n\ge 1\}$ convergeix en llei si, i només si, $\lim_{n\to\infty}\alpha_n=0$ i, en aquest cas, calculeu-ne el límit.

4.3 *Sigui X una variable aleatòria simètrica, ço és, X i $-X$ tenen la mateixa llei (per exemple X amb distribució $\mathrm{N}(0,1)$). Definim $X_n=(-1)^nX$, $n\ge 1$. Estudieu la convergència en llei i en probabilitat de la successió $\{X_n, n\ge 1\}$.*

4.4 *Sigui* $\{X_n, n \geq 1\}$ *una successió de variables aleatòries independents, idènticament distribuïdes tals que*

$$P\{X_n = 1\} = P\{X_n = -1\} = \frac{1}{2}.$$

Definim

$$Z_n = \sum_{j=1}^{n} \frac{X_j}{2^j}, \quad n \geq 1, \quad Z = \sum_{j=1}^{\infty} \frac{X_j}{2^j}.$$

Demostreu que Z *està ben definida, és a dir, que la sèrie* $\sum_{j=1}^{\infty} \frac{X_j}{2^j}$ *convergeix quasi segurament. Demostreu també que la successió* $\{Z_n, n \geq 1\}$ *convergeix cap a* Z *en* $L^1(\Omega, \mathcal{A}, P)$ *(mitjana d'ordre* 1*).*

4.5 *Sigui* U *una variable aleatòria amb llei uniforme en* $[0, 1]$*. Per a tot* $n \geq 1$*, definim*

$$X_n = \begin{cases} 1, & si\ U > \frac{1}{n}, \\ 0, & altrament. \end{cases}$$

Estudieu la convergència de la successió $\{X_n, n \geq 1\}$ *en mitjana d'ordre* p*, en probabilitat, en llei i quasi segurament.*

4.6 *Sigui* f *una funció de densitat en* $[0, \infty)$*, contínua, tal que*

$$\lambda := \lim_{x \to \infty} f(x) > 0.$$

1. *Demostreu* $\lim_{n \to \infty} n \int_0^{\frac{x}{n}} f(y)\, \mathrm{d}y = \lambda x, x \geq 0.$
2. *Sigui* $\{Z_n, n \geq 1\}$ *una successió de variables aleatòries independents, idènticament distribuïdes, amb densitat* f*. Demostreu que la successió* $\{W_n, n \geq 1\}$ *definida per*

   $$W_n = n \min(Z_1, \cdots, Z_n), \quad n \geq 1,$$

   convergeix en distribució cap a una llei exponencial amb paràmetre λ*.*

4.7 *Sigui* $\{X_n, n \geq 1\}$ *una successió de variables aleatòries tal que cada* X_n *té una llei geomètrica amb paràmetre* $\frac{\lambda}{n}$*,* $n \geq 1$*. Considerem la successió* $\{Y_n = \frac{X_n}{n}, n \geq 1\}$*. Estudieu la convergència en llei de la successió* $\{Y_n, n \geq 1\}$ *i determineu la llei de la variable aleatòria límit.*

4.8 *Sigui* $\{Z_n, n \geq 1\}$ *una successió de variables aleatòries independents, idènticament distribuïdes amb llei uniforme en* $[0,1]$. *Definim*

$$\begin{aligned} Y_n &= n \min(X_1, \cdots, X_n), \\ Z_n &= n[\max(X_1, \cdots, X_n) - 1], \end{aligned}$$

$n \geq 1$.
Estudieu la convergència en llei de les successions $\{Y_n = \frac{X_n}{n}, n \geq 1\}$ *i* $\{Z_n, n \geq 1\}$.
Estudieu la convergència en probabilitat de la successió $\{Z_n, n \geq 1\}$ *cap a la variable aleatòria* $Z = 0$.

4.9 *Sigui* $p \in (0,1)$. *Considerem una successió* $\{X_n, n \geq 1\}$ *de variables aleatòries, independents, idènticament distribuïdes, tal que*

$$P(X_1 = 1) = p, \quad P(X_1 = 0) = 1 - p.$$

Definim $Y_n = X_n X_{n+1}$, $U_n = Y_1 + \cdots + Y_n$, $n \geq 1$.

1. *Determineu la llei de* Y_n, $n \geq 1$.
2. *Descriviu les parelles* n, m *per a les quals les variables aleatòries* Y_n, Y_m *són independents.*
3. *Proveu que, per a tot* $\varepsilon > 0$,

$$\lim_{n \to \infty} P\left(\left|\frac{U_n}{n} - p^2\right| \geq \varepsilon\right) = 0.$$

4.10 *Sigui* $\{X_n, n \geq 1\}$ *una successió de variables aleatòries amb distribució uniforme en l'interval* $[1, 1 + \frac{1}{n}]$.
Estudieu el límit en distribució, probabilitat i quasi segurament d'aquesta successió.
Definim $Y_n = nX_n$, $n \geq 1$. *Hi ha convergència en distribució de la successió* $\{Y_n, n \geq 1\}$ *?*

Capítol 5

Lleis dels grans nombres

Els fenòmens aleatoris tenen un cert caràcter irregular. Ara bé, en una llarga sèrie d'observacions, aquesta irregularitat té una tendència a suavitzar-se. Per exemple, si observem el resultat d'un gran nombre de llançaments d'una moneda, la proporció de vegades que surt cara és molt similar a la de les creus. Quan ens referim a aquesta propietat parlarem de la *llei dels grans nombres*. Aquest fet empíric, que pot ser utilitzat per a la determinació de probabilitats d'esdeveniments, té una justificació matemàtica no trivial, de la qual ens ocuparem al llarg d'aquest capítol.

Considerem una successió $\{X_n, n \geq 1\}$ de variables aleatòries definides en un mateix espai de probabilitat $(\Omega, \mathcal{A}, P)$. Les *lleis dels grans nombres* estudien el comportament asimptòtic de la successió de les mitjanes aritmètiques $\{\frac{S_n}{n}, n \geq 1\}$, on $S_n = X_1 + \cdots + X_n$. S'analitzen dos tipus diferents de convergències: en probabilitat i quasi segura. En el primer cas es parla de *lleis febles* i en el segon, de *lleis fortes*.

Les diverses versions de lleis dels grans nombres actualment al nostre abast difereixen en les hipòtesis de l'enunciat. En aquest capítol se'n troba una mostra representativa. La llei forta de Kolmogorov, establerta en el teorema 5.4, correspon a la versió més elaborada. La seva demostració requereix un cert domini dels tecnicismes de la teoria de la probabilitat. Hem cregut convenient incloure-la a causa de la seva àmplia utilització.

5.1 Lleis febles dels grans nombres

Comencem aquest apartat amb una versió de *llei feble* que estableix la convergència en mitjana quadràtica, i per tant en probabilitat, de la successió de mitjanes aritmètiques d'una successió de variables aleatòries incorrelacionades, tot precisant la velocitat de convergència. Com a cas particular, obtindrem el resultat clàssic de Jakob Bernoulli, publicat en l'*Ars Conjectandi* el 1713 i millorat posteriorment per De Moivre i Laplace. Justificarem la interpretació de l'*histograma de freqüències* en l'anàlisi estadística de dades i, finalment, donarem una demostració probabilista del teorema de Weierstrass sobre aproximació de funcions contínues per polinomis.

Teorema 5.1 *Sigui $\{X_n, n \geq 1\}$ una successió de variables aleatòries incorrelacionades i amb moment de segon ordre fitat per una constant C independent de n. Aleshores,*

$$E\left(\left|\frac{S_n - E(S_n)}{n}\right|^2\right) \leq \frac{C}{n}, \tag{5.1}$$

i, en conseqüència,

$$\lim_{n\to\infty} \frac{S_n - E(S_n)}{n} = 0, \tag{5.2}$$

en el sentit de la convergència en mitjana quadràtica.

Demostració: Com que la successió està formada per variables aleatòries incorrelacionades, es compleix que

$$E\left(\left|\frac{S_n - E(S_n)}{n}\right|^2\right) = \frac{\sum_{i=1}^n E\big(X_i - E(X_i)\big)^2}{n^2}. \tag{5.3}$$

D'altra banda, per hipòtesi,

$$E\big(X_i - E(X_i)\big)^2 = \mathrm{Var}(X_i) \leq C.$$

Per tant, el segon membre de (5.3) està fitat per $\frac{C}{n}$ i això prova (5.1). □

Una situació corrent en problemes d'inferència estadística és la d'una successió de variables aleatòries independents i amb la mateixa distribució (abreujadament i.i.d.). En aquest cas, podem formular el resultat següent.

Corol·lari 5.1 *Sigui* $\{X_n, n \geq 1\}$ *una successió de variables aleatòries i.i.d. amb moment de segon ordre finit. Sigui* m *el valor de* $E(X_1)$*. Es compleix*

$$\lim_{n\to\infty} \frac{S_n}{n} = m, \tag{5.4}$$

en mitjana quadràtica.

Amb la terminologia actual, la llei feble dels grans nombres de J. Bernoulli que hem citat abans s'enuncia d'aquesta manera.

Teorema 5.2 (Llei feble de Bernoulli) *Sigui* $\{X_n, n \geq 1\}$ *una successió de variables aleatòries independents, totes amb la distribució* $B(1,p)$*,* $p \in (0,1)$*. La successió* $\{\frac{S_n}{n}, n \geq 1\}$ *convergeix cap a* p *quan* n *tendeix cap a infinit en el sentit de la convergència en probabilitat.*

Les variables aleatòries del teorema anterior compleixen les hipòtesis del teorema 5.1 amb $C = p(1-p)$. Una aplicació immediata de la desigualtat de Txebixev i la fitació (5.1) permeten d'obtenir que

$$P\left\{\left|\frac{S_n}{n} - p\right| \geq \varepsilon\right\} \leq \frac{p(1-p)}{\varepsilon^2 n}, \tag{5.5}$$

per a tot $\varepsilon > 0$.

Com hem comentat breument en la introducció d'aquest capítol, el teorema 5.2 s'utilitza per a determinar empíricament la probabilitat d'un esdeveniment A. Amb tal finalitat, considerem una successió de repeticions independents d'una experiència aleatòria amb dos resultats possibles, A i A^c, amb probabilitats respectives, p i $1-p$. Fixem n i considerem la variable aleatòria S_n que dóna el nombre de vegades que ha aparegut el resultat A en les n repeticions de l'experiència. La llei de S_n és $B(n,p)$ i es compleix (5.5). Per tant, la freqüència relativa de l'esdeveniment A, $\frac{S_n}{n}$, s'acosta, en el sentit de la convergència en probabilitat (i també en mitjana quadràtica) a la probabilitat de A.

La desigualtat (5.5) serveix també per a la determinació de l'error comès en aproximar p mitjançant $\frac{S_n}{n}$ per a n fix.

En efecte, fixem un marge d'error donat per $\varepsilon > 0$ i una probabilitat α. El segon membre de la desigualtat (5.5) està fitat per $\frac{1}{4n\varepsilon^2}$ i permet de trobar el valor de n perquè

$$P\left\{\left|\frac{S_n}{n} - p\right| \geq \varepsilon\right\} < \alpha.$$

La resposta és que n ha de ser superior o igual a $\frac{1}{4\varepsilon^2\alpha}$.

Una aplicació directa d'aquest teorema justifica el fet, generalment assumit, que les gràfiques conegudes amb el nom d'*histogrames de freqüències* donen informació sobre la funció de densitat de les variables aleatòries absolutament contínues.

Amb aquest objectiu considerem una successió $\{X_n, n \geq 1\}$ de variables aleatòries i.i.d. absolutament contínues amb funció de densitat f. Fixem un interval de la recta real on la funció f no s'anul·la, $[a, b]$, $a \leq b$. Subdividim l'interval $[a, b]$ en subintervals I_k, $k = 1, \ldots, r$, i definim per a cada k

$$Y_k^n = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{I_k}(X_i).$$

Les variables aleatòries $\mathbb{1}_{I_k}(X_i)$, $i \geq 1$ són i.i.d. amb distribució de Bernoulli de paràmetre $p = P\{X_i \in I_k\}$. El teorema 5.2 afirma que $\{Y_k^n, n \geq 1\}$ convergeix en probabilitat (i en mitjana quadràtica) cap a $P\{X_i \in I_k\} = \int_{I_k} f(x)\,\mathrm{d}x$.

Si els intervals I_k són suficientment petits i la funció f és contínua, podrem aproximar $\int_{I_k} f(x)\,\mathrm{d}x$ pel producte del valor de la funció f en l'extrem esquerre de l'interval I_k i la longitud d'aquest interval.

Considerem ara n observacions independents d'una variable aleatòria X amb densitat f, ço és, fixem $\omega \in \Omega$ i sigui $x_i = X_i(\omega)$, $i = 1, \ldots, n$. Posem

$$y_k^n = Y_k^n(\omega) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{I_k}(x_i).$$

La gràfica de la funció que és constant en els intervals I_k i pren els valors y_k^n, $k = 1, \ldots, r$, s'anomena *histograma.*

Els arguments anteriors justifiquen l'aproximació de la funció de densitat f per l'histograma. Això permet de sondejar l'ajustament de dades a una distribució de probabilitat. Per exemple, les dades següents han estat obtingudes a partir de l'observació del temps de durada (en hores) d'un component electrònic d'un sistema d'alarma:

262,8	1,0	36,4	4,0	59,4	35,3	70,5	22,6	3,7	5,8
32,1	0,5	17,4	77,6	46,7	182,4	76,7	3,5	13,4	29,7
6,1	15,1	110,5	45,9	31,7	22,4	27,8	10,0	33,0	26,7
8,0	6,8	63,0	70,9	30,0	12,2	29,6	3,3	32,2	12,3
128,2	24,6	7,0	39,8	71,1	19,4	5,4	4,4	54,4	24,8

La mitjana aritmètica d'aquestes dades és $41,1$. En la figura 5.1 hem dibuixat l'histograma de freqüències basat en 6 intervals de longitud 20 (hores), el qual suggereix una distribució de tipus exponencial.

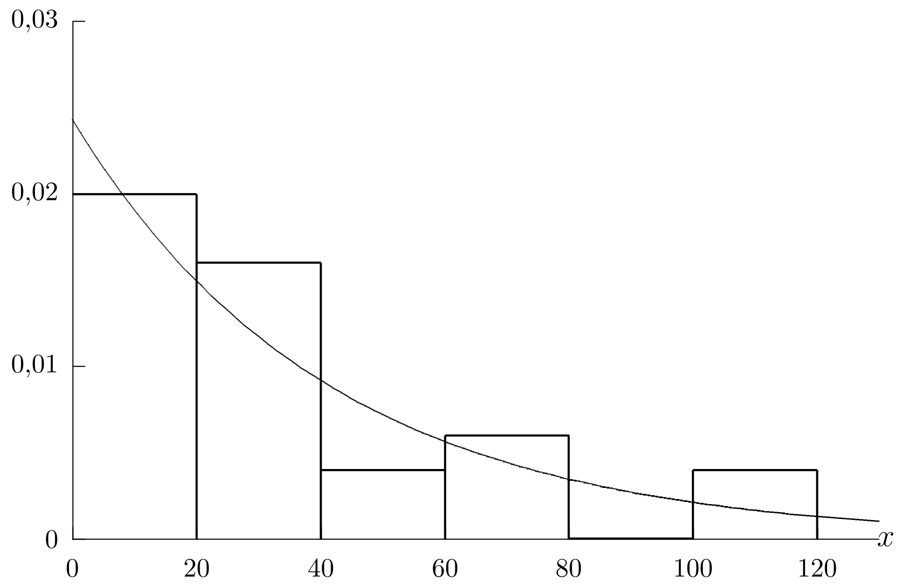

Figura 5.1: Histograma de freqüències i gràfica d'una densitat exponencial amb $\lambda = 41,1$

Exemple 5.1 Com a aplicació interessant de la llei feble dels grans nombres que acabem de veure i, concretament, de la desigualtat (5.5), demostrarem que tota funció contínua $f : [0,1] \to \mathbb{R}$ es pot aproximar uniformement per polinomis. És a dir, donarem una demostració del *teorema de Weierstrass*, proposada per Bernstein, que es basa en tècniques probabilístiques.

Considerem una funció contínua $f : [0,1] \to \mathbb{R}$. Per a cada nombre $x \in [0,1]$ considerem una successió $\{X_n, n \geq 1\}$ de variables aleatòries i.i.d. amb lleis de Bernoulli de paràmetre x, definides en un cert espai de probabilitat $(\Omega, \mathcal{A}, P)$. Sabem que S_n té una llei binomial $B(n,x)$, és a dir,

$$P\{S_n = k\} = \binom{n}{k} x^k (1-x)^{n-k}, \quad k = 0, 1, \ldots, n.$$

Aleshores,

$$E\left[f\Big(\frac{S_n}{n}\Big)\right] = \sum_{k=0}^{n} f\left(\frac{k}{n}\right) \binom{n}{k} x^k (1-x)^{n-k} := p_n(x).$$

Els polinomis $\{p_n(x), n \geq 1\}$ que acabem d'introduir s'anomenen *polinomis de Bernstein* associats a f. Demostrarem que

$$\lim_n \sup_{x \in [0,1]} \big|f(x) - p_n(x)\big| = 0.$$

En efecte, per la propietat de monotonia de l'esperança,

$$\big|f(x) - p_n(x)\big| = \left|f(x) - E\Big[f\Big(\frac{S_n}{n}\Big)\Big]\right| \leq E\left(\left|f(x) - f\Big(\frac{S_n}{n}\Big)\right|\right). \tag{5.6}$$

Com que f és uniformement contínua, fixat $\varepsilon > 0$, existeix $\delta > 0$ tal que, si $|x - y| \leq \delta$, $0 \leq x, y \leq 1$, aleshores $|f(x) - f(y)| < \varepsilon$. Descomponem l'esperança en (5.6) en dues parts i majorem fent servir (5.5):

$$
\begin{aligned}
E\left(\left|f(x) - f\Big(\frac{S_n}{n}\Big)\right|\right) &= E\left(\left|f(x) - f\Big(\frac{S_n}{n}\Big)\right| \mathbb{1}_{\{|\frac{S_n}{n}-x|>\delta\}}\right) \\
&\quad + E\left(\left|f(x) - f\Big(\frac{S_n}{n}\Big)\right| \mathbb{1}_{\{|\frac{S_n}{n}-x|\le\delta\}}\right) \\
&\le 2\|f\|_\infty P\Big\{\Big|\frac{S_n}{n} - x\Big| > \delta\Big\} + \varepsilon \\
&\le 2\|f\|_\infty \frac{1}{\delta^2}\sigma^2\Big(\frac{S_n}{n}\Big) + \varepsilon \\
&= 2\|f\|_\infty \frac{1}{\delta^2}\frac{x(1-x)}{n} + \varepsilon \le \frac{\|f\|_\infty}{2\delta^2 n} + \varepsilon.
\end{aligned}
$$

En conseqüència,

$$
\limsup_{n\to\infty} \sup_{x\in[0,1]} |f(x) - p_n(x)| \le \varepsilon,
$$

i com que ε és arbitrari, obtenim el resultat que buscàvem.

5.2 Lleis fortes dels grans nombres

El primer resultat d'aquest apartat és una millora del teorema 5.1; amb les mateixes hipòtesis que allí, provem una *llei forta*. Ja hem vist en el capítol anterior que la convergència en probabilitat és estrictament més feble que la quasi segura. Per tant, la versió forta del teorema 5.1 requereix una prova més complexa.

Un dels objectius en l'estudi de problemes límit en probabilitats ha estat el d'afeblir al màxim les hipòtesis dels seus enunciats. Aquestes són molt sovint formulades en termes dels moments de les variables de la successió. Pel que fa a les lleis dels grans nombres, en el cas de successions i.i.d., l'existència de moment de primer ordre és una condició necessària i suficient per a la validesa de la llei forta dels grans nombres. Aquest resultat fou establert per Kolmogorov. La segona part d'aquest apartat es dedica a desenvolupar tots els ingredients necessaris per a la seva demostració.

5.2.1 Llei forta dels grans nombres per a variables incorrelacionades

Teorema 5.3 *Sigui $\{X_n, n \geq 1\}$ una successió de variables aleatòries incorrelacionades amb moments de segon ordre fitats uniformement per una constant C. Aleshores*

$$\lim_{n\to\infty} \frac{S_n - E(S_n)}{n} = 0, \tag{5.7}$$

en la convergència q.s.

Demostració: Podem suposar, sense pèrdua de generalitat, que les variables aleatòries són centrades, és a dir, $E(X_n) = 0$ per a tot $n \geq 1$, i demostrar

$$\lim_{n\to\infty} \frac{S_n}{n} = 0.$$

La propietat d'incorrelació implica que

$$E(S_n^2) \leq nC.$$

Per la desigualtat de Txebixev, per a tot $\varepsilon > 0$ i $n \geq 1$,

$$P\{|S_{n^2}| > n^2\varepsilon\} \leq \frac{n^2 C}{n^4\varepsilon^2} = \frac{C}{n^2\varepsilon^2}.$$

Per tant,

$$\sum_{n\geq 1} P\{|S_{n^2}| > n^2\varepsilon\} \leq \sum_{n\geq 1} \frac{C}{n^2\varepsilon^2} < +\infty.$$

Aleshores, pel primer lema de Borel-Cantelli (vegeu lema 4.2),

$$P\Big\{\limsup_n \{|S_{n^2}| > n^2\varepsilon\}\Big\} = 0$$

o, equivalentment,

$$P\Big\{\liminf_n \{|S_{n^2}| \leq n^2\varepsilon\}\Big\} = 1. \tag{5.8}$$

Considerem la successió $\{\frac{S_{n^2}}{n^2}, n \geq 1\}$. Ja que es compleix (5.8), podem aplicar la proposició 4.1. D'on resulta

$$\lim_{n\to\infty} \frac{S_{n^2}}{n^2} = 0, \tag{5.9}$$

quasi segurament.

Acabem, doncs, de demostrar que (5.7) és cert per a la subsuccessió $\{\frac{S_{n^2}}{n^2}, n \geq 1\}$.

Es tracta ara de controlar els termes de la successió que corresponen a índexs k entre n^2 i $(n+1)^2$.

Per a tot $n \geq 1$ definim

$$D_n = \max_{n^2 \leq k \leq (n+1)^2} |S_k - S_{n^2}|.$$

Es compleix que

$$\begin{aligned}
E(D_n^2) &\leq \sum_{k=n^2}^{(n+1)^2} E|S_k - S_{n^2}|^2 \\
&= \sum_{k=n^2}^{(n+1)^2} E(X_{n^2+1} + X_{n^2+2} + \cdots + X_k)^2 \\
&= \sum_{k=n^2}^{(n+1)^2} (EX_{n^2+1}^2 + \cdots + EX_k^2) \\
&\leq \sum_{k=n^2}^{(n+1)^2} (EX_{n^2+1}^2 + \cdots + EX_{(n+1)^2}^2) = (2n+2)E(S_{(n+1)^2} - S_{n^2})^2 \\
&\leq (2n+2)^2 C.
\end{aligned}$$

Per tant,

$$P\{D_n > n^2\varepsilon\} \leq \frac{(2n+2)^2 C}{n^4\varepsilon^2}.$$

Els mateixos arguments utilitzats abans proven que

$$\lim_{n\to\infty} \frac{D_n}{n^2} = 0, \tag{5.10}$$

q.s.

Les convergències (5.9) i (5.10) demostren el resultat del teorema. En efecte, sigui $n^2 \leq k < (n+1)^2$. Aleshores,

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

i aquesta desigualtat permet de concloure la demostració. □

Exemple 5.2 (Els nombres normals de Borel) Un cas interessant d'aplicació de la llei forta dels grans nombres apareix en l'estudi dels anomenats *nombres normals* de Borel (Borel, 1909).

Per a tot nombre $x \in [0,1)$ designarem per $X_n(x)$ la n-èsima xifra decimal de x. Cal indicar que els nombres del conjunt $M = [0,1) \cap \{m10^{-n},\, m,n \geq 0\}$ tenen dues expressions decimals diferents (per exemple, 0,2 i 0,1999...) i cal escollir-ne una, per exemple, la que té zeros a partir d'un cert lloc.

Posem $X_n^{(k)}(x) = \mathbb{1}_{\{X_n = k\}}(x)$ on $k \in \{0,1,2,\ldots,9\}$ és una xifra que fixem.

Aleshores, $\frac{1}{n}\sum_{i=1}^{n} X_i^{(k)}(x)$ representa la freqüència relativa amb què la xifra k apareix en les n primeres xifres decimals de x.

Es diu aleshores que el nombre x és *normal* si

$$\frac{1}{n}\sum_{i=1}^{n} X_i^{(k)}(x) \to \frac{1}{10}, \tag{5.11}$$

quan n tendeix a infinit, per a tota xifra k.

No se sap si nombres irracionals ben coneguts com per exemple $e-2$ o $\pi - 3$ són normals, però en canvi el *teorema de Borel* ens diu que, llevat d'un subconjunt de $[0,1)$ de longitud zero, tots els nombres són normals.

Donem, tot seguit, una demostració d'aquest resultat. Considerem l'espai de probabilitat $([0,1), \mathcal{B}([0,1)), P)$ on P és la llei uniforme en $[0,1]$. En aquest espai, $\{X_n, n \geq 1\}$ és una successió de variables aleatòries independents amb distribució uniforme sobre el conjunt $\{0,1,2,\ldots,9\}$. En efecte, fixem n xifres $k_1, k_2, \ldots, k_n$ i calculem $P\{X_1 = k_1, \ldots, X_n = k_n\}$, que és la probabilitat dels conjunts de $x \in [0,1]$ tals que les primeres n xifres decimals de x són $k_1, k_2, \ldots, k_n$.

Obtenim que

$$P\{X_1 = k_1, \dots, X_n = k_n\} = P\{[0, k_1 \cdots k_n,\ 0, k_1 \cdots k_n + 10^{-n})\} = 10^{-n}.$$

Les variables $\{X_n^{(k)}, n \geq 1\}$ (fixada una xifra k) també seran independents i amb lleis de Bernoulli de paràmetre $P\{X_n^{(k)} = 1\} = P\{X_n = k\} = 10^{-1}$. Per tant, la convergència (5.11) serà certa per a tots els $x \in [0,1)$ excepte en un conjunt de longitud zero, a causa de la llei forta dels grans nombres. Queda així provat el resultat.

5.2.2 Llei forta dels grans nombres de Kolmogorov

En aquest apartat considerarem una successió $\{X_n, n \geq 1\}$ de variables aleatòries independents, idènticament distribuïdes. L'objectiu que ens proposem és demostrar que l'existència de moment de primer ordre per a la llei comuna de les variables de la successió és una condició necessària i suficient per a la validesa de la llei forta dels grans nombres.

Teorema 5.4 *Sigui $\{X_n, n \geq 1\}$ una successió de variables aleatòries i.i.d,*

1. *Suposem $E(|X_1|) < \infty$. Aleshores,*

$$\lim_{n\to\infty} \frac{S_n}{n} = E(X_1), \quad q.s.$$

2. *Recíprocament, si $E(|X_1|) = \infty$,*

$$\limsup_{n\to\infty} \frac{S_n}{n} = \infty, \quad q.s.$$

La idea de la demostració de l'apartat (1) d'aquest teorema consisteix a utilitzar un procediment de truncació que permet de suposar existència de moments de qualsevol ordre per a les variables aleatòries de la successió. Ara bé, la successió truncada ja no és idènticament distribuïda. Per a això cal disposar d'una llei forta dels grans nombres adaptada a aquesta situació. Els resultats que donem a continuació proporcionen els ingredients necessaris.

Proposició 5.1 (Desigualtat de Kolmogorov) *Siguin $X_1, \dots, X_n$ variables aleatòries independents, de quadrat integrable i centrades. Posem $S_k = X_1 + \cdots + X_k, 1 \le k \le n$. Aleshores, per a tot $\varepsilon > 0$ tenim que*

$$P\Big\{ \max_{1\le k\le n} |S_k| > \varepsilon \Big\} \le \frac{\sigma^2(S_n)}{\varepsilon^2}.$$

Demostració: Fixem $\varepsilon > 0$ i posem $A = \big\{ \max_{1\le k\le n} |S_k| > \varepsilon \big\}$. Definim

$$\begin{aligned} A_1 &= \{|S_1| > \varepsilon\} \\ A_k &= \Big\{ \max_{1\le j\le k-1} |S_j| \le \varepsilon, |S_k| > \varepsilon \Big\}, \quad 2 \le k \le n. \end{aligned}$$

Els conjunts A_k, $k = 1, \dots, n$ formen una partició del conjunt A.
Es compleix que

$$\begin{aligned} \sigma^2(S_n) &= E\left(S_n^2\right) \ge E\left(S_n^2 \mathbb{1}_A\right) = \sum_{k=1}^n E\left(S_n^2 \mathbb{1}_{A_k}\right) \\ &= \sum_{k=1}^n E\left((S_k + (S_n - S_k))^2 \mathbb{1}_{A_k}\right) \ge \sum_{k=1}^n E\left(\left(S_k^2 + 2S_k(S_n - S_k)\right) \mathbb{1}_{A_k}\right). \end{aligned} \tag{5.12}$$

D'altra banda, per la propietat d'independència,

$$E\left(S_k(S_n - S_k)\mathbb{1}_{A_k}\right) = E\big(\mathbb{1}_{A_k} S_k\big) E(S_n - S_k) = 0. \tag{5.13}$$

Finalment, com que en el conjunt A_k, S_k^2 és més gran que ε^2, a partir de (5.12) i (5.13) obtenim que

$$\sigma^2(S_n) \ge \sum_{k=1}^n E\left(S_k^2 \mathbb{1}_{A_k}\right) \ge \varepsilon^2 \sum_{k=1}^n P(A_k) = \varepsilon^2 P(A).$$

□

Per a $n = 1$, la desigualtat de Kolmogorov coincideix amb el que obtindríem aplicant la desigualtat de Txebixev. Per a $n > 1$, és un refinament d'aquesta. El pas següent consisteix a establir un criteri de convergència quasi segura per a sèries aleatòries.

Proposició 5.2 *Sigui* $\{X_n, n \geq 1\}$ *una successió de variables aleatòries independents, de quadrat integrable i centrades. Si la sèrie* $\sum_{n=1}^{\infty} \sigma^2(X_n)$ *convergeix, aleshores la sèrie* $\sum_{n=1}^{\infty} X_n$ *convergeix quasi segurament.*

Demostració: Aplicant la desigualtat de Kolmogorov a les variables X_{m+1}, $\ldots, X_{m+k}$, $m \geq 0$, s'obté que

$$\begin{aligned} P\Big\{\sup_{j\geq 1} |S_{m+j} - S_m| > \varepsilon\Big\} &= \lim_{k\to\infty} P\Big\{\sup_{1\leq j\leq k} |S_{m+j} - S_m| > \varepsilon\Big\} \\ &\leq \lim_{k\to\infty} \frac{1}{\varepsilon^2} \sum_{j=1}^{k} \sigma^2(X_{m+j}) = \frac{1}{\varepsilon^2} \sum_{j=1}^{\infty} \sigma^2(X_{m+j}). \end{aligned}$$

Per hipòtesi, aquesta darrera expressió tendeix a zero quan m tendeix a infinit, i això implica la convergència quasi segura de la successió S_n. En efecte, la convergència en probabilitat, quan $n \to \infty$,

$$sup_{j\geq 1} |S_{m+j} - S_m| \to 0$$

implica l'existència d'una successió estrictament creixent de nombres naturals m_i tal que $\sup_{j\geq 1} |S_{m_i+j} - S_{m_i}| \xrightarrow[i\to\infty]{} 0, q.s.$ i, per tant, la successió $\{S_n,\, n \geq 1\}$ és de Cauchy amb probabilitat 1. $\square$

Exemple 5.3 És ben sabut que la sèrie harmònica $\sum_{n=1}^{\infty} \frac{1}{n}$ és divergent mentre que la sèrie alternada $\sum_{n=1}^{\infty} \frac{(-1)^n}{n}$ convergeix. Considerem una successió $\{\varepsilon_n, n \geq 1\}$ de variables aleatòries i.i.d. tals que $P\{\varepsilon_n = 1\} = P\{\varepsilon_n = -1\} = 1/2$. Aleshores, la sèrie aleatòria $\sum_{n=1}^{\infty} \frac{\varepsilon_n}{n^p}$ convergeix q.s. si $p > 1/2$ ja que la sèrie de variàncies és igual a $\sum_{n=1}^{\infty} \frac{1}{n^{2p}} < \infty$. Observeu que per a $1/2 < p \leq 1$ aquesta sèrie aleatòria no convergeix absolutament.

Lema 5.1 (Lema de Kronecker) *Sigui* $\{x_n, n \geq 1\}$ *una successió de nombres reals i* $\{a_n, n \geq 1\}$ *una altra successió de nombres reals tal que* $0 < a_n \uparrow \infty$. *Aleshores, si* $\sum_{n=1}^{\infty} \frac{x_n}{a_n}$ *convergeix, es té que*

$$\lim_n \frac{1}{a_n} \sum_{j=1}^{n} x_j = 0.$$

Demostració: Per a cada $n \geq 1$ posem $b_n = \sum_{j=1}^{n} \frac{x_j}{a_j}$. També escrivim, per conveni, $a_0 = b_0 = 0$. Llavors, $x_n = a_n(b_n - b_{n-1})$, per a $n \geq 1$ i, mitjançant el mètode de sumació parcial d'Abel, obtenim que

$$\begin{aligned}
\frac{1}{a_n}\sum_{j=1}^{n} x_j &= \frac{1}{a_n}\sum_{j=1}^{n} a_j(b_j - b_{j-1}) \\
&= \frac{1}{a_n}\big(a_1b_1 + a_2b_2 - a_2b_1 + a_3b_3 - a_3b_2 + \cdots \\
&+ a_{n-1}b_{n-1} - a_{n-1}b_{n-2} + a_nb_n - a_nb_{n-1}\big) \\
&= b_n - \frac{1}{a_n}\sum_{j=0}^{n-1} b_j(a_{j+1} - a_j).
\end{aligned}$$

Sigui $b_\infty = \lim_n b_n$ que, per hipòtesi, és finit. Cal provar que

$$\lim_n \frac{1}{a_n}\sum_{j=0}^{n-1} b_j(a_{j+1} - a_j) = b_\infty.$$

Observem que $\frac{1}{a_n}\sum_{j=0}^{n-1}(a_{j+1} - a_j) = 1$. En conseqüència,

$$\begin{aligned}
\left|\frac{1}{a_n}\sum_{j=0}^{n-1} b_j(a_{j+1} - a_j) - b_\infty\right| &= \left|\frac{1}{a_n}\sum_{j=0}^{n-1}(b_j - b_\infty)(a_{j+1} - a_j)\right| \\
\leq \frac{1}{a_n}\left|\sum_{j=0}^{m-1}(b_j - b_\infty)(a_{j+1} - a_j)\right| \\
+ \frac{1}{a_n}\left|\sum_{j=m}^{n-1}(b_j - b_\infty)(a_{j+1} - a_j)\right|,
\end{aligned}$$

per a $n > m$.

Fixat un $\varepsilon > 0$, sigui m tal que $|b_j - b_\infty| < \varepsilon$, per a tot $j \geq m$. Tenint en compte que $a_{j+1} - a_j \geq 0$, el segon sumand de l'expressió anterior estarà fitat per ε per a aquest valor de m. Per tant, si fixem ε i m, tindrem que

$$\limsup_n \left|\frac{1}{a_n}\sum_{j=0}^{n-1} b_j(a_{j+1} - a_j) - b_\infty\right| \leq \varepsilon,$$

ja que el primer sumand té límit zero quan $n \to \infty$. Com que ε és arbitrari obtenim el resultat desitjat. $\square$

El lema de Kronecker i el criteri de convergència q.s. per a sèries aleatòries ens permeten demostrar la versió següent de la llei forta dels grans nombres per a variables de quadrat integrable, no necessàriament idènticament distribuïdes.

Teorema 5.5 *Sigui* $\{X_n, n \geq 1\}$ *una successió de variables aleatòries independents, de quadrat integrable i centrades. Considerem una successió* $\{a_n, n \geq 1\}$ *de nombres reals tal que* $0 < a_n \uparrow \infty$. *Aleshores, si* $\sum_{n=1}^{\infty} \frac{\sigma^2(X_n)}{a_n^2} < \infty$*, es compleix que*

$$\frac{S_n}{a_n} \to 0,$$

quasi segurament, quan n *tendeix a infinit.*

Demostració: Per hipòtesi,

$$\sum_{n=1}^{\infty} \sigma^2 \left(\frac{X_n}{a_n} \right) = \sum_{n=1}^{\infty} \frac{\sigma^2(X_n)}{a_n^2} < \infty.$$

La proposició 5.2 aplicada a la successió $\frac{X_n}{a_n}$ implica la convergència quasi segura de la sèrie $\sum_{n=1}^{\infty} \frac{X_n}{a_n}$ i, pel lema de Kronecker,

$$\lim_{n\to\infty} \frac{1}{a_n} S_n = \lim_{n\to\infty} \frac{1}{a_n} \sum_{j=1}^{n} X_j = 0, q.s.$$

$\square$

En el cas particular de la successió definida per $a_n = n$, el teorema anterior estableix que la condició $\sum_{n=1}^{\infty} \frac{\sigma^2(X_n)}{n^2} < \infty$ és suficient per a assegurar la convergència quasi segura

$$\frac{S_n}{n} \to 0 \tag{5.14}$$

quan n tendeix a infinit.

En particular, si $\{X_n, n \geq 1\}$ és una successió de variables aleatòries i.i.d, de quadrat integrable, amb $E(X_1) = m$ i $\sigma^2(X_1) = \sigma^2$ retrobem el resultat

$$\lim_{n\to\infty} \frac{S_n}{n} = m, \tag{5.15}$$

quasi segurament.

El darrer ingredient abans de poder donar la demostració del teorema 5.4 és el següent:

Lema 5.2 *Per a tota variable aleatòria Y es compleix que*

$$\sum_{n=1}^{\infty} P\big\{|Y| \geq n\big\} \leq E\big(|Y|\big) \leq 1 + \sum_{n=1}^{\infty} P\big\{|Y| \geq n\big\}.$$

En altres paraules, la variable Y té esperança finita si, i només si, la sèrie $\sum_{n=1}^{\infty} P\big\{|Y| \geq n\big\}$ és convergent.

Demostració: Podem escriure

$$\begin{aligned}
\sum_{n=1}^{\infty} P\{|Y| \geq n\} &= \sum_{n=1}^{\infty} \sum_{k=n}^{\infty} P\{k \leq |Y| < k+1\} \\
&= \sum_{k=1}^{\infty} kP\{k \leq |Y| < k+1\} = \sum_{k=0}^{\infty} kP\{k \leq |Y| < k+1\}.
\end{aligned}$$

Aquesta darrera sèrie està majorada per $\sum_{k=0}^{\infty} E\left(|Y| 1\!\!1_{\{k\leq|Y|<k+1\}}\right)$, per la propietat de monotonia de la integral. Obtenim així la primera desigualtat. D'altra banda, fent una partició del conjunt Ω,

$$\begin{aligned}
E(|Y|) &= \sum_{k=0}^{\infty} E\left(|Y| 1\!\!1_{\{k\leq|Y|<k+1\}}\right) \\
&\leq \sum_{k=0}^{\infty} (k+1)P\{k \leq |Y| < k+1\} \leq \sum_{n=1}^{\infty} P\{|Y| \geq n\} + 1.
\end{aligned}$$

□

Demostració del teorema 5.4: Definim $Y_n = X_n \mathbb{1}_{\{|X_n|<n\}}$. Pel lema 5.2,

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

Pel primer lema de Borel-Cantelli (vegeu el lema 4.2),

$$P\left\{\lim_n \sup\{X_n \neq Y_n\}\right\} = 0,$$

és a dir, si $G = \liminf_n \{X_n = Y_n\}$, tindrem $P(G) = 1$.
Sigui $\omega \in G$. Existeix un $n_0(\omega)$ tal que $X_n(\omega) = Y_n(\omega)$ per a tot $n \geq n_0(\omega)$. Per tant, per a tot $n \geq n_0(\omega), \omega \in G$.

$$\frac{1}{n}\sum_{i=1}^{n} X_i(\omega) = \frac{1}{n}\sum_{i=1}^{n} Y_i(\omega) + \frac{1}{n}\sum_{i=1}^{n_0(\omega)} (X_i - Y_i)(\omega).$$

Per tant, és suficient demostrar que

$$\frac{1}{n}\sum_{i=1}^{n} Y_i \to E(X_1), \tag{5.16}$$

quan n tendeix a infinit, q.s.
Pel teorema de la convergència dominada (proposició 3.4(2)),

$$E(Y_n) = E(X_n \mathbb{1}_{\{|X_n|<n\}}) = E(X_1 \mathbb{1}_{\{|X_n|<n\}}) \to E(X_1),$$

quan n tendeix a infinit.
Així, doncs,

$$\frac{1}{n}\sum_{i=1}^{n} E(Y_i) \to E(X_1), \quad n \to \infty.$$

Per tant, per provar (5.16) és suficient demostrar que

$$\frac{1}{n}\sum_{i=1}^{n}\big(Y_i - E(Y_i)\big) \to 0, \quad n \to \infty. \tag{5.17}$$

Amb aquest objectiu, comprovem que la successió $\{Y_n - E(Y_n), n \geq 1\}$ de variables independents, centrades i de quadrat integrable, satisfà les hipòtesis del teorema 5.5, és a dir, $\sum_{n=1}^{\infty} \frac{\sigma^2(Y_n)}{n^2} < \infty$.
En efecte,

$$\begin{aligned}
\sum_{n=1}^{\infty} \frac{\sigma^2(Y_n)}{n^2} &= \sum_{n=1}^{\infty} \frac{E(Y_n^2)}{n^2} = \sum_{n=1}^{\infty} \frac{1}{n^2} E\big(X_n^2 1\!\!1_{\{|X_n|<n\}}\big) \\
&= \sum_{n=1}^{\infty} \frac{1}{n^2} E\big(X_1^2 1\!\!1_{\{|X_1|<n\}}\big) = \sum_{n=1}^{\infty} \frac{1}{n^2} \sum_{k=1}^{n} E\big(X_1^2 1\!\!1_{\{k-1\leq |X_1|<k\}}\big) \\
&= \sum_{k=1}^{\infty} E\big(X_1^2 1\!\!1_{\{k-1\leq |X_1|<k\}}\big) \sum_{n=k}^{\infty} \frac{1}{n^2} \\
&\leq \sum_{k=1}^{\infty} kE\big(|X_1| 1\!\!1_{\{k-1\leq |X_1|<k\}}\big) \frac{2}{k} \\
&\leq 2E(|X_1|) < \infty,
\end{aligned}$$

on hem fet servir, entre d'altres, la fitació

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

Això acaba la demostració de l'apartat (1).

Sigui $K > 0$. El lema 5.2 ens permet d'escriure

$$\sum_{n=1}^{\infty} P\big\{|X_n| \geq Kn\big\} = \sum_{n=1}^{\infty} P\left\{\frac{|X_1|}{K} \geq n\right\} \leq \frac{1}{K} E(|X_1|) = \infty.$$

El segon lema de Borel-Cantelli (vegeu lema 4.3) ens assegura $P(G_K) = 1$, on

$$G_K = \limsup \big\{|X_n| \geq Kn\big\}.$$

Posem $G = \bigcap_{K=1}^{\infty} G_K$. Clarament, $P(G) = 1$.

Sobre el conjunt G, es compleix $\frac{|X_n|}{n} \geq K$ per a infinits valors de n, per a tot natural $K \geq 1$. Això implica

$$\frac{|S_n|}{n} + \frac{|S_{n-1}|}{n-1} > \frac{|S_n| + |S_{n-1}|}{n} \geq \frac{|S_n - S_{n-1}|}{n} = \frac{|X_n|}{n} \geq K$$

per a infinits valors de n, és a dir, $\frac{|S_n|}{n} \geq \frac{K}{2}$ per a infinits valors de n i, en conseqüència, $\limsup_n \frac{|S_n|}{n} = \infty$ en el conjunt G. Com que aquest conjunt té probabilitat 1, això acaba la demostració del teorema. $\square$

Exercicis

5.1 *Es considera una successió de variables aleatòries i.i.d, $\{X_n, n \geq 1\}$, amb distribució $N(0,1)$. Demostreu*

$$\lim_{n\to\infty} \frac{X_1^2 + \cdots + X_n^2}{n} = 1,$$

en el sentit de la convergència quasi segura.

Per a $n \geq 1$ i $\varepsilon > 0$, definim $Y^{(n)} = (X_1, \ldots, X_n)$ i la corona

$$B_{n,\varepsilon} = \{(x_1, \ldots, x_n) \in \mathbb{R}^n : (1-\varepsilon)n \leq x_1^2 + \cdots + x_n^2 \leq (1+\varepsilon)n\}.$$

Demostreu

$$\lim_{n\to\infty} P\{Y^{(n)} \notin B_{n,\varepsilon}\} = 0.$$

5.2 *Sigui $\{X_n, n \geq 1\}$ una successió de variables aleatòries independents amb distribució comuna del tipus exponencial amb paràmetre $\lambda > 0$. Demostreu*

$$\lim_{n\to\infty} S_n = \infty.$$

Per a cada $t > 0$ definim

$$N(t) = \sum_{n=1}^{\infty} 1\!\!1_{\{S_n \le t\}}.$$

Per exemple, els diferents S_n poden representar els temps en què es produeixen determinats esdeveniments i, aleshores, $N(t)$ és el nombre d'esdeveniments que s'han produït en l'interval $(0, t]$ (procés de Poisson). Demostreu

$$\lim_{n\to\infty} \frac{N(t)}{t} = \lambda, \quad q.s.$$

Indicació: Demostreu primer

$$\lim_{n\to\infty} \frac{N(S_n)}{S_n} = \lambda, \quad q.s.$$

i, després, utilitzeu la primera part de l'exercici.

5.3 *Considerem una successió de variables aleatòries independents $X_n, n \ge 1$, tals que*

$$P\{X_n = n^\alpha\} = \frac{1}{2} = P\{X_n = 1 - n^\alpha\}, \quad \alpha > 0.$$

Estudieu per a quins valors de α es pot esperar convergència quasi segura de la successió $\frac{1}{n}\sum_{i=1}^{n} X_i, n \ge 1$.

5.4 *Sigui $X_n, n \ge 1$, una successió de variables aleatòries independents, amb llei absolutament contínua $\exp(\lambda_n)$, $\lambda_n > 0, n \ge 1$. Doneu una condició suficient sobre la successió $\lambda_n, n \ge 1$, perquè la sèrie aleatòria $\sum_{i=1}^{n} X_i$ convergeixi quasi segurament.*

5.5 *Considerem una successió $X_n, n \ge 1$, de variables aleatòries independents, amb distribució de Bernoulli amb paràmetre p. Definim*

$$Y_n = 1\!\!1_{\{X_n = X_{n+2} = 1\}}, \quad n \ge 1.$$

Calculeu el límit en la convergència quasi segura de la successió $\{\frac{1}{n}\sum_{i=1}^{n} Y_i, n \ge 1\}$.

5.6 *Sigui $X_n, n \ge 1$, una successió de variables aleatòries centrades, amb moments de segon ordre uniformement fitats, que compleixen*

$$\lim_{|k-l|\to\infty} \mathrm{Cov}(X_k, X_l) = 0.$$

Demostreu

$$\lim_{n\to\infty} \frac{\sum_{i=1}^{n} X_i}{n} = 0,$$

quasi segurament.

5.7 *Sigui* $X_n, n \geq 1$, *una successió de variables aleatòries independents, totes amb llei uniforme en l'interval (0,1), i* $f : [0,1] \longrightarrow \mathbb{R}_+$ *una funció contínua. Demostreu*

$$\frac{1}{n}\sum_{i=1}^{n} f(X_i) \to \int_0^1 f(x)\,\mathrm{d}x,$$

$n \to \infty$, *q.s.*

Capítol 6

El teorema del límit central

L'objectiu d'aquest capítol és demostrar algunes versions elementals de l'anomenat teorema del límit central. Aquest resultat es refereix al comportament asimptòtic, en la convergència en distribució, de la successió $\{S_n, n \geq 1\}$ de les sumes parcials d'una família $\{X_n, n \geq 1\}$ de variables aleatòries independents, convenientment normalitzades. Així, doncs, el problema que plantegem té un caràcter diferent del que hem estudiat en el capítol 5, encara que té en comú l'estudi del comportament en el límit de sumes de variables aleatòries independents.

Sense parlar de moment de manera gaire precisa, podem dir que el teorema del límit central estableix la convergència en llei cap a la distribució normal de la successió de sumes de variables aleatòries independents (no és necessari que tinguin la mateixa distribució), amb moments de segon ordre finit, estandarditzades pel factor $\sqrt{n}$. El fet sorprenent és que el límit sigui independent de la llei de les variables de la successió. Això és el que vol dir *límit central*: el paper central que juga la llei normal o de Gauss.

El teorema del límit central té moltíssimes aplicacions en la pròpia teoria de la probabilitat. Ens interessa, però, destacar aquí el seu paper crucial en l'estadística. Gairebé tots els resultats sobre la distribució límit d'estimadors i d'estadístics emprats en els tests d'hipòtesis es basen en aquest resultat.

Hi ha versions molt diverses d'aquest teorema i també poden utilitzar-se tècniques diferents en les seves demostracions. Potser la més usual és la proporcionada per les *funcions característiques* o transformades de Fourier de la llei de probabilitat. Tenint present el nivell elemental d'aquest llibre, hem optat per no presentar el mètode de les funcions característiques i per donar una versió particular, però prou interessant, del teorema.

6.1 Convergència de la llei binomial

Històricament, el problema que va motivar la primera versió del teorema del límit central va ser l'estudi del límit, quan n tendeix a infinit, de les probabilitats associades a la llei binomial. Més precisament: Sigui $p \in (0,1)$, $q = 1 - p$. Definim, per a $k = 0, 1, \dots, n$,

$$P_n(k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Recordem que, si p depèn de n de la manera particular descrita en l'exemple 4.10,

$$\lim_{n\to\infty} P_n(k) = \frac{\lambda^k}{k!} \exp(-\lambda).$$

Aquest resultat és el que s'anomena de convergència de la binomial cap a la llei de Poisson. Aquí estem interessats en la convergència de la successió $\{P_n, n \geq 1\}$ en altres circumstàncies.

Si representem gràficament l'aplicació

$$x \longrightarrow P_n(np + x\sqrt{npq}),$$

on x varia en el conjunt dels nombres reals tals que $np + x\sqrt{npq} \in \{0, 1, \dots, n\}$, trobarem una gràfica com la de la figura 6.1.

Veiem que els extrems dels bastonets s'ajusten bé a la representació de la funció $f(x) = \frac{1}{\sqrt{2\pi pq}} \exp(-\frac{x^2}{2})$. Això fa pensar en un tipus d'aproximació diferent del donat per la llei de Poisson i ens porta, de manera natural, a parlar del teorema de De Moivre.

La primera versió del teorema del límit central fou demostrada per A. De Moivre en *The Doctrine of Chances* (1718). Tot seguit en donem l'enunciat.

Teorema 6.1 (teorema del límit central de De Moivre-Laplace) *Considerem una successió $\{X_n, n \geq 1\}$ de variables aleatòries independents, totes amb la mateixa distribució de Bernoulli amb paràmetre p. Aleshores, qualssevol que siguin $-\infty < a \leq b < \infty$, es compleix la propietat següent:*

$P_n(np + x\sqrt{npq})$

$\frac{1}{\sqrt{2\pi npq}} \mathrm{e}^{-x^2/2}$

0 x

Figura 6.1: Aproximació de la binomial per la normal

$$\lim_{n\to\infty} P\left\{a < \frac{S_n - np}{\sqrt{np(1-p)}} \leq b\right\} = \int_a^b \frac{1}{\sqrt{2\pi}} \exp\Big(-\frac{x^2}{2}\Big)\,\mathrm{d}x. \tag{6.1}$$

Observem, en primer lloc, la diferència entre aquest resultat i la llei forta dels grans nombres. Denotem per F la funció de distribució de la llei $\mathcal{N}(0,1)$ i per F_n la de la variable aleatòria $(S_n - np)/\sqrt{np(1-p)}$, $n \geq 1$. El resultat (6.1) expressa una convergència de les funcions de distribució F_n cap a F. És, doncs, un resultat de *convergència en llei o en distribució*. D'altra banda, en el teorema 6.1 les sumes centrades $S_n - np$ es normalitzen mitjançant el factor $\sqrt{n}$, mentre que en les lleis dels grans nombres es fa per n. Aquesta diferència pot explicar-se així: En el context de les lleis dels grans nombres, la variable aleatòria límit seria una constant, és a dir, una variable aleatòria discreta. Aquí, en canvi, la variable aleatòria límit (en el sentit de la convergència en llei) és absolutament contínua; és clar que això ha de suposar una novetat en la successió que hem d'estudiar. La diferència s'introdueix en el factor de normalització.

No donem en aquest apartat la demostració del teorema 6.1. La nostra intenció és de provar més endavant una versió del teorema del límit central que englobarà la presentada en el teorema anterior.

Un objectiu important en la teoria de la probabilitat ha estat el de generalitzar el resultat de De Moivre. Per exemple, a una successió qualsevol de variables aleatòries independents i idènticament distribuïdes (en aquest cas el factor de normalització és $\sqrt{n}\sigma$, on denotem per σ^2 la variància comuna de la distribució de les variables de la successió). Una extensió posterior consisteix a no suposar distribució idèntica. En aquest cas, cal imposar una restricció sobre les variàncies de les variables aleatòries de la successió. Volem, però, insistir en el fet que, en totes les extensions, la hipòtesi que s'imposa és l'existència de moment de segon ordre per a les variables aleatòries de la successió, i que aquest forma part del factor de normalització.

6.2 El teorema del límit central de Lévy-Lindeberg

En aquest apartat demostrarem una versió del teorema del límit central per a variables aleatòries independents, idènticament distribuïdes i amb moment de segon ordre finit. Com a cas particular trobarem el teorema de De Moivre presentat en la secció anterior.

El primer resultat que donem és auxiliar. Permet d'establir la velocitat de convergència en problemes de convergència en llei.

Teorema 6.2 *Siguin $\{X_i, i = 1, \ldots, n\}$ una família de variables aleatòries independents, centrades i fitades per una constant $c \in (0, \infty)$. Definim $\sigma_i^2 = E(X_i)^2$, $\sigma^2 = \sum_{i=1}^n \sigma_i^2$. Sigui f una funció real de classe $\mathcal{C}^3$ fitada i amb derivades fitades. Si Y és una variable aleatòria amb llei $\mathcal{N}(0, \sigma^2)$ es té la fitació següent:*

$$|Ef(X_1 + \cdots + X_n) - Ef(Y)| \leq kc_3c\sigma^2, \tag{6.2}$$

on k és una constant independent de n i $c_3 = \sup_{x\in\mathbb{R}} |f'''(x)|$.

Demostració: Considerem una família $\{Y_i, i = 1, \ldots, n\}$ de variables aleatòries independents amb lleis $\mathcal{N}(0, \sigma_i^2)$, independent de la família $\{X_i, i = 1, \ldots, n\}$. En l'apèndix d'aquest capítol demostrem que la variable aleatòria $\sum_{i=1}^n Y_i$ té llei $\mathcal{N}(0, \sigma^2)$. Per tant, sense pèrdua de generalitat, per demostrar (6.2) podem suposar que $Y = \sum_{i=1}^n Y_i$.

Es compleix que

$$
\begin{aligned}
&|Ef(X_1 + \cdots + X_n) - Ef(Y)| \\
&\quad \le \sum_{i=1}^{n} \big| E\big(f(X_1 + \cdots + X_{i-1} + X_i + Y_{i+1} + \cdots + Y_n)\big) \\
&\quad - E\big(f(X_1 + \cdots + X_{i-1} + Y_i + Y_{i+1} + \cdots + Y_n)\big)\big|.
\end{aligned} \tag{6.3}
$$

Un sumand típic del segon membre de (6.3) és

$$|Ef(X_i + U_i) - Ef(Y_i + U_i)|,$$

on U_i és una variable aleatòria que és independent de X_i i Y_i.
Fixem $\omega \in \Omega$. Per la fórmula de Taylor tenim que

$$f(X_i + U_i)(\omega) = \big(f(U_i) + f'(U_i)X_i + \frac{1}{2} f''(U_i)X_i^2 + \frac{1}{6} f'''(\xi)X_i^3\big)(\omega), \tag{6.4}$$

on $\xi(\omega)$ és un punt intermedi entre $X_i(\omega) + U_i(\omega)$ i $U_i(\omega)$.
Prenent esperances en (6.4), i tenint en compte que $E(X_i) = 0$ per a tot $i = 1, \ldots, n$ s'obté que

$$Ef(X_i + U_i) = Ef(U_i) + \frac{1}{2} Ef''(U_i)\sigma_i^2 + \frac{1}{6} E\big(f'''(\xi)X_i^3\big).$$

Anàlogament,

$$Ef(Y_i + U_i) = Ef(U_i) + \frac{1}{2} Ef''(U_i)\sigma_i^2 + \frac{1}{6} E\big(f'''(\bar{\xi})Y_i^3\big),$$

on $\bar{\xi}$ es defineix de manera anàloga a ξ.
En conseqüència,

$$|Ef(X_i + U_i) - Ef(Y_i + U_i)| \le \frac{1}{6} c_3 \big(E|X_i|^3 + E|Y_i|^3\big). \tag{6.5}$$

Les hipòtesis sobre les variables aleatòries X_i impliquen

$$E|X_i|^3 \le c\sigma_i^2.$$

D'altra banda, fent una integració per parts resulta que

$$E|Y_i|^3 = \frac{2}{\sqrt{2\pi}\sigma_i} \int_0^\infty x^3 \mathrm{e}^{-\frac{x^2}{2\sigma_i^2}} \, \mathrm{d}x = \frac{4\sigma_i^3}{\sqrt{2\pi}}.$$

Finalment, com que $|X_i| \leq c$, es compleix $\sigma_i \leq c$.
Utilitzant aquestes estimacions en (6.5) s'obté que

$$|Ef(X_i + U_i) - Ef(Y_i + U_i)| \leq \frac{1}{6} c_3 (c\sigma_i^2 + ck'\sigma_i^2),$$

amb $k' = \frac{4}{\sqrt{2\pi}}$. Això demostra (6.2) i dóna com a valor de k, $k = \frac{1}{6} + \frac{1}{6}\frac{4}{\sqrt{2\pi}}$.
□

En la proposició següent donem una versió particular del teorema del límit central que es dedueix de manera immediata del teorema 6.2.

Proposició 6.1 (Teorema del límit central, cas fitat) *Sigui* $\{Z_n, n \geq 1\}$ *una successió de variables aleatòries independents, idènticament distribuïdes i fitades. Designem per* σ^2 *la variància de* Z_n*, per a tot* $n \geq 1$*, i suposem* $\sigma^2 > 0$*. Si* $S_n = Z_1 + \cdots + Z_n$*, es compleix que*

$$\mathcal{L} - \lim_{n\to\infty} \frac{S_n - ES_n}{\sigma\sqrt{n}} = Y, \tag{6.6}$$

on Y *és una variable amb distribució* $\mathcal{N}(0,1)$*.*

Demostració: Sigui $f \in \mathcal{C}_b^\infty$. Apliquem el teorema 6.2 a les variables aleatòries

$$X_i = \frac{Z_i - EZ_i}{\sqrt{n}\sigma},$$

que són independents i centrades.
Es compleix que

$$\sup_{1\leq i\leq n} |X_i| \leq \frac{C}{\sigma\sqrt{n}},$$

on

$$C = \sup_n |Z_n - EZ_n|$$

i $\sum_{i=1}^n EX_i^2 = 1$. La desigualtat (6.2) ens dóna

$$\left| Ef\left(\frac{S_n - ES_n}{\sigma\sqrt{n}}\right) - Ef(Y)\right| \leq kc_3 \frac{C}{\sigma\sqrt{n}}. \tag{6.7}$$

El segon terme de (6.7) convergeix cap a zero quan n tendeix cap a infinit. El teorema 4.1 (vegeu, en particular, l'observació 4.1) ens permet de concloure la demostració. □

Una aplicació senzilla de la proposició 6.1 ens permet de demostrar el teorema de De Moivre-Laplace. En efecte, suposem que les variables aleatòries $\{Z_n, n \geq 1\}$ del teorema anterior tenen la distribució $B(1,p)$. Òbviament, les hipòtesis del teorema es compleixen amb $C = 1$. Per tant, hi ha convergència en llei de la successió $\frac{S_n - np}{\sqrt{np(1-p)}}$ cap a una variable aleatòria amb llei $\mathcal{N}(0,1)$.

L'objectiu final d'aquest apartat és millorar la proposició 6.1, afeblint les hipòtesis. Més precisament, volem treure la condició de fitació sobre les variables de la successió. La idea que utilitzarem serà ben natural: truncarem les variables aleatòries. D'aquesta manera el problema es reduirà a demostrar un resultat per a les variables aleatòries truncades i a controlar la part no truncada. Aquesta és una tècnica usual en la teoria de la probabilitat. Observem, per exemple, la seva utilització en el capítol 5 per a demostrar la llei forta dels grans nombres amb les hipòtesis menys restrictives possibles (vegeu el teorema 5.4).

Teorema 6.3 (Teorema del límit central de Lévy-Lindeberg) *Sigui $\{Z_n, n \geq 1\}$ una successió de variables aleatòries independents, idènticament distribuïdes i de quadrat integrable. Designem per m i σ^2 els valors de la mitjana i la variància de la distribució comuna. Suposem $\sigma^2 > 0$. Si $S_n = Z_1 + \cdots + Z_n$, es compleix que*

$$\mathcal{L} - \lim_{n\to\infty} \frac{S_n - nm}{\sigma\sqrt{n}} = Y, \tag{6.8}$$

on Y és una variable amb distribució $\mathcal{N}(0,1)$.

Demostració: Per a tot $j \geq 1$, $n \geq 1$, definim

$$Y_j^{(n)} = X_j \mathbb{1}_{\{|X_j| \leq n^{1/3}\}},$$

$m_n = E(Y_j^{(n)})$, $\sigma_n = \sqrt{\operatorname{Var}(Y_j^{(n)})}$.
Tenim la descomposició següent:

$$\frac{\sum_{j=1}^n (X_j - m)}{\sigma\sqrt{n}} = A_1^n + A_2^n,$$

amb

$$\begin{aligned} A_1^n &= \frac{\sum_{j=1}^n (Y_j^{(n)} - m_n)}{\sigma\sqrt{n}}, \\ A_2^n &= \frac{\sum_{j=1}^n \left(X_j \mathbb{1}_{\{|X_j| > n^{1/3}\}} + m_n - m \right)}{\sigma\sqrt{n}}. \end{aligned}$$

Definim

$$Z_j^{(n)} = \frac{Y_j^{(n)} - m_n}{\sigma_n \sqrt{n}}, \quad j \geq 1.$$

Aleshores,

$$A_1^n = \sum_{j=1}^n Z_j^{(n)} \frac{\sigma_n}{\sigma}.$$

Volem ara demostrar les propietats següents.

$$\lim_{n\to\infty} \frac{\sigma_n}{\sigma} = 1, \tag{6.9}$$

$$\mathcal{L} - \lim_{n\to\infty} \sum_{j=1}^n Z_j^{(n)} = Y, \tag{6.10}$$

on Y és una variable amb distribució $\mathcal{N}(0,1)$,

$$L^2 - \lim_{n\to\infty} A_2^n = 0.$$

Una vegada provades aquestes propietats, la convergència expressada en (6.8) s'obté a partir del teorema de Slutsky (vegeu la proposició 4.10).

Demostrem (6.9). Per la definició de $Y_j^{(n)}$,

$$E\left(Y_j^{(n)}\right)^2 = E\left(X_j^2 \mathbb{1}_{\{|X_j|\le n^{1/3}\}}\right).$$

Pel teorema de la convergència monòtona (proposició 3.4(1)), aquesta successió tendeix a $E(X_j^2)$ quan $n\to\infty$.

De manera anàloga, pel teorema de la convergència dominada (proposició 3.4(2)),

$$\lim_{n\to\infty} E\left(Y_j^{(n)}\right) = \lim_{n\to\infty} E\left(X_j \mathbb{1}_{\{|X_j|\le n^{1/3}\}}\right) = E(X_j).$$

En conseqüència,

$$\lim_{n\to\infty}\sigma_n^2 = \lim_{n\to\infty}\left(E(Y_j^{(n)})^2 - (E(Y_j^{(n)}))^2\right) = \sigma^2.$$

La demostració de (6.10) utilitza el teorema 6.2. En efecte, les variables aleatòries $\{Z_j^{(n)}, j\ge 1\}$ són independents, centrades i

$$|Z_j^{(n)}| \le \frac{|Y_j^{(n)}| + |m_n|}{\sigma_n\sqrt{n}} \le \frac{2n^{\frac{1}{3}}}{\sigma_n\sqrt{n}}.$$

Endemés, $\mathrm{Var}(Z_j^{(n)}) = \frac{1}{n}$. Per tant, per a tota f de classe $\mathcal{C}^3$ fitada i amb derivades fitades,

$$|E(f(Z_1^{(n)} + \cdots + Z_n^{(n)}) - Ef(Y))| \le kc_3\frac{2n^{\frac{1}{3}}}{\sigma_n\sqrt{n}},$$

que tendeix a zero quan n tendeix a infinit, ja que $\lim_{n\to\infty}\sigma_n = \sigma$. Aleshores, (6.10) és una conseqüència del teorema 4.1.

Observem que

$$E\left(X_j \mathbb{1}_{\{|X_j|\ge n^{1/3}\}}\right) = m - m_n.$$

Per tant, pel teorema de convergència dominada,

$$L^2 - \lim_{n\to\infty} A_2^n = \lim_{n\to\infty} \frac{\mathrm{Var}\left(X_1 \mathbb{1}_{\{|X_1|\geq n^{1/3}\}}\right)}{\sigma^2} = 0.$$

Això acaba la demostració del teorema. □

6.3 Algunes aplicacions del teorema del límit central

Ja hem comentat en la introducció que les aplicacions del teorema del límit central són nombroses. En aquesta secció en presentarem una petita mostra.

6.3.1 Una demostració de la llei feble dels grans nombres de J. Bernoulli

Sigui $\{X_n, n \geq 1\}$ una successió de variables aleatòries independents, totes amb la distribució $B(1,p)$. Utilitzant el teorema de De Moivre-Laplace resulta que

$$\begin{aligned}\lim_{n\to\infty} P\left\{\left|\frac{S_n}{n} - p\right| < \varepsilon\right\} &= \lim_{n\to\infty} P\left\{-\varepsilon\sqrt{\frac{n}{pq}} < \frac{S_n - np}{\sqrt{npq}} < \varepsilon\sqrt{\frac{n}{pq}}\right\} \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \mathrm{e}^{-\frac{x^2}{2}}\,\mathrm{d}x = 1\end{aligned}$$

o, equivalentment,

$$\lim_{n\to\infty} P\left\{\left|\frac{S_n}{n} - p\right| > \varepsilon\right\} = 0.$$

6.3.2 Estimació de l'error comès en aproximar la freqüència relativa per la probabilitat

Amb les mateixes notacions que en 6.3.1 tenim que, per a n prou gran,

$$
\begin{aligned}
P\left\{\left|\frac{S_n}{n}-p\right|<\varepsilon\right\} &= P\left\{-\varepsilon\sqrt{\frac{n}{pq}}<\frac{S_n-np}{\sqrt{npq}}<\varepsilon\sqrt{\frac{n}{pq}}\right\} \\
&\sim \frac{1}{\sqrt{2\pi}}\int_{-\varepsilon\sqrt{\frac{n}{pq}}}^{\varepsilon\sqrt{\frac{n}{pq}}} \mathrm{e}^{-\frac{x^2}{2}}\,\mathrm{d}x = \frac{2}{\sqrt{2\pi}}\int_0^{\varepsilon\sqrt{\frac{n}{pq}}} \mathrm{e}^{-\frac{x^2}{2}}\,\mathrm{d}x.
\end{aligned}
$$

Hi ha taules que donen els valors de la funció $\phi(x)=\int_0^x \mathrm{e}^{-\frac{x^2}{2}}\,\mathrm{d}x$. Això permet d'obtenir un valor aproximat de la probabilitat de l'error.

Ens pot també interessar conèixer quin és el mínim nombre de proves que cal fer perquè, amb probabilitat no més petita que $\alpha\in(0,1)$, la freqüència relativa d'un esdeveniment no difereixi de la seva probabilitat més de ε. Pels càlculs que acabem de fer, podem respondre a aquesta pregunta determinant n tal que

$$
\frac{2}{\sqrt{2\pi}}\int_0^{\varepsilon\sqrt{\frac{n}{pq}}} \mathrm{e}^{-\frac{x^2}{2}}\,\mathrm{d}x \geq \alpha.
$$

6.3.3 Determinació del nombre d'enquestats necessaris per obtenir resultats fiables

Una aplicació possible de la qüestió tractada en l'apartat anterior és la de determinar el nombre d'enquestats en una consulta per obtenir resultats fiables.

Suposem que es fa una enquesta entre una població de votants per predir el percentatge de vots a favor d'un cert candidat. Suposem que hi ha una proporció desconeguda p que li dóna suport i que els diversos votants decideixen de forma independent. Volem determinar quin ha de ser el nombre d'enquestats perquè p es pugui predir amb error més petit que 0.04 i probabilitat superior a 0.95.

Sobre la població Ω de votants, definim variables aleatòries de la manera següent:

$$
X_i=\begin{cases} 1, & \text{si el } i\text{-èsim individu enquestat vota } A,\\ 0, & \text{en cas contrari.}\end{cases}
$$

Les variables aleatòries X_i tenen distribució $B(1,p)$ i poden suposar-se independents. La predicció del resultat de la votació es fa a partir de la mitjana aritmètica $\frac{S_n}{n}$. Cal, doncs, trobar n tal que

$$P\left\{\left|\frac{S_n}{n}-p\right|\leq 0.04\right\}\geq 0.95.$$

Equivalentment, si denotem per Φ la funció de distribució de la llei $\mathcal{N}(0,1)$, hem de trobar el n més petit que compleix que

$$\Phi\left(0.04\sqrt{\frac{n}{p(1-p)}}\right)\geq\frac{1+0.95}{2}. \tag{6.11}$$

Com que $p(1-p)\leq\frac{1}{4}$, si n compleix que

$$\Phi(0.08\sqrt{n})\geq 0.975,$$

també tindrem (6.11). Aquesta darrera condició dóna n de l'ordre de 600.

En la deducció de (6.11) hem utilitzat l'expressió $\Phi(x)-\Phi(-x)=2\Phi(x)-1$, que s'obté per la simetria de la llei normal estàndard.

6.3.4 Errors en les mesures físiques

Una característica física es mesura diverses vegades per donar un resultat ben acurat. Cada mesura està sotmesa a error i suposem que aquest es distribueix segons una llei uniforme en $[-1,1]$. Com a mesura definitiva es proposa la mitjana aritmètica de les diverses mesures realitzades. Podem estar interessats en qüestions com les que hem plantejat en els exemples anteriors. Per exemple, quina és la probabilitat que aquest valor proposat difereixi del veritable valor menys de $\delta\in(0,1)$.

Denotem per m el veritable valor i per $X_j,\quad j=1,\ldots,n$ els resultats obtinguts en la sèrie de mesures efectuades. Podem suposar que

$$X_j=m+\xi_j,\quad j=1,\ldots,n,$$

on $\{\xi_j,\quad j=1,\ldots,n\}$ són variables aleatòries independents, amb distribució $\mathcal{U}[-1,1]$.

És fàcil deduir que $EX_j = m$ i $\operatorname{Var} X_j = \frac{1}{3}$. Utilitzant el teorema del límit central, podem fer l'aproximació

$$P\left\{\left|\frac{X_1 + \cdots + X_n}{n} - m\right| \leq \delta\right\} \sim 2\Phi(\delta\sqrt{3n}) - 1.$$

Problemes relacionats amb la distribució dels errors de mesura a la Física i l'Astronomia foren àmpliament estudiats per Gauss i Laplace.

6.4 Apèndix

Aquest apartat es dedica a la demostració del resultat següent:

Lema 6.1 *Siguin X, Y dues variables aleatòries independents amb lleis $\mathcal{N}(m_1, \sigma_1^2)$, $\mathcal{N}(m_2, \sigma_2^2)$, respectivament. La llei de la variable aleatòria $X+Y$ és una $\mathcal{N}(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$.*

Aquest resultat ha estat utilitzat en la demostració del teorema 6.2, que era un ingredient essencial en la prova del teorema del límit central. La manera més ràpida de demostrar el lema és utilitzant funcions característiques. Aquí en donem una demostració directa, una mica feixuga, però que utilitza solament tècniques elementals d'integració.

Demostració del lema 6.1: Com que X i Y són independents, la funció de distribució de $X + Y$ és

$$\begin{aligned} &P\{X + Y \leq t\} \\ &\quad = \iint_{\{(x,y)\in\mathbb{R}^2 : x+y\leq t\}} \frac{1}{2\pi\sigma_1\sigma_2} \exp\left(-\left(\frac{(x-m_1)^2}{2\sigma_1^2} + \frac{(y-m_2)^2}{2\sigma_2^2}\right)\right) \mathrm{d}x\,\mathrm{d}y. \end{aligned}$$

Utilitzant la transformació bijectiva de $\mathbb{R}^2$ en $\mathbb{R}^2$ definida per

$$\begin{cases} u = x + y \\ v = y, \end{cases}$$

i fent el corresponent canvi de variable en la integral anterior, obtenim que

$$P\{X+Y \le t\}$$
$$= \int_{-\infty}^{t} \mathrm{d}u \int_{\mathbb{R}} \mathrm{d}v \frac{1}{2\pi\sigma_1\sigma_2} \exp\left(-\left(\frac{(u-v-m_1)^2}{2\sigma_1^2} + \frac{(v-m_2)^2}{2\sigma_2^2}\right)\right).$$

De la igualtat anterior es dedueix que $X+Y$ és una variable aleatòria absolutament contínua amb densitat

$$f_{X+Y}(u) = I(u),$$

on

$$I(u) = \int_{\mathbb{R}} \frac{1}{2\pi\sigma_1\sigma_2} \exp\left(-\left(\frac{(u-v-m_1)^2}{2\sigma_1^2} + \frac{(v-m_2)^2}{2\sigma_2^2}\right)\right) \mathrm{d}v.$$

El problema es redueix ara a fer manipulacions amb aquesta integral per tal de provar que

$$I(u) = \frac{1}{\sqrt{2\pi(\sigma_1^2+\sigma_2^2)}} \exp\left(-\frac{(u-(m_1+m_2))^2}{2(\sigma_1^2+\sigma_2^2)}\right). \tag{6.12}$$

La idea és transformar l'integrand de la integral $I(u)$ en una densitat gaussiana. Per a això, expressarem

$$\frac{(u-v-m_1)^2}{2\sigma_1^2} + \frac{(v-m_2)^2}{2\sigma_2^2} \tag{6.13}$$

com la suma de dos termes, un quadrat perfecte en v i un segon terme independent de v.

Més precisament, desenvolupant els quadrats en (6.13), hom obté que

$$\frac{(u-v-m_1)^2}{2\sigma_1^2} + \frac{(v-m_2)^2}{2\sigma_2^2} = A(v,u) + B(u),$$

amb

$$
\begin{aligned}
A(v,u) &= \left(v \left(\frac{\sigma_1^2+\sigma_2^2}{2\sigma_1^2\sigma_2^2} \right)^{\frac{1}{2}} + \frac{\sigma_2^2(m_1-u)-\sigma_1^2 m_2}{(\sigma_1^2+\sigma_2^2)^{\frac{1}{2}}(2\sigma_1^2\sigma_2^2)^{\frac{1}{2}}} \right)^2, \\
B(u) &= \frac{(\sigma_1^2+\sigma_2^2)(\sigma_2^2(m_1-u)^2+\sigma_1^2 m_2^2)-(\sigma_2^2(m_1-u)-\sigma_1^2 m_2)^2}{(\sigma_1^2+\sigma_2^2)2\sigma_1^2\sigma_2^2}.
\end{aligned}
$$

Resulta, doncs, que

$$
I(u) = \frac{1}{2\pi\sigma_1\sigma_2} \exp(-B(u)) \int_{\mathbb{R}} \exp(-A(u,v))\, \mathrm{d}v. \tag{6.14}
$$

Desenvolupant $B(u)$ és fàcil veure que

$$
B(u) = \frac{(u-(m_1+m_2))^2}{2(\sigma_1^2+\sigma_2^2)}. \tag{6.15}
$$

Definim la transformació de $\mathbb{R}$ en $\mathbb{R}$ donada per

$$
w = \sqrt{2}\sqrt{A(v,u)}.
$$

Aleshores,

$$
\begin{aligned}
\frac{1}{2\pi\sigma_1\sigma_2} \int_{\mathbb{R}} \exp(-A(u,v))\, \mathrm{d}v &= \frac{1}{\sqrt{2\pi(\sigma_1^2+\sigma_2^2)}} \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{w^2}{2}\right) \mathrm{d}w \\
&= \frac{1}{\sqrt{2\pi(\sigma_1^2+\sigma_2^2)}}.
\end{aligned} \tag{6.16}
$$

Les identitats (6.14)–(6.16) ens donen la validesa de la fórmula (6.12) i això és el que volíem demostrar. □

Exercicis

6.1 *Es suposa que el pes dels infants en néixer segueix una llei normal. En una determinada població el pes de les nenes és una llei* $\mathcal{N}(3.1, 0.3)$ *i la dels nens una* $\mathcal{N}(3.3, 0.5)$. *Escollim dos nadons a l'atzar, una nena i un nen. Quina és la probabilitat que el nen pesi més que la nena?*

6.2 *En un procés de producció de iogurts, es produeixen lots que tenen una caducitat, mesurada en dies, que es distribueix segons una llei* $\mathcal{N}(20, 4)$. *Es serveix un lot amb n iogurts. Determineu el valor mínim de n perquè, amb probabilitat* $0,95$, *la caducitat mitjana sigui superior a* 18 *dies.*

6.3 *En fer una operació aritmètica, una calculadora aproxima el resultat a l'enter més pròxim. Suposem que els errors d'aproximació són independents i es distribueixen segons una llei uniforme en l'interval* $[-0.5, 0.5]$. *Trobeu el nombre màxim d'operacions que es poden fer perquè, amb probabilitat més gran que* 0.9, *l'error total comès sigui, en valor absolut, inferior a* 10.

6.4 *En unes eleccions el candidat A va obtenir el* 45 *per cent dels vots. En el procés de recompte, es varen extreure d'una urna dos blocs de* 100 *paperetes cadascun. Determineu la probabilitat que la diferència de vots del candidat A entre els dos blocs fos més gran o igual al* 5 *per cent.*

Apèndix A

Simulació de variables aleatòries

La simulació de variables aleatòries o *simulació de Montecarlo* és una tècnica numèrica per generar distribucions de probabilitat a partir de nombres a l'atzar. El nom de *Montecarlo* fou introduït per Neumann i Ulam com a codi secret per al seu treball a Los Alamos durant la segona guerra mundial, que consistia en simulació directa del comportament de la difusió aleatòria de neutrons en material fissionable. Actualment, el mètode de Montecarlo s'aplica en l'estudi de molts problemes per als quals es veu difícil o impossible trobar una solució analítica.

A.1 Generadors de nombres aleatoris

Amb el nom de *nombres aleatoris* entenem observacions de variables aleatòries independents amb distribució uniforme en $[0, 1]$. Una qüestió natural que pot plantejar-se és la següent: donada una successió de nombres entre 0 i 1, quan acceptarem que són *nombres aleatoris*? Des del punt de vista matemàtic, aquesta és una qüestió difícil. L'any 1966, Martin-Löf va donar una resposta teòrica satisfactòria a aquesta pregunta. És, però, una resposta molt frustrant, car exclou tots aquells casos en què tals nombres són obtinguts mitjançant un algorisme. Malgrat això, és ben conegut que la majoria d'ordinadors tenen incorporats programes per generar nombres que podem qualificar de *pseudoaleatoris*. I el resultat és bo, en el sentit que els nombres obtinguts passen satisfactòriament tests estadístics d'uniformitat i d'independència.

Una forma lenta i dificultosa d'obtenir nombres aleatoris és mitjançant un mostreig amb reposició en una urna que conté boles numerades del 0 al 9. Una ruleta dividida en 10 sectors iguals pot resultar igualment útil. Però aquests no són els sistemes que s'utilitzen habitualment.

Els ordinadors produeixen, com hem dit abans, nombres *pseudoaleatoris*. Són nombres obtinguts de manera totalment determinista per fórmules recursives basades en el càlcul de residus mòdul un cert nombre enter.

Més precisament, fixem un valor inicial X_0 i enters no negatius a, c i m. Considerem la successió obtinguda recursivament mitjançant la fórmula

$$X_{i+1} = aX_i + c \pmod{m},$$

$i = 0, 1, \ldots, n$.

Aleshores, es prenen com a nombres pseudoaleatoris els obtinguts fent

$$U_i = \frac{X_i}{m}.$$

La seqüència obtinguda així és periòdica i conté com a màxim m nombres diferents. Per exemple, si $a = c = X_0 = 3$ i $m = 5$, obtenim $X_i = 3, 2, 4, 0, 3$. Quan un nombre es repeteix entrem en un bucle.

L'objectiu és triar m prou gran per assegurar una successió llarga de nombres diferents. Sigui p el període de la successió. Pot demostrar-se que $p = m$ si, i només si, es compleixen les tres condicions següents:

1. $\text{mcd}(c, m) = 1$.
2. $a \equiv 1 \mod g$, per a tot factor g primer de m.
3. $a \equiv 1 \mod 4$, si m és un múltiple de 4.

La demostració es basa en resultats d'aritmètica.

Aquest és un exemple de generadors pseudoaleatoris. Altres mètodes bastant similars a aquest es troben descrits en [R].

Alguns sistemes operatius, com ara el Linux, utilitzen esdeveniments externs aleatoris, com ara les polsacions del teclat, els accessos al disc, i el moviments del ratolí, per a poder generar nombres amb un grau d'aleatorietat major.

A.2 Mètodes generals per a la simulació de variables aleatòries

En aquesta secció presentem diversos mètodes per generar lleis de probabilitat.

A.2.1 Mètode de la transformació inversa

Sigui F una funció de distribució. En el capítol 4 vàrem definir la funció quantil o *pseudoinversa* de F de la manera següent:

$$Q(y) = \inf\{x : F(x) \geq y\},$$

$y \in [0, 1]$.

També vàrem observar que, a causa de l'equivalència expressada en (4.6), si U és una variable aleatòria amb distribució uniforme en $(0, 1)$, la variable aleatòria $X = Q(U)$ té F com a funció de distribució.

Amb aquest mètode podem generar variables aleatòries a partir de nombres aleatoris sempre que Q tingui una expressió analítica senzilla. Però aquest no és el cas en moltes de les distribucions usuals.

Exemple A.1 Generació d'una variable aleatòria X absolutament contínua amb densitat

$$f(x) = 2x\mathbb{1}_{[0,1]}(x).$$

La funció de distribució corresponent és

$$F(x) = \begin{cases} 0, & \text{si } x < 0 \\ \int_0^x 2y\,\mathrm{d}y = x^2, & \text{si } x \in [0, 1] \\ 1, & \text{si } x > 1 \end{cases}$$

Aleshores, $Q(U) = F^{-1}(U) = U^{\frac{1}{2}} = X$. Per tant, X pot generar-se a partir de l'arrel quadrada d'una variable aleatòria amb distribució uniforme en $[0, 1]$.

Exemple A.2 Simulació d'una variable aleatòria X amb llei $\exp(\lambda)$.

La funció de densitat és

$$f(x) = \lambda \mathrm{e}^{-\lambda x} \mathbb{1}_{[0,\infty)}(x),$$

i la funció de distribució,

$$F(x) = 1 - \mathrm{e}^{-\lambda x}.$$

La funció inversa és donada per

$$F^{-1}(y) = -\frac{1}{\lambda} \log(1-y).$$

Per tant, segons la discussió anterior, la variable aleatòria $X = -\frac{1}{\lambda}\log(1-U)$ té llei $\exp(\lambda)$. Observeu que $1-U$ té la mateixa llei que U. En conseqüència, tenim la igualtat següent en llei: $X = -\frac{1}{\lambda}\log U$.

Exemple A.3 Sigui X una variable aleatòria discreta que pren els valors $a_1, a_2, \ldots, a_n, \ldots$ amb probabilitats $p_1, p_2, \ldots, p_n, \ldots$, respectivament. La funció de distribució de X és com la donada en la figura 2.3. La corresponent funció quantil és donada per

$$\begin{aligned} Q(y) = {} & a_1 \mathbb{1}_{\{0 \le y \le p_1\}} + a_2 \mathbb{1}_{\{p_1 < y \le p_1+p_2\}} + \cdots \\ & + a_n \mathbb{1}_{\{p_1+\cdots+p_{n-1} < y \le p_1+\cdots+p_{n-1}+p_n\}} + \cdots . \end{aligned}$$

Considerem una variable aleatòria amb distribució uniforme en $[0,1]$. Es compleix la igualtat següent (en llei):

$$\begin{aligned} X = {} & a_1 \mathbb{1}_{\{U \le p_1\}} + a_2 \mathbb{1}_{\{p_1 < U \le p_1+p_2\}} + \cdots \\ & + a_n \mathbb{1}_{\{p_1+\cdots+p_{n-1} < U \le p_1+\cdots+p_{n-1}+p_n\}} + \cdots . \end{aligned}$$

A.2.2 Mètode de composició

Com el seu nom indica, aquest mètode consisteix a generar variables aleatòries a partir de la composició o integració de lleis conegudes. L'explicarem mitjançant un exemple.

Sigui X una variable aleatòria absolutament contínua amb densitat

$$f(x) = \frac{1 - \mathrm{e}^{-ax}(ax+1)}{ax^2} \mathbb{1}_{\{x>0\}}.$$

Mitjançant una integració per parts és fàcil veure que

$$f(x) = \frac{1}{a} \int_0^a y\mathrm{e}^{-yx}\,\mathrm{d}y.$$

A la vista d'aquesta expressió es dedueix que $f(x)$ és la densitat marginal d'una parella (X, L) amb densitat conjunta

$$\frac{1}{a} \mathbb{1}_{[0,a]}(y) y\mathrm{e}^{-yx}.$$

Així, doncs, X pot simular-se a partir d'una variable aleatòria amb llei $\mathcal{U}([0,a])$, $a > 0$ i una variable aleatòria amb distribució $\exp(y)$, $y > 0$.

Aquest mètode s'aplica també a densitats de la forma

$$f(x) = \int f(x,y)g(y)\,\mathrm{d}y,$$

on g i $f(\cdot, y)$ són funcions de densitat per a tot y.

També pot utilitzar-se en el cas discret, és a dir, si

$$f(x) = \sum_{k=1}^{\infty} a_k g_k(x),$$

on $a_k \geq 0$, $\sum_{k\geq 1} a_k = 1$ i g_k són funcions de densitat per a tot $k \geq 1$.

A.2.3 Mètode d'acceptació-rebuig

Considerem una funció de densitat f contínua que té suport compacte contingut en $[a, b]$. Denotem per k una fita de $\sup_x f(x)$. Generem un punt aleatori $P = (U, V)$ amb distribució uniforme en el rectangle $[a, b] \times [0, k]$. Si el punt P està situat en el gràfic de f, que anomenarem A, l'acceptem i posem $X = U$. En cas contrari, el rebutgem i tornem a generar un altre punt P independent del primer i amb la mateixa distribució.

Volem establir els fets següents:

1. Hi haurà únicament un nombre finit de proves abans que es produeixi l'esdeveniment $\{P \in A\}$.

2. La variable aleatòria X així definida té per funció de densitat f.

En efecte, sigui $\{P_n, n \geq 1\}$ una successió de vectors aleatoris independents amb distribució uniforme en el rectangle $[a, b] \times [0, k]$. La probabilitat de l'esdeveniment $\cap_{n\geq 1}\{P_n \notin A\}$ és

$$\lim_{n\to\infty} \left(\frac{k(b-a) - |A|}{k(b-a)} \right)^n = 0,$$

on $|A|$ designa l'àrea de A. Per tant,

$$P\Big\{ \bigcup_{n\geq 1} \{P_n \in A\} \Big\} = 1,$$

la qual cosa demostra (1).
Per demostrar (2), observem que la variable aleatòria X es pot definir de la manera següent

$$\begin{array}{ll}
X = U_1, & \text{si } P_1 \in A \\
X = U_2, & \text{si } P_1 \notin A, P_2 \in A \\
\cdots & \\
X = U_n, & \text{si } P_1 \notin A, P_2 \notin A, \dots, P_{n-1} \notin A, P_n \in A \\
\cdots &
\end{array}$$

Considerem la successió de conjunts

$$B_1 = \{P_1 \in A\}, \dots, B_n = \{P_1 \notin A, \dots, P_{n-1} \notin A, P_n \in A\}, \dots,$$

que formen una partició del conjunt $C = \Omega \cap (\cap_{n\geq 1}\{P_n \notin A\})^c$. Hem demostrat abans que $P(C) = 1$.
Aleshores, donats $c, d \in \mathbb{R}$, $c < d$, tenim que

$$
\begin{aligned}
P\{c \le X \le d\} &= \sum_{n\ge 1} P\{c \le X \le d, B_n\} \\
&= \sum_{n\ge 1} P\{c \le U_n \le d, B_n\} \\
&= \sum_{n\ge 1} \big(P\{P_1 \notin A\}\big)^{n-1} P\{P_n \in [c,d]\times[0,k]\cap A\} \\
&= \frac{\int_c^d f(t)\,\mathrm{d}t}{\left(1-\frac{k(b-a)-|A|}{k(b-a)}\right)k(b-a)} = \frac{\int_c^d f(t)\,\mathrm{d}t}{|A|}.
\end{aligned}
$$

Observeu que $|A| = \int_a^b f(t)\,\mathrm{d}t = 1$. Això acaba la demostració de (2).

A.3 Simulació d'algunes variables aleatòries particulars

A.3.1 Simulació de la distribució binomial

Es pot utilitzar el mètode de l'exemple A.3, però també pot fer-se utilitzant l'observació següent: Sigui X una variable aleatòria amb distribució $B(n,p)$. Considerem n variables aleatòries independents amb distribució uniforme en $[0,1]$. La llei de X coincideix amb la de la variable aleatòria $\sum_{i=1}^n \mathbb{1}_{\{U_i \le p\}}$.

A.3.2 Simulació de la llei de Poisson

Sigui $\{E_n, n \ge 1\}$ una successió de variables aleatòries independents amb llei $\exp(\lambda)$, $\lambda > 0$. Un càlcul estàndard amb densitats conjuntes permet de demostrar que

$$P\{E_1 + \cdots + E_n \le 1 < E_1 + \cdots + E_{n+1}\} = \exp(-\lambda)\frac{\lambda^n}{n!}.$$

D'aquí es dedueix la igualtat següent (en llei)

$$X = \sum_{n=1}^{\infty} n \mathbb{1}_{\{E_1+\cdots+E_n \le 1 < E_1+\cdots+E_{n+1}\}}.$$

Simulem les variables aleatòries $\{E_n, n \geq 1\}$ segons hem descrit en l'exemple A.2. La comparació $1 < E_1 + \cdots + E_n$ equival a $\exp(-\lambda) > \prod_{i=1}^{n} U_i$.

A.3.3 Simulació de la llei $\mathcal{N}(0,1)$

Malgrat la seva àmplia utilització, la llei normal no és fàcil de simular. Aquest és un exemple típic en què el mètode de la transformació inversa no dóna cap resultat satisfactori, a causa de la complexitat de la funció de distribució.

La manera usual de simular la llei normal és mitjançant un canvi de variable tal com expliquem a continuació.

Sigui R una variable aleatòria tal que R^2 segueix una llei exponencial amb paràmetre $\lambda = \frac{1}{2}$ i sigui ζ una variable aleatòria amb distribució uniforme en $[0, 2\pi]$ independent de R. Utilitzant el mètode del canvi de variable (vegeu la secció 2.9) es pot demostrar que, si

$$X_1 = R\cos\zeta, \quad X_2 = R\sin\zeta,$$

el vector aleatori $X = (X_1, X_2)$ té una funció de densitat

$$f(x,y) = \frac{1}{2\pi} \mathrm{e}^{-\frac{x^2+y^2}{2}}.$$

En altres paraules, X té llei normal bidimensional i els seus components són variables aleatòries amb llei $\mathcal{N}(0,1)$.

Com que sabem simular lleis uniformes en qualsevol interval i també lleis exponencials, de la discussió anterior resulta que, si U_1 i U_2 són dues variables aleatòries independents amb distribució $\mathcal{U}[0,1]$, la variable aleatòria

$$X = (-2\log U_1)^{\frac{1}{2}} \cos 2\pi U_2$$

té llei $\mathcal{N}(0,1)$.

Bibliografia

[A] ASH, R. B., *Basic Probability Theory*. John Wiley and Sons, Inc., New York, London, Sydney, Toronto, 1970.

[Ba] BALDI, P., *Calcolo delle probabilità e statistica*. McGraw-Hill Libri Italia srl., Milano, 1992.

[Bo] BOULEAU, N., *Probabilités de l'ingénieur: Variables aléatoires et simulation*. Hermann, Paris, 1986.

[Br] BRÉMAUD, P., *Introduction aux Probabilités*. Springer-Verlag, Berlin, Heidelberg, 1984.

[C] CERDÀ, J., *Análisis real. Col·lecció U.B., 23*, Edicions Universitat de Barcelona, Barcelona, 1996.

[Ch 1] CHUNG, K.L., *Teoría elemental de la probabilidad y de los procesos estocásticos*. Ed. Reverté, Barcelona, 1983.

[Ch 2] CHUNG, K.L., *A Course in Probability Theory*. Academic Press, New York, 1974.

[K] KARR, A.F., *Probability. Springer Texts in Statistics*, Springer-Verlag, New York, 1993.

[N-S] NUALART, D. SANZ SOLÉ, M., *Curs de Probabilitats. Colección Estadística y Análisis de datos*, PPU, Barcelona, 1990.

[Pf] PFEIFFER, P. E., *Probability for Applications. Springer Texts in Statistics*, Springer-Verlag, New York, 1990.

[P] PITMAN, J., *Probability. Springer Texts in Statistics*, Springer-Verlag, New York, 1993.

[R] RUBINSTEIN, R., *Simulation and the Monte Carlo Method*. Wiley, New York, 1981.

[S] SANZ SOLÉ, M., *Lliçons de càlcul de probabilitats. Textos docents, 9*, Edicions Universitat de Barcelona, Barcelona, 1997.

Índex alfabètic