UNIVERSITAT DE BARCELONA
Facultat de Matemàtiques i Enginyeria Informàtica

APUNTS DE CLASSE

Grau en Matemàtiques

Curs 2023-2024 / Setè Semestre

Probabilitats (PROB)

Autor:
Mario VILAR

Professor/a:
Dr. David MÁRQUEZ

PRESENTACIÓ DE L'ASSIGNATURA

Aborda els fonaments de la teoria de la probabilitat, incloent espais de probabilitat i variables aleatòries. A més, es tracten conceptes avançats com vectors aleatoris i successions de variables aleatòries. L'assignatura també explora la Llei dels Grans Nombres i el Teorema del Límit Central, que són fonamentals en l'anàlisi probabilístic i estadístic. *Revisats per Alex CAÑAS.*

CLASSIFICACIÓ AMS (2020): 03B48, 60-00, 60-01, 60A05, 60B12, 60F05.

Índex

Introducció

Tuvimos, hombre, tiempo para que nuestra sed fuera saciándose, el ancestral deseo de enumerar las cosas y sumarlas, de reducirlas hasta hacerlas polvo, arenales de números. Fuimos empapelando el mundo con números y nombres, pero las cosas existían, se fugaban del número, enloquecían en sus cantidades, se evaporaban dejando su olor o su recuerdo y quedaban los números vacíos.

Pablo NERUDA, *Oda a los números*

Primer de tot, trobareu que hi ha un índex, on hi distingim els diferents apartats ordenats seguint el meu propi criteri i, de tant en tant, seguint l'ordre cronològic del curs. Hi ha capítols, seccions, subseccions (i fins i tot subsubseccions). Us faig cinc cèntims de com he organitzat els encapçalaments de cada pàgina:

1. el número de l'últim capítol/secció/subsecció, depèn de la profunditat que hi hagi definida en aquell moment, figurarà en cada cantonada superior de pàgina parella (per exemple, *1.2*);
2. el nom del capítol es trobarà a la part dreta de la capçalera de les pàgines parelles (per exemple, «Divisibilitat i nombres primers»);
3. el nom de l'última secció/subsecció de la pàgina, a la cantonada dreta superior de les pàgines parelles (per exemple, «Polinomis: algorisme d'Euclides»);
4. el número de l'últim teorema, definició... de la pàgina en qüestió es trobarà a les pàgines senars, a la cantonada superior dreta, *destacat en el color de la seva capçalera corresponent* (per exemple, **1.2.3**).

A més, hi ha una taula, la taula de continguts. En aquest sentit, tal com acabem de dir, es veu fàcilment que s'ha seguit una mena de *sorting-by-color* per poder treballar de manera més eficient amb els diferents tipus d'enunciats matemàtics. D'aquesta manera, si busqueu una definició, un teorema... podreu distingir que estan destacats amb colors diferents (ara els introduïm) i trobar-los molt ràpidament:

1. Teoremes, proposicions, lemes, corol·laris, propietats, conjectures, processos i exercicis tindran aquest format (capçalera destacada amb color gris fosc):

   Teorema. *Compte! L'enunciat del teorema també serà en cursiva! Jove xef, porti whisky amb quinze glaçons d'hidrogen, coi!*

2. Les definicions i notacions tindran aquest format (capçalera de color gris clar):

   Definició. Aqueix betzol, Jan, comprava whisky de figa.

3. Les remarques i exemples tindran aquest format:

 Observació. Zel de grum: quetxup, whisky, cafè, bon vi; ja!

Després de molts anys fent-ho malament, ara sol quedaran numerades aquelles equacions a les quals em referiré més endavant. És la *filosofia dominant* en la majoria de textos matemàtics (té nom i tot, es diu la *regla d'Occam*).

Per últim, m'estalviaré de comentar l'índex terminològic perquè el seu propòsit és clar i, en efecte, paral·lel al de l'organització d'aquest document: poder facilitar-vos al màxim la feina per localitzar qualsevol concepte que desitgeu. Espero que us serveixin d'alguna cosa aquests apunts, els he fet amb tot l'amor del món. Sort!

Mario Vilar
Sitges, Barcelona
21 de gener de 2024

Taula de continguts

CAPÍTOL 1

CAPÍTOL 2

I

Espai de probabilitat

1.1

Definició i propietats de la probabilitat

Definició 1.1.1 (Espai de probabilitat). Un espai de probabilitat és una terna $(\Omega, \mathcal{F}, P)$, on Ω és l'espai mostral (el conjunt de possibles resultats), $\mathcal{F}$ és una família de parts d'Ω que té estructura de σ-àlgebra[1]; és a dir:

1. $\Omega \in \mathcal{F}$;
2. $\mathcal{F}$ és estable per pas al complementari: si $A \in \mathcal{F}$, també $A^C \in \mathcal{F}$;
3. $\mathcal{F}$ és estable per unions numerables: si $\{A_n \mid n \geq 1\} \subset \mathcal{F}$, es compleix: $\bigcup_{n\geq 1} A_n \in \mathcal{F}$.

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

$$\begin{array}{rccc} P: & \mathcal{A} & \longrightarrow & [0,1] \\ & A & \longmapsto & P(A) \end{array}$$

que té les següents propietats:

1. $P(\Omega) = 1$.
2. (*o de σ-additivitat*). Si $\{A_n\}_{n\geq 1}$ és una successió de conjunts d'$\mathcal{F}$ disjunts dos a dos (recordem, $A_i \cap A_j = \emptyset, i \neq j$), aleshores:

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1\infty} P(A_n).$$

Propietat 1.1.2 (Propietats de la probabilitat P).

1. $P(\emptyset) = 0$.
2. *La propietat de σ-additivitat implica l'additivitat finita.*
3. *Per a tot $A \in \mathcal{F}$, $P(A^c) = 1 - P(A)$.*
4. *Si $A, B \in \mathcal{F}$, aleshores:*

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

5. *(de monotonia). Siguin $A, B \in \mathcal{F}$. Si $A \subseteq B$, es dona $P(A) \leq P(B)$.*
6. *Tota probabilitat és subadditiva; és a dir, per a una família de conjunts $\{A_i\}_{1\leq i\leq n}$ de $\mathcal{F}$:*

$$P\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} P(A_i).$$

[1] Aquesta σ-àlgebra $\mathcal{F}$ serveix per descriure tots els *esdeveniments* possibles relacionats amb l'experiència aleatòria.

7. *Per a una família de conjunts $\{A_i\}_{1\leq i\leq n}$ de $\mathcal{F}$ es compleix el següent:*

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{\substack{i=1\\ i\neq j}}^{n} P(A_i\cap A_j) + \sum_{i\neq j\neq k} P(A_i\cap A_j\cap A_k) + \cdots + (-1)^{n-1} P(A_1\cap\cdots\cap A_n).$$

Demostració.

1. En efecte, prenem $A_n = \emptyset$ per a tot $n \geq 1$, de manera que $\bigcup_{n=1}^{\infty} A_n = \emptyset$. Tindrem:

$$P(\emptyset) = P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n) = \sum_{n=1}^{\infty} P(\emptyset) \implies P(\emptyset) = 0.$$

2. Siguin $A, B \in \mathcal{F}, A\cap B = \emptyset$.
3. 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 l'apartat anterior, tenim que

$$P\left(\bigcup_{j=1}^{n} A_j\right) = P\left(\left(\cup_{j=1}^{n} A_j\right) \cup (\emptyset)\right) = \sum_{j=1}^{n} P(A_j) + \sum P(\emptyset) = \sum_{j=1}^{n} P(A_j)$$

4. Sigui $A \in \mathcal{F}$. Aquesta igualtat és conseqüència de la descomposició $\Omega = A \cup A^c$ i la propietat d'additivitat finita. En efecte, $P(\Omega) = P(A\cup A^c) = P(A) + P(A^c)$ ja que $A\cap A^c = \emptyset$.
5. Si $A, B \in \mathcal{F}$, podem escriure la fórmula de Grassmann perquè $P(A) = P(A\cap B) + P(A\cap B^c)$. Si sumem $P(B)$ a banda i banda, obtenim:

$$P(A) + P(B) = P(A\cap B) + P(A\cap B^c) + P(B)$$

   El resultat se segueix de la igualtat $(A\cap B^c)\cup B = A\cup B$.
6. Usem el resultat de teoria de conjunts: $B = A\cup(B\cap A^c)$, per a A, B qualssevol. Tenim:

$$P(B) = P(A) + P(B\cap A^c) \geq P(A).$$

7. La demostració d'aquesta propietat es fa per inducció, no la farem.
8. Tampoc la farem. ■

Proposició 1.1.3. *Sigui $\{A_n\}_{n\geq 1}$ una successió de conjunts de $\mathcal{F}$. Aleshores:*

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

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

2. *Anàlogament, si la successió és decreixent i denotem per A el conjunt $\bigcap_{n=1}^{\infty} A_n$, es compleix que:*

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

Demostració. Prenem la successió $B_n = A_n \cap A_{n-1}^c \forall n \geq 1$, i tal que $A_0 = \emptyset$. Adonem-nos que $B_n \in \mathcal{F}$ i $B_n \cap B_k = \emptyset$ per a $n \neq k$. A més:

$$\bigcup_{k=1}^{n} B_k = A_n = \bigcup_{k=1}^{n} A_k \quad \bigcup_{k=1}^{\infty} A_k = \bigcup_{k=1}^{\infty} B_k.$$

Aleshores, per la propietat de σ-additivitat de P, tenim que:

$$P(A) = P\left(\bigcup_{n=1}^{\infty} A_n\right) = P\left(\bigcup_{n=1}^{\infty} B_n\right) = \sum_{n=1}^{\infty} P(B_n) = \lim_{n\to\infty} \sum_{i=1}^{n} P(B_i) = \lim_{n\to\infty} P(A_n).$$

Per a la segona part, la successió $\{A_n^c\}_{n\geq 1}$ és creixent i $\bigcup_{i=1}^{\infty} A_n^c = A^c$, aleshores, pel primer apartat:

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

1.2

Espais de probabilitats finits

Si el conjunt és finit, l'espai de probablitat $(\Omega, \mathcal{F}, P)$ s'anomena *espai de probabilitat finit.*

Definició 1.2.1 (Partició). Una família de conjunts d'$\mathcal{F}$, $\{A_1, \ldots, A_n\}$, és una partició del conjunt Ω si són disjunts dos a dos i $\bigcup_{i=1}^{n} A_i = \Omega$.

Teorema 1.2.2. *Tota àlgebra finita $\mathcal{F}$ de parts de Ω és l'àlgebra generada per una partició finita.*

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

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

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

$$P(A) = \frac{\#A}{\#\Omega} = \frac{\text{casos favorables}}{\text{casos totals}},$$

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

1.3

Combinatòria i binomials

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 $(n_1, \ldots, n_k)$, on $n_j \in \{1, \ldots, n\}, j = 1, \ldots, k$. El resultat és n^k. Aquest nombre és també el cardinal del conjunt de les aplicacions $f : K \longrightarrow N$.

   Definició 1.3.1 (Muestreo ordenado con reposición). Hacemos k extracciones de las m bolas de manera ordenada retornando las bolas que hemos encontrado en la urna. Por tanto, hemos de suponer $k \leq m$. Se trata de un *muestreo ordenado con reposición.* El conjunto de resultados es, entonces:

$$\{(a_1, \ldots, a_k) \bigm| a_i \in \{1, 2, \ldots, m\}\}.$$

   Ahora, en cada una de las elecciones tenemos m posibilidades. Aplicando la regla del producto podemos extraer las bolas de $VR_m^k = m^k$ maneras diferentes. VR_m^k quiere decir *variaciones con repetición de m elementos agrupados de k en k.*

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).

   Definició 1.3.2 (Muestreo ordenado sin reposición). Hacemos k extracciones de las m bolas de manera ordenada y sin retornar las bolas que hemos encontrado en la urna. Por tanto, hemos de suponer $k \leq m$. Se trata de un *muestreo ordenado sin reposición.* El conjunto de resultados es, entonces:

$$\{(a_1, \ldots, a_k) \bigm| a_i \in \{1, \ldots, m\}, a_i \neq a_j, i \neq j\}.$$

Siguiendo por el mismo caso, en la primera extracción tendremos m posibilidades; en la segunda, $m-1$ y así sucesivamente. Luego, podemos extraer los elementos de

$$V_m^k = m \cdot (m-1) \cdots (m-(k-1)) = \frac{m!}{(m-k)!}$$

formas distintas. V_m^k quiere decir *variaciones de m elementos agrupados de k en k.*

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:

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

Definició 1.3.3 (Mostreo sin orden ni reposición**).** Ahora, hagamos k extracciones de las m bolas sin tener en cuenta el orden y sin retornar las bolas a la urna. Podemos considerar, en efecto, *que sacamos todas las k bolas a la vez.* Como no hay reposición hemos de suponer $m \geq k$. El conjunto de resultados será:

$$\{\{a_1, \ldots, a_k\} \mid a_i \in \{1, \ldots, m\}, a_i \neq a_j, i \neq j\}.$$

El número de conjuntos de k elementos seleccionados de un conjunto de m elementos sin importar el orden y sin reposición, es:

$$C_m^k = \frac{V_m^k}{k!} = \frac{m}{k!(m-k)!} = \binom{m}{k}.$$

C_m^k indica las combinaciones de m elementos cogidos de k en k y usaremos, normalmente, la notación de los números combinatorios.

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: $\binom{n-k_1}{k_2}, \ldots$ 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 $\binom{n+k-1}{n}$, 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.

Definició 1.3.4 (Binomial). Para $m \geq k$ se define el número combinatorio $\binom{m}{k}$ como el número natural:

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

Estos números tienen una serie de propiedades la mayoría de las cuales son triviales. Escribimos unas pocas que no lo son tanto a simple vista:

1. $\binom{m}{k} = \binom{m}{m-k}$.
2. $\binom{m}{k} + \binom{m}{k+1} = \binom{m+1}{k+1}$.
3. $\sum_{k=0}^{m} \binom{m}{k} = 2^m$.

Observació 1.3.5. La quinta propiedad nos permite definir el triángulo de Pascal. Cada número es la suma de los dos que tiene directamente encima. Además, por la última propiedad anunciada tenemos que los elementos de la fila m-ésima suman 2^m.

Definició 1.3.6 (Binomio de Newton). Dados x, y reales y m natural, entonces:

$$(x+y)^m = \sum_{k=0}^{m} \binom{m}{k} x^{m-k} y^k.$$

$$(x-y)^m = \sum_{k=0}^{m} (-1)^k \binom{m}{k} x^{m-k} y^k.$$

1.4
Probabilitat condicionada

Definició 1.4.1 (Probabilitat condicionada). La probabilitat d'un conjunt $A \in \mathcal{F}$ condicionada per B es defineix per:

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

Proposició 1.4.2 (Fórmula de les probabilitats compostes). *Siguin $A_1, \ldots, A_n$ elements d'$\mathcal{F}$ tal que $P(A_1 \cap \cdots \cap A_{n-1}) > 0$. Aleshores:*

$$P(A_1 \cap \cdots \cap A_{n-1} \cap A_n) = P(A_1) \cdot P(A_2|A_1)P(A_1) \cdot P(A_3|A_1 \cap A_2) \cdots P(A_n|A_1 \cap A_{n-1}). \quad (1.1)$$

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:

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

i apliquem la hipòtesi d'inducció, de manera que queda (1.1), com volíem. ∎

Proposició 1.4.3 (Fórmula de les probabilitats totals). *Considerem una partició finita del conjunt Ω, $\mathcal{P} = \{A_1, \ldots, A_n\} \subset \mathcal{F}$, on $P(A_i) > 0$, i $1 \leq i \leq n$. Donat un esdeveniment $A \in \mathcal{F}$, tenim que:*

$$P(A) = \sum_{i=1}^{n} P(A \cap A_i) = \sum_{i=1}^{n} P(A|A_i)P(A_i). \quad (1.2)$$

Demostració. Com que $A = A \cap \Omega$, i $\Omega = \bigcup_{i=1}^{n} A_i$:

$$A \cap \left(\bigcup_{i=1}^{n} A_i \right) = \bigcup_{i=1}^{n} (A \cap A_i),$$

la propietat d'addivitat de les probabilitats i la definció de probabilitat condicionada impliquen (1.2). ∎

1.5
Esdeveniments independents

Definició 1.5.1 (Esdeveniments independents). Dos esdeveniments $A, B \in \mathcal{F}$ són independents si $P(A \cap B) = P(A)P(B)$; equivalentment, si $P(A) > 0$, $P(B|A) = P(B)$. L'esdeveniment A no condiciona l'esdeveniment de B de cap manera.

Proposició 1.5.2.

1. *$\emptyset$ i Ω són independents a qualsevol esdeveniment.*
2. *Un esdeveniment és independent a sí mateix si, i només si $P(A) = 0$ o $P(A) = 1$.*
3. *A i B són independents si, i només si A^c és independent a B si, i només si, A és independent a B^c si, i només si, A^c és independent a B^c.*

Demostració.

1. Evident.
2. Ja que $P(A) = P(A \cap A) = P(A)P(A)$, i tenim dues possibilitats: $P(A) = 0$ o bé $P(A) = 1$.
3. En fem alguna de la cadena d'equivalències, la resta es fa de manera anàloga:

$$P(A^c \cap B) + P(A \cap B) = P(B) \iff P(A^c \cap B) = P(B) - P(A)P(B) = P(B)P(A^c). \blacksquare$$

Ara anem a fer una generalització.

Definició 1.5.3. Un conjunt d'esdeveniments $\{A_i \mid i \in I\} \subset \mathcal{F}$ són dits independents si per a qualsevol subconjunt finit $\{A_{i_1}, \dots, A_{i_n}\}$, se satisfà que:

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

Aquesta definició d'independència és més forta que la independència entre cada parella d'esdeveniments del conjunt $\{A_i\}_{i \in I}$.

Observació 1.5.4. Per exemple, en el cas de tres esdeveniments A_1, A_2 i A_3, pot ocórrer que

$$P(A_1 \cap A_2) = P(A_1) P(A_2),$$
$$P(A_1 \cap A_3) = P(A_1) P(A_3),$$
$$P(A_2 \cap A_3) = P(A_2) P(A_3),$$

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.

II

Variables i vectors aleatoris

2.1

Variables aleatòries

Definició 2.1.1 (Àlgebra de Borel)**.** La σ-àlgebra de Borel 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. La denotarem per $\mathcal{B}$.

Definició 2.1.2 (Variable aleatòria)**.** Una variable aleatòria és una aplicació $X : \Omega \longrightarrow \mathbb{R}$ que és *mesurable*; és a dir, $\forall B \in \mathcal{B}$ es dona que $X^{-1}(B) = \{\omega \in \Omega \mid X(\omega) \in B\} \in \mathcal{F}$. Una variable aleatòria proporciona, doncs, una assignació numèrica als elements de l'espai mostral. Equivalentment, per a tot $t \in \mathbb{R}$:

$$\begin{aligned} X^{-1}\left((-\infty, t]\right) &\in \mathcal{F} \\ X^{-1}\left((-\infty, t)\right) &\in \mathcal{F} \\ X^{-1}\left((t, \infty)\right) &\in \mathcal{F} \\ X^{-1}\left([t, \infty)\right) &\in \mathcal{F} \end{aligned}$$

Definició 2.1.3 (Llei d'una variable aleatòria)**.** Sigui $B \in \mathcal{B}$. La llei d'una variable aleatòria $X : (\Omega, \mathcal{F}) \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ és la probabilitat associada:

$$P_X(B) = P(X^{-1}(B)) = P \circ X^{-1}(B) = P(X \in B) = P(\{\omega \in \Omega \mid X(\omega) \in B\}).$$

És clar que $P_X(\mathbb{R}) = 1$, de manera que és una probabilitat. Finalment, sigui $\{B_n\}_{n\geq 1}$ una col·lecció de conjunts disjunts de B. Aleshores:

$$X^{-1}\left(\bigcup_{n\geq 1} B_n\right) = \bigcup_{n\geq 1} X^{-1}(B_n).$$

Exemple 2.1.4 (Funció indicatriu)**.** Sigui $A \in \mathcal{F}$. Definim $\mathbb{1}_A : \Omega \longrightarrow \mathbb{R}$ per:

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

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. En aquest cas, $X^{-1}(B) = \emptyset$.
2. B conté el 0 i l'1. En aquest cas, $X^{-1}(B) = \Omega$.
3. B conté el 0 però no l'1. En aquest cas, $X^{-1}(B) = A^c$.

4. B conté el 1 però no el 0. En aquest cas, $X^{-1}(B) = A$.

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

Proposició 2.1.5. *Siguin X, Y dues variables aleatòries definida 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 és una variable aleatòria.*

Demostració. Només hem de saber com es fa el segon, que se segueix de les igualtats:

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

I amb això ja hem acabat. ■

Proposició 2.1.6. *Sigui $\{X_n\}_n$ una successió de variables aleatòries definides en el mateix espai de probabilitat $(\Omega, \mathcal{F}, P)$. Aleshores, $\inf X_n$, $\sup X_n$, $\underline{\lim} X_n$ i $\overline{\lim} X_n$ són també variables aleatòries.*

Definició 2.1.7 (Variable aleatòria elemental). Siguin $a_1, \ldots, a_n \in \mathbb{R}$ i $A_1, \ldots, A_n$ és una partició d'Ω. Definim la variable aleatòria elemental com:

$$X(\omega) = \sum_{i=1}^{n} a_i \mathbb{1}_{A_i}(\omega), \; A_1 \cup \cdots \cup A_n = \Omega. \tag{2.1}$$

La seva esperança (cf. 3) és $E(X) = \sum_{i=1}^{r} a_i P(A_i)$.

Proposició 2.1.8. *Sigui $X : \Omega \longrightarrow [0, +\infty)$ una variable aleatòria no negativa. Aleshores, existeix una successió de variables aleatòries simples i positives, $(x_n)_n$, $n \geq 1$, tal que convergeix cap a X.*

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

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

Com que X és una variable simple 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 $\omega \in \Omega$ 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 que $|X(\omega) - X_n(\omega)| \leq 2^{-n}$, que tendeix a zero quan $n \to \infty$. ■

Notació 2.1.9. Veiem tot seguit un conjunt de notacions que farem servir contínuament en el context de variables aleatòries:

1. $\{X = a\} := X^{-1}(\omega \in \Omega \mid X(\omega) = a)$.
2. $\{X \leq a\} = \{\omega \in \Omega \mid X(\omega) \leq a\}$.
3. $\{a < X \leq b\} := \{X \leq b\} \cap \{a < X\} = \{\omega \in \Omega \mid a < X(\omega) \leq b\}$.

2.2 Variables aleatòries canòniques

Ara introduirem una sèrie d'exemples importants, canònics si es vol, de variables aleatòries, des de 2.2.4 fins a 2.2.7, són totes variables aleatòries *discretes*.

Definició 2.2.1 (Variable aleatòria discreta). Una variable aleatòria X direm que és discreta si la seva llei està concentrada en un conjunt numerable $B_0 \in \mathcal{B}$.

Observació 2.2.2. Les variables aleatòries simples (cf. l'equació (2.1)) 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,

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

Definició 2.2.3 (Funció de probabilitat). 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 2.2.4 (Variable aleatòria de Bernoulli). Considerem una experiència aleatòria amb dos resultats possibles a_1, a_2 amb probabilitats respectives p i $q = 1 - p$, $p \in (0,1)$. Associem a aquesta experiència la variable aleatòria definida per $X(a_1) = 1$ i $X(a_2) = 0$. La llei d'aquesta variable aleatòria és una probabilitat concentrada en els punts 1,0 amb intensitats p, q, respectivament. S'anomena *llei de Bernoulli* de paràmetre p i es denota per $B(1, p)$. Formalment:

$$\begin{array}{rccc} X: & \{a_1, a_2\} & \longrightarrow & \{0,1\} \\ & \{a_1\} & \longmapsto & 1 \\ & \{a_0\} & \longmapsto & 0 \end{array} \qquad F(X) = \begin{cases} 0, & \text{si } x < 0; \\ 1-p, & \text{si } 0 \leq x < 1; \\ 1, & \text{si } x \geq 1. \end{cases}$$

Aleshores, $P_X(\{1\}) = P(X = 1) = p$ i $P_X(\{0\}) = P(X = 0) = q$. L'esperança és $E(X) = 1 \cdot P(X = 1) + 0 \cdot P(X = 0) = p$.

Exemple 2.2.5 (Variable aleatòria binomial)**.** Considerem aquí la repetició, n vegades *de manera independent*, d'una experiència aleatòria com la descrita en l'exemple 2.2.4. Ens interessa estudiar el nombre de vegades que hem obtingut el resultat a_1. L'espai mostral és

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

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

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

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

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

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 $\sum_{k=0}^{n} Q(\{k\}) = 1$. La llei de la variable aleatòria X s'anomena llei binomial amb paràmetres (n, p) i es denota per $B(n, p)$.

Pel que fa a l'esperança, tenim:

$$\begin{aligned} E(X) &= \sum_{k=0}^{n} kP(X = k) = \sum_{k=0}^{n} k \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} = \sum_{k=1}^{n} \frac{n!}{(k-1)! \cdot (n-k)!} p^k (1-p)^{n-k} \\ &= np \sum_{k=1}^{n} \frac{(n-1)!}{k-1)! \cdot (n-k)!} p^{k-1} (1-p)^{n-k} \overset{\ell=k-1}{=} \sum_{\ell=0}^{n-1} \frac{(n-1)!}{\ell! \cdot (n-1-\ell)!} p^{\ell} (1-p)^{n-1-\ell} \\ &= np(p + 1 - p)^{n-1} = np. \end{aligned}$$

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.

Exemple 2.2.6 (Variable aleatòria geomètrica)**.** Considerem una successió de repeticions independents d'una experiència aleatòria com la de l'exemple 2.2.4. Ens aturem la primera vegada que obtenim el resultat a_1 (*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:

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

La llei de X és una probabilitat Q concentrada en $\{1, 2, \dots\}$ tal que $P(X = k) = Q(\{k\}) = p(1-p)^{k-1}$. S'anomena *llei geomètrica de paràmetre p*. Observem que $\sum_{k=1}^{\infty} Q(\{k\}) = 1$ car, pels resultats de la suma de la sèrie geomètrica:

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

Pel que fa l'esperança (cf. secció 3):

$$E(X) = \sum_{k=1}^{\infty} k \cdot P(X = k) = p\sum_{k=1}^{\infty} k(1-p)^{k-1} = p\frac{1}{(1-(1-p))^2} = \frac{1}{p}.$$

En la penúltima igualtat hem usat que si tenim la sèrie $f(x) = \sum_{k=0}^{\infty} x^k = \frac{1}{1-x}$, $|x| < 1$ i f derivable, tenim que $f'(x) = \sum_{k=1}^{\infty} kx^{k-1} = \frac{1}{(1-x)^2}$.

Exemple 2.2.7 (Variable aleatòria de Poisson)**.** És la probabilitat de tenir èxit amb un sol experiment. Tenim $P(X = k) = \frac{\lambda^k}{k!}e^{-\lambda}$, $k = 0, 1, 2, \dots$ i $\lambda > 0$. En efecte, la suma de totes les probabilitats ens dona 1:

$$\sum_{k=0}^{\infty} e^{-\lambda}\frac{\lambda^k}{k!} \overset{\text{Taylor}}{=} e^{-\lambda}e^{\lambda} = 1.$$

A més, l'esperança és:

$$E(X) = \sum_{k=1}^{\infty} k \cdot P(X = k) = e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^k}{(k-1)!} = e^{-\lambda}\lambda\sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!} \overset{\text{Taylor}}{=} e^{-\lambda} \cdot \lambda \cdot e^{\lambda} = \lambda.$$

Matemàticament, la distribució de Poisson definida 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 $\lim_n np_n = \lambda > 0$. Observem que això implica que $\lim_n p_n = 0$. Aleshores, per a tot nombre natural $k \geq 0$ es té que:

$$\begin{aligned}\lim_n B(n, p_n)(\{k\}) &= \lim_n \binom{n}{k} p_n^k(1-p_n)^{n-k} = \lim_n \frac{n!}{k!(n-k)!}p_n^k(1-p_n)^{n-k} \\ &= \lim_n \frac{(np_n)^k}{k!} \cdot \frac{n!}{n^k(n-k)!}\left((1-p_n)^{\frac{1}{p_n}}\right)^{np_n}(1-p_n)^{-k} = \frac{\lambda^k}{k!}e^{-\lambda},\end{aligned}$$

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

2.3

Funcions de distribució

Definició 2.3.1 (Funcions de distribució). La funció de distribució associada a una variable aleatòria X és la funció $F : \mathbb{R} \longrightarrow [0,1]$ definida per:

$$\begin{array}{rccl} F_X : & \mathbb{R} & \longrightarrow & [0,1] \\ & x & \longmapsto & F_X := P(X \leq x) = P \circ X^{-1}((-\infty, x]). \end{array}$$

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

Proposició 2.3.2. *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, és a dir, $\lim_{x_n \to x^+} F(x_n) = F(x)$.*
3. *$\lim_{x \to +\infty} F(x) = 1$ i $\lim_{x \to -\infty} F(x) = 0$.*

La implicació contrària també és certa (però no ho demostrarem).

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 el mateix que demostrar la continuïtat per successions. Fixem $x \in \mathbb{R}$ i considerem una successió decreixent de nombres reals, $(x_n)_n$ tal que $\lim_n x_n = x$. La successió de conjunts $\{X \leq x_n\}$ decreix i $\bigcap_{n=1}^{\infty} \{X \leq x_n\} = \{X \leq x\}$. Aleshores, resulta que:
$$\lim_{n \to \infty} F(x_n) = F(x).$$
3. Considerem una successió creixent de nombres reals $(x_n)_n$ tal que $\lim_n x_n = +\infty$. Considerem també la successió de conjunts associada, $\{X \leq x_n\}$ que creix cap a $\{X < \infty\} = \Omega$. Aleshores, tenim 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, $(x_n)_n$ tal que $\lim_{n \to \infty} x_n = -\infty$. En aquest cas, els conjunts $\{X \leq x_n\}$ formen una successió decreixent cap a $\emptyset$. Per tant, novament:
$$\lim_{x \to \infty} F(x) = P(\emptyset) = 0.$$ ■

Observació 2.3.3. 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.*

Demostració. Sigui $A_c = \{x \in \mathbb{R}, F_X(x^+) - F_X(x^-) > \varepsilon\}$ i $\varepsilon = \frac{1}{n}$, de manera que $\#A_{\frac{1}{n}} \leq n$. Suposem que $\#A_{\frac{1}{n}} > n$, aleshores, existeixen $x_1, \ldots, x_{n+1}$ tals que $F_X(x_i) - F_X(x_i^-) > \frac{1}{n}$.

$$\begin{aligned} F_X(x_{n+1}) &= F_X(x_{n+1}) - F_X(x_{n+1}^-) + F_X(x_{n+1}^-) \geqslant F_X(x_{n+1}) - F_X(x_{n+1}^-) + F_X(x_n) \\ &\geqslant \sum_{k=1}^{n} F_X(x_{k+1}) - F_X(x_{k+1}^-) + F_X(x_1) > 1 \end{aligned}$$

la qual cosa és una contradicció, i ja hem acabat. ■

Propietat 2.3.4.

1. $P(X < a) = F(a^-)$,
2. $P(a \leq X \leq b) = F(b) - F(a^-)$,
3. $P(a < X \leq b) = F(b) - F(a)$,
4. $P(a \leq X < b) = F(b^-) - F(a^-)$,
5. $P(a < X < b) = F(b^-) - F(a)$.
6. $P(X = a) = F(a) - F(a^-) \geq 0$.

Demostració. Farem la demostració de les dues primeres identitats. Sigui $(a_n)_n$ una successió de nombres reals creixent cap a a. Aleshores, els conjunts $\{X \leq a_n\}$ formen una successió creixent cap a $\{X < a\}$. Sigui $A = \bigcup_n \{X \leq a_n\} \subset \{X \leq a\}$. Aleshores, per a tot $\omega \in A$ existeix n_0 tal que $X \leq a_n < a$ i això implica que $\omega \in \{X < a\}$. Per tant, per a tot $\omega \in \{X < a\}$ existeix n_0 tal que $X(\omega) \leq a_{n_0}$ i $\omega \in A$: hem demostrat l'altra inclusió i ja hem acabat.

Alternativament, tenim:

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

2.4 Variables aleatòries absolutament contínues

Les variables aleatòries que tenen una funció de distribució contínua, anomenades variables aleatòries contínues, compleixen que $P(\{X = x\}) = 0$, per a tot $x \in \mathbb{R}$. En aquest apartat descriurem un subconjunt d'aquestes variables aleatòries contínues.

Definició 2.4.1 (Densitat). Una funció $f : \mathbb{R} \longrightarrow \mathbb{R}$ s'anomena *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 $\int_{-\infty}^{+\infty} f(x)\,dx = 1$.

Definició 2.4.2 (Variable aleatòria absolutament contínua)**.** Es diu que una variable aleatòria X és absolutament contínua amb densitat f si la seva funció de distribució F es pot escriure com:

$$F(x) = \int_{-\infty}^{x} f(y)\,dy,$$

per a tot $x \in \mathbb{R}$, on la funció f satisfà les condicions de la definició 2.4.1.

Proposició 2.4.3 (Conseqüències del Teorema Fonamental del Càlcul)**.** *Suposem que la funció de distribució F compleix:*

1. *És contínua a 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.

Observació 2.4.4. Tenim:

$$\begin{aligned} P(a \leq X \leq b) = P(a < X \leq b) = P(a \leq X < b) = P(a < X < b) &= \int_{-\infty}^{+\infty} f(x)\mathbb{1}_{[a,b]}(x)\,dx \\ &= \int_{a}^{b} f(x)\,dx\,. \end{aligned}$$

Exemple 2.4.5 (Llei uniforme)**.** Siguin $a, b \in \mathbb{R}$ de manera que X té llei $U(a,b)$ (això és, $X \sim U(a,b)$). Definim:

$$f(x) = \frac{1}{b-a}\mathbb{1}_{[a,b]}(x) \geq 0. \tag{2.2}$$

És evident que (2.2) defineix una funció de densitat. Notem que:

$$P(X = x) = \int_{-\infty}^{x} f(y)\,dy = \int_{-\infty}^{x} \frac{1}{b-a}\mathbb{1}_{[a,b]}(y)\,dy = \int_{a}^{x} \frac{1}{b-a}\,dy = \frac{x-a}{b-a}.$$

La funció de distribució associada és, doncs:

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

També tenim:

$$P(X \leqslant x) = \int_{-\infty}^{x} f(y)\,dy = \int_{-\infty}^{x} \frac{1}{b-a}\mathbb{1}_{[a,b]}\,dy = \int_{a}^{x} \frac{1}{b-a}\,dy$$

L'esperança és
$$E(X) = \int_{-\infty}^{\infty} xf(x)\,dx = \int_a^b x\frac{1}{b-a}\,dx = \frac{b+a}{2}.$$
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 $\mathcal{U}[a,b]$.

Exemple 2.4.6 (Llei exponencial)**.** Fixem un $\lambda > 0$, $\lambda \in \mathbb{R}$. La llei *exponencial* amb paràmetre λ, $\exp(\lambda) \equiv e^{\lambda}$, correspon a la funció de densitat següent:

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

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 nul·la.
2. No hi ha envelliment. Més precisament, suposem que:

$$P\{X \geq x + y | X \geq x\} = P\{X \geq y\}, \forall x, y \geq 0.$$

   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, $\varphi(x+y) = \varphi(x)\varphi(y)$. Tenint en compte que $\varphi(0) = 1$, resulta que $\varphi(x) = \exp(-\lambda x)$, $\lambda > 0$ car $\lim_{n\to\infty} \varphi(x) = 0$.

Exemple 2.4.7 (Llei Gamma)**.** Fixem $p \in (0, +\infty)$. Es defineix la funció gamma mitjançant la integral.
$$\Gamma(p) = \int_0^{\infty} x^{p-1} e^{-x}\,dx, \tag{2.3}$$
en particular, $\Gamma(1) = 1$ i $\Gamma(\frac{1}{2}) = \sqrt{\pi}$. Fixem ara $p \in (0, +\infty)$ i $\lambda \in (0, +\infty)$. Definim, al seu torn:
$$f(x) = \begin{cases} \frac{\lambda^p}{\Gamma(p)} x^{p-1} e^{-\lambda x}, & \text{si } x \geq 0; \\ 0, & \text{en cas contrari.} \end{cases}$$
Fent un canvi de variables i utilitzant (2.3) és fàcil veure que $f(x)$ defineix una funció de densitat. La llei corresponent s'anomena la *llei gamma* amb paràmetres (p, λ). La llei exponencial amb paràmetre λ es pot també descriure com una llei $\Gamma(1, \lambda)$. Pel que fa a l'esperança (cf. secció 3):

$$J = \int_0^{\infty} x \frac{\lambda^p}{\Gamma(p)} x^{p-1} e^{-\lambda x}\,dx = \frac{\Gamma(p+1)}{\lambda\Gamma(p)}$$

Exemple 2.4.8 (Llei normal estàndard). Considerem la funció $g(x) = e^{-\frac{x^2}{2}}$, amb $x \in \mathbb{R}$. 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,

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}. \tag{2.4}$$

Una variable aleatòria amb distribució normal estàndard és la que té per funció de densitat la donada en (2.4). Denotarem la llei normal estàndard per $\mathcal{N}(0,1)$. L'esperança (cf. secció 3) d'X és finita i $E(X) = 0$, ja que:

$$\int_{-\infty}^{+\infty} |x| \cdot e^{-\frac{x^2}{2}}\, dx \overset{1}{=} 2 \int_{0}^{+\infty} x e^{-\frac{x^2}{2}}\, dx = 2.$$

A partir d'aquí, és clar que $\int_{-\infty}^{+\infty} x \cdot e^{-\frac{x^2}{2}}\, dx = 0$ per la paritat de l'integrand.

2.5

Densitat de transformacions de variables aleatòries

Sigui X una variable aleatòria que pren valors en un interval obert I finit o no. Sigui $g : I \longrightarrow J$ una transformació bijectiva i de classe $\mathcal{C}^1$ 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 $f_X(x)$. La funció g 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 $\mathcal{C}^1$ en J, per a tot $y \in J$,

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

Proposició 2.5.1. *Y té una densitat donada per:*

$$f_Y(y) = f_X(g^{-1}(y)) \cdot |(g^{-1})'(y)| = f_X(g^{-1}(y)) \frac{1}{|g'(g^{-1}(y))|}, y \in J. \tag{2.5}$$

Demostració. Suposem que g és creixent i que $I = (a,b)$, $J = (c,d)$, de manera que tindrem:

$$F_Y(y) = P(\{Y \leq y\}) = P(\{X \leq g^{-1}(y)\}) = \int_a^{g^{-1}(y)} f_X(x)\, dx = \int_c^y f_X(g^{-1}(z))(g^{-1})'(z)\, dz,$$

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

[1] Hem usat que $|x| \cdot e^{-\frac{x^2}{2}}$ és parella.

Exemple 2.5.2 (Canvi de variable a una variable aleatòria normal). 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 $g : \mathbb{R} \longrightarrow \mathbb{R}$, $g(x) = \sigma x + m$. Podem, doncs, aplicar la fórmula (2.5) i obtenim:

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

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 $\mathcal{N}(m, \sigma^2)$.

2.6 Vectors aleatoris

Definició 2.6.1 (Vector aleatori). Un vector aleatori m-dimensional és una aplicació $X : \Omega \longrightarrow \mathbb{R}^m$, $X = (X_1, \ldots, X_m)$ tal que cadascun dels components $X_i : \Omega \longrightarrow \mathbb{R}$ és una variable aleatòria.

1. Per a definir la llei d'un vector aleatori cal introduir una σ-àlgebra associada a $\mathbb{R}^n$. La llei d'un vector X és una probabilitat sobre $(\mathbb{R}^n, B \in \mathbb{R}^n)$ definida per a tot $B \in \mathcal{B}(\mathbb{R}^n)$ de la manera següent: $P_X(B) = P(X^{-1}(B)) = P(\omega \in \Omega \mid X(\omega) = B)$.
2. La funció de distribució d'un vector X és:

$$\begin{array}{rccl} F : & \mathbb{R}^n & \longrightarrow & [0,1] \\ & x & \longmapsto & P(X = x) = P(X_1 \leq x_1, \ldots, X_n \leq x_n) = P(\{X_1 \leq x_1\} \cap \cdots \cap \{X_n \leq x_n\}). \end{array}$$

Teorema 2.6.2. *La funció de distribució d'un vector aleatori compleix les següents propietats:*

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

$$\lim_{y \downarrow x} F(y) = F(x),\ y_i \downarrow x_i,\ \forall i = 1, \ldots, m.$$

3. *F té les propietats asimptòtiques següents:*

$$\lim_{x \to \infty} F(x) = 1,\ \lim_{x_i \to -\infty} F(x) = 0,\ \forall i = 1, \ldots, m.$$

4. *Funció de distribució marginal i-èsima. Siguin $a, b \in \mathbb{R}^m$, $a \leq b$. Definim:*

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

tal que $a = (a_1, \ldots, a_m)$ i $b = (b_1, \ldots, b_m)$. Aleshores, $\Delta_{a,b} F = P(\{X \in (a,b]\})$, on $(a,b]$ denota el rectangle de $\mathbb{R}^m$ definit per:

$$(a,b] = \{x \in \mathbb{R}^m \mid a_i < x_i \leq b_i, i = 1, \ldots, m\} \implies \Delta_{a,b} F \geq 0.$$

Podem demostrar l'expressió $\Delta_{a,b}F = P(\{X \in (a,b]\})$ per a $m = 2$. En efecte:

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

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 2.6.3. *Per a tot $i = 1, \ldots, m$,*

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

Demostració. Quan $x_j \to \infty$, $j \neq i$, els conjunts $\{X_1 \leq x_1, \ldots, X_m \leq x_m\}$ creixen cap al conjunt $\{X_i \leq x_i\}$. La propietat de continuïtat seqüencial 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 2.6.4 (Vector discret). Un vector aleatori X s'anomena discret si les variables aleatòries components són discretes, això és, tenen la seva llei concentrada en un conjunt numerable $B_0 \in \mathcal{B}$.

Observació 2.6.5 (Probability mass function, cas discret i bidimensional). If discrete random variables X and Y are defined on the same sample space S, then their joint probability mass function (joint pmf) is given by $p(x,y) = P(X = x, Y = y)$, where (x,y) is a pair of possible values for the pair of random variables (X,Y), and $p(x,y)$ satisfies the following conditions:

1. $0 \leq p(x,y) \leq 1$.
2. $\sum_x \sum_y p(x,y) = 1$.
3. $P((X,Y) \in A) = \sum_x \sum_y p(x,y)$, $(x,y) \in A$.

In the discrete case, we can obtain the joint cumulative distribution function (joint cdf) of X and Y by summing the joint pmf:

$$F(x,y) = P(X \leq x, Y \leq y) = \sum_{x_i \leq x} \sum_{y_j \leq y} p(x_i, y_j).$$

Exemple 2.6.6 (Llei multinomial). Considerem una experiència aleatòria amb m resultats possibles $A_1, \ldots, A_m$ de probabilitats $p_1, \ldots, p_m$, respectivament, $\sum_{j=1}^m p_j = 1$. Repetim l'experiència n vegades de manera independent i denotem per X_i, $i = 1, \ldots, m$ la variable aleatòria que dona el nombre de vegades que s'ha produït el resultat A_i. Cadascuna de les variables aleatòries X_i té llei $B(n, p_i)$. Considerem el vector aleatori discret $X = (X_1, \ldots, X_m)$ que pren valors en el conjunt:

$$\mathcal{C} = \left\{ (n_1, \ldots, n_m) \;\middle|\; n_i \in \mathbb{N}, i = 1, \ldots, m, \sum_{i=1}^{m} n_i = n \right\},$$

Al seu torn, si $(n_1, \ldots, n_m) \in \mathcal{C}$, tenim:

$$P\{X_1 = n_1, \ldots, X_m = n_m\} = \frac{n!}{n_1! \cdots n_m!} p_1^{n_1} \cdots p_m^{n_m}.$$

Es diu que el vector aleatori X té *llei multinomial* amb paràmetres $n, p_1, \ldots, p_m$ i escriurem $M(n; p_1, \ldots, p_m)$. Observem que per a $m = 2$ tenim $B(n, p_1)$.

Definició 2.6.7 (Densitat, d'un vector)**.** 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
$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x)\, dx$$
en el sentit de Riemann, *i*
$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x)\, dx = 1$$

Definició 2.6.8 (Vector absolutament continu)**.** 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

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

on f és una funció de densitat en $\mathbb{R}^m$, és a dir, compleix les condicions de la definició de densitat d'un vector, 2.6.7, i $x = (x_1, \ldots, x_m)$. 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 2.6.9. *Sigui X un vector aleatori amb llei absolutament contínua. Aleshores, cadascuna de les variables aleatòries components X_i, $i = 1, \ldots, m$, són també absolutament contínues i les seves densitats respectives, que denotarem per f_i (densitats marginals), s'expressen com:*

$$f_i(y) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_m)\, dx_1 \cdots dx_{i-1}\, dx_{i+1} \cdots dx_m, \tag{2.6}$$

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

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

$$\begin{aligned} P(\{X_i \le x\}) = P(\{X_i \le x; X_j \in \mathbb{R}, i \neq j\}) &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{x} \cdots \int_{-\infty}^{+\infty} f(y)\, dy \\ &\overset{2}{=} \int_{-\infty}^{x_i} \left(\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(y_1, \dots, y_m)\, dy_1 \cdots dy_{i-1}\, dy_{i+1} \cdots dy_m \right) dy_i\,. \end{aligned}$$

D'on resulta que la variable aleatòria X_i té una densitat donada per (2.6); per tant, definint $\tilde{f}$ com correspon (solament depèn de y_i):

$$\int_{-\infty}^{x_i} \left(\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(y_1, \dots, y_m)\, dy_1 \cdots dy_{i-1}\, dy_{i+1} \cdots dy_m \right) dy_i = \int_{-\infty}^{x_i} \tilde{f}(y_i)\, dy_i,$$

tenim $\int_{-\infty}^{x_i} \tilde{f}(y_i)\, dy_i$ i X_i és absolutament contínua amb densitat $\tilde{f}$. ■

Aquestes densitats $\tilde{f}$ s'anomenen densitats marginals i no determinen la densitat del vector aleatori, f.

Observació 2.6.10 (Densitats, vector bidimensional)**.** If continuous random variables X and Y are defined on the same sample space S, then their joint probability density function (joint pdf) is a piecewise continuous function, denoted $f(x, y)$, that satisfies the following.

1. $f(x, y) \ge 0$, $(x, y) \in \mathbb{R}^2$.
2. $\iint_{\mathbb{R}^2} f(x, y)\, dx\, dy = 1$.
3. $P((X, Y) \in A) = \iint_A f(x, y)\, dx\, dy$ per a cada $A \subset \mathbb{R}^2$.

The joint cdf for continuous random variables X and Y is obtained by integrating the joint density function over a set A of the form $A = \{(x, y) \in \mathbb{R}^2 \mid X \le a, Y \le b\}$, where a and b are constants. Specifically, if A is given as above, then the joint cdf of X and Y, at the point (a, b), is given by:

$$F(a, b) = P(X \le a, Y \le b) = \int_{-\infty}^{b} \int_{-\infty}^{a} f(x, y)\, dx\, dy\,.$$

Suppose that continuous random variables X and Y have joint density function $f(x, y)$. The marginal pdf's of X and Y are respectively given by the following.

$$f_X(x) = \int_{\mathbb{R}} f(x, y)\, dy, \;\; f_Y(y) = \int_{\mathbb{R}} f(x, y)\, dx\,.$$

Exemple 2.6.11. Sigui X i $f_{(X,X)}(x, y)$ que suposem absolutament contínua. Volem demostrar que tenint les dues marginals ($f_{(X,X)} \equiv f_X, f_X$), no ens implica la conjunta. Usant:

$$P((X, Y) \subset A) = \iint_A f_{(X,Y)}(x, y)\, dx\, dy,$$

[2] Podem treure x_i aplicant Tonelli: com totes les components són positives, podem fer servir Tonelli i reordenem les integrals.

veiem:

$$1 = P(X = x) = \iint_{\{x=y\}} f_{(X,X)}(x,y)\,dx\,dy = 0.$$

Que evidentment és una contradicció, que ve de suposar que és absolutament contínua.

Exemple 2.6.12. Densitats diferents poden donar densitats marginals idèntiques. En efecte, considerem les densitats en el pla definides per:

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

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

Tenim:

$$f_X(x) = \int_{-\infty}^{+\infty} f_{(X,Y)}(x,y)\,dy = \frac{1}{4}\int_{-1}^{1}(1+xy)\,dy = \frac{1}{4}\left(y + x\frac{y^2}{2}\right)_{-1}^{1} = \frac{1}{4}\Big(1 + \frac{x}{2} - \frac{x}{2} - (-1)\Big) = \frac{1}{2}.$$

Per tant, $f_{(X,Y)}(x,y) = \frac{1}{4}\mathbb{1}_{(-1,1)}(x)\mathbb{1}_{(-1,1)}(y)$ i:

$$f_Y(x) = \int_{-\infty}^{+\infty} \frac{1}{4}\mathbb{1}_{(-1,1)}(x)\cdot\mathbb{1}_{(-1,1)}(y)\,dy = \frac{1}{2}\mathbb{1}_{(-1,1)}(x).$$

Exemple 2.6.13 (Llei uniforme en un recinte de $\mathbb{R}^2$)**.** Considerem el quadrat unitat en el pla, $T = [0,1]^2$. Definim $f(x,y) = \mathbb{1}_T(x,y)$. És obvi que aquesta 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

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

$|R|$ 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 2.6.14 (Llei normal bidimensional).

$$f_{(X,Y)}(x,y) = \frac{1}{2\pi} e^{-\frac{1}{2}(x^2+y^2)}, \ x, y \in \mathbb{R}.$$

En efecte, podem calcular $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x,y)\,dx\,dy = 1$. Les densitats marginals corresponen a lleis $\mathcal{N}(0,1)$.

2.7 Canvi de variable

Sigui X un vector aleatori m-dimensional i $g : \mathbb{R}^m \longrightarrow \mathbb{R}^d$ una funció contínua. Aleshores, $Y = g(X)$ és un vector aleatori d-dimensional. Suposem que X té llei absolutament contínua i considerem el cas $m = d$: U, V dos oberts de $\mathbb{R}^m$ i $g : U \longrightarrow V$ és un difeomorfisme. Aleshores, el vector aleatori Y té llei absolutament contínua i la seva densitat és:

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

$J_{g^{-1}}$ és el jacobià de la transformació g^{-1}.

Exemple 2.7.1. Sigui $(X,Y) = f_{(X,Y)}(x,y) = \mathbb{1}_{(0,1)}(x) \cdot \mathbb{1}_{(0,1)}(y)$ que val 1 quan $0 \leq x, y \leq 1$ i 0 en cas contrari. Posem $Z = X + Y$, i $g(X,Y) = X + Y$:

$$\begin{array}{rccc} g: & (0,1)\times(0,1) & \longrightarrow & D \\ & (x,y) & \longmapsto & (x+y, y) \\ & (z-t, t) & \longleftarrow & (z,t) \end{array}$$

Per tant, tenim el domini $D = \{(z,t) \in (0,2)\times(0,1) \mid t < z < t+1\}$. O, equivalentment, $0 < z - t < 1$.

$$f_{(Z,T)}(z,t) = f_{(X,Y)}(z-t,t)|J_{g^{-1}}(z,t)| = \mathbb{1}_{(0,1)\times(0,1)}(z-t,t)\cdot 1 = \mathbb{1}_D(z,t).$$

Aquest 1, és resultat de calcular el jacobià:

$$|J_{g^{-1}}(z,t)| = \left|\det\begin{pmatrix} \frac{\partial g_1^{-1}}{\partial z} & \frac{\partial g_2^{-1}}{\partial z} \\ \frac{\partial g_1^{-1}}{\partial t} & \frac{\partial g_2^{-1}}{\partial t} \end{pmatrix}\right| = \left|\det\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}\right| = 1.$$

Ens falta la densitat:

$$f_Z(z) = \int_{-\infty}^{+\infty} f_{(Z,T)}(z,t)\,dt = \begin{cases} 0, & \text{si } z \notin (0,2) \\ \int_0^z 1\,dt = z\mathbb{1}_{[0,1]}(z), & \text{si } z \in (0,1) \\ \int_{z-1}^1 1\,dt = (2-z)\mathbb{1}_{[0,1]}(z) & \text{si } z \in (1,2) \end{cases}$$

$$\Longrightarrow f_Z(z) = z\mathbb{1}_{[0,1]}(z) + (2-z)\mathbb{1}_{[0,1]}(z).$$

2.8 Variables aleatòries independents

Definició 2.8.1 (Variables aleatòries independents). Siguin $X_1, \ldots, X_m$ variables aleatòries definides en un espai de probabilitat $(\Omega, \mathcal{F}, P)$. Direm que són *independents* si, per a conjunts qualssevol $B_1, \ldots, B_m \in \mathcal{B}$ es compleix que:

$$P_1(\{X_1 \in B_1, \ldots, X_m \in B_m\} = \prod_{i=1}^{m} P(\{X_i \in B_i\});$$

en altres paraules, si els esdeveniments $\{X_1 \in B_1\}, \ldots, \{X_m \in B_m\}$ són independents. Un conjunt $\{X_n \mid n \geq 1\}$ infinit té esdeveniments independents si qualsevol subcol·lecció finita ho és.

Proposició 2.8.2. *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:*

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

Demostració.

$\boxed{\Rightarrow}$ És evident, perquè prenent $B_i = \{x_i\}$, $P(X_i \in B_i) = P(X_i \in \{x_i\}) = P(X_i = x_i)$.

$\boxed{\Leftarrow}$ Recíprocament, siguin $B_1, \ldots, B_m \in \mathcal{B}$. Aleshores,

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

Per tant, $P(\{X_1 \in B_1, \ldots, X_m \in B_m\}) = \prod_{i=1}^{m} P(\{X_i \in B_i\})$. ■

Teorema 2.8.3. *Considerem un vector aleatori $X = (X_1, \ldots, X_m)$. Denotem per F la seva funció de distribució i per F_i, $i = 1, \ldots, m$ les funcions de distribució respectives de les variables aleatòries components. Suposem que les variables aleatòries $X_1, \ldots, X_m$ són independents. Aleshores, es compleix que $F(x_1, \ldots, x_m) = F_1(x_1) \cdots F_m(x_m)$, per a tot $x = (x_1, \ldots, x_m) \in \mathbb{R}^m$. El recíproc també és cert.*

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

$$P(\{X_1 \in B_1, \ldots, X_m \in B_m\}) = P(\{X_1 \in B_1\}) \cdots P(\{X_m \in B_m\}).$$

En particular, prenent $B_i = (-\infty, x_i]$ per a tot $1 \leq i \leq m$, obtenim que:

$$P(\{X_1 \in B_1\}) \cdots P(\{X_m \in B_m\}) = \prod_{i=1}^{n} P(\{x_i \in (-\infty, x_i]\}) = \prod_{i=1}^{n} F_{x_i}(x).$$ ■

Proposició 2.8.4. *Siguin $X_1, \dots, X_m$ variables aleatòries. Suposem que el vector aleatori $X = (X_1, \dots, X_m)$ té llei absolutament contínua amb densitat f i densitats marginals $f_1, \dots, f_m$. Suposem que es compleix:*

$$f(x_1, \dots, x_m) = f_1(x_1) \cdots f_m(x_m).$$

Aleshores, les variables $X_1, \dots, X_m$ 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, usant $f(x_1, \dots, x_m) = f_1(x_1) \cdots f_m(x_m)$ tenim:

$$\begin{aligned} F(y) &= \int_{-\infty}^{y_1} \cdots \int_{-\infty}^{y_m} f(x_1, \dots, x_m)\, dx_1\, dx_m \\ &= \int_{-\infty}^{y_1} f_1(x_1)\, dx_1 \cdots \int_{-\infty}^{y_m} f_m(x_m)\, dx_m = F_1(y_1) \cdots F_m(y_m). \end{aligned}$$

Per tant, es compleix que $F(x_1, \dots, x_m) = F_1(x_1) \cdots F_m(x_m)$ i les variables aleatòries $X_1, \dots, X_m$ són, en efecte, independents. ■

Proposició 2.8.5. *Considerem un vector aleatori $X = (X_1, \dots, X_m)$ 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 $f(x_1, \dots, x_m) = f_1(x_1) \cdots f_m(x_m)$.*

Demostració. Sigui F la funció de distribució del vector aleatori X. Com estem suposant la independència dels components, podem implicar que $F(x_1, \dots, x_m) = F_1(x_1) \cdots F_m(x_m)$. Així doncs, resulta:

$$F(x_1, \dots, x_m) = \int_{-\infty}^{x_1} f_1(y)\, dy \cdots \int_{-\infty}^{x_m} f_m(y)\, dy = \int_{-\infty}^{x_1} \int_{-\infty}^{x_m} f_1(y_1) \cdots f_m(y_m)\, dy_1 \cdots dy_m$$ ■

2.9 Distribucions condicionades

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 2.9.1 (Llei de X condicionada). La llei de X condicionada a $\{Y = y\}$ és la probabilitat sobre $\mathbb{R}$ definida per:

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

Si les variables aleatòries X, Y són independents, les lleis condicionades coincideixen amb les respectives marginals.

Observació 2.9.2. La llei de X condicionada a Y i la llei de Y determinen la llei de X, això és conseqüència directa de la fórmula habitual,

$$P(\{X=x\}) = \sum_{\{y \mid p_Y(y)>0\}} p(x,y) = \sum_y \overline{p}_X(x|Y=y)p_Y(y).$$

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

Exemple 2.9.3. 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 $n \in \{0, 1, \ldots, n_1+n_2\}$ i $0 \le k \le n$. Aleshores:

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

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

III

Esperança matemàtica

3.1 Construcció

Volem definir una noció de mitjana dels valors presos per una variable aleatòria. 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.

Definició 3.1.1 (Esperança matemàtica, v.a. simples). Sigui X una variable aleatòria simple elemental, $X = \sum_{i=1}^{n} a_i \mathbb{1}_{A_i}(w)$, A_i disjunts dos a dos, $a_i \in \mathbb{R}$. L'esperança és la mitjana ponderada dels x:

$$E(X) = \sum_{i=1}^{n} a_i P(A_i) = \sum_{i=1}^{n} a_i P(X = a_i) = \sum_{i} a_i \cdot f(a_i). \tag{3.1}$$

La definició no ha de dependre de la variable simple escollida; és a dir:

$$X = \sum_{i=1}^{n} a_i \mathbb{1}_{A_i} = \sum_{i=1}^{n} a_i' \mathbb{1}_{A_i'}.$$

Proposició 3.1.2. *Siguin X, 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.

1. 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:

$$E(aX + bY) = \sum_{i=1}^{r} \sum_{j=1}^{s} (aa_i + bb_j) \mathbb{1}_{A_i \cap B_j}$$

i és, doncs, simple. Per tant:

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

2. Com $Y - X$ és una variable aleatòria simple i positiva, és obvi, a partir de la definició 3.1.3, que l'esperança ha de ser positiva. ■

Definició 3.1.3 (Esperança d'una variable aleatòria no negativa). L'esperança d'una variable aleatòria no negativa és $E(X) = \lim_n E(X_n)$, on $\{X_n\}_{n \geq 1}$ és una successió de v.a. simples i positives que s'aproximen a X en el sentit que acabem de precisar. Observem que $E(X) \in [0, +\infty]$. La definició és independent de la successió que s'aproxima a X i és consistent amb 3.1.1.

Proposició 3.1.4. *Siguin $\{X_n, n \geq 1\}$, $\{Y_n, n \geq 1\}$ dues successions de v.a. simples i positives que s'aproximen a la variable aleatòria X. Aleshores, els límits quan $n \to \infty$ de $E(X_n)$ i $E(Y_n)$ coincideixen.*

Demostració. Sigui Z una variable aleatòria simple tal que $Z \leq X$. Es compleix que $E(Z) \leq \lim_n E(X_n)$, si apliquem $Z = Y_n$ obtindrem la desigualtat $\lim_n E(X_n) \geq \lim_n E(Y_n)$ i, intercanviant els papers de les successions, $\lim_n E(X_n) \leq \lim_n E(Y_n)$. Sigui $\varepsilon > 0$. Definim $A_n = \{X_n > Z - \varepsilon\}$, $n \geq 1$. Els conjunts A_n creixen quan $n \to \infty$, cap a Ω i $X_n \geq (Z - \varepsilon)\mathbb{1}_{A_n}$. En conseqüència,

$$E(X_n) \geq E((Z-\varepsilon)\mathbb{1}_{A_n}) = E(Z\mathbb{1}_{A_n}) - \varepsilon P(A_n) \geq E(Z) - E(Z\mathbb{1}_{A_n^c} - \varepsilon \geq E(Z) - P(A_n^c)\max_\omega Z - \varepsilon.$$

Com que $\lim_n P(A_n^c) = 0$, el darrer terme d'aquesta desigualtat tendeix a $E(Z) - \varepsilon$, quan $n \to \infty$, i com ε és arbitrari, ja hem acabat. ■

Observació 3.1.5. Si X és una v.a. simple, podem considerar com a successió *aproximadora* la definida per $X_n = X$, per a tot n. Per tant, a la vista de 3.1.4, les definicions de 3.1.1 i 3.1.3 són coherents.

Sigui X una variable aleatòria arbitrària. Considerem $X = X^+ - X^-$ en termes de les seves parts positiva i negativa respectivament. Com que X^+ i X^- són variables aleatòries positives, podem definir $E(X^+)$ i $E(X^-)$.

Definició 3.1.6. Donada una variable aleatòria X qualsevol, direm que es té esperança matemàtica finita si, i només si, $E(|X|) < \infty$ i, en aquest cas, definim $E(X) = E(X^+) - E(X^-)$.

Notació 3.1.7. A partir d'ara denotarem $L^1(\Omega)$ el conjunt de les variables aleatòries definides en un espai de probabilitat $(\Omega, \mathcal{F}, P)$ que tenen esperança matemàtica finita.

Proposició 3.1.8. *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. *Si $X \leq Y$, aleshores $E(X) \leq E(Y)$.*
3. *Propietat de convexitat. Es compleix que $|E(X)| \leq E(|X|)$.*

Teorema 3.1.9 (de la convergència monòtona)**.** *Sigui $\{X_n\}_n$ una successió creixent de variables aleatòries positives que convergeix puntualment cap a una variable aleatòria X. Aleshores, $\lim_n E(X_n) = E(X)$.*

Teorema 3.1.10 (de la convergència dominada). *Sigui $\{X_n\}_n$ una successió de variables aleatòries que convergeix puntualment cap a una variable aleatòria X. Suposem que $|X_n| \leq Y$, on $Y \in L^1(\Omega)$. Aleshores, $\lim_n E(X_n) = E(X)$. Si no tinc convergència creixent, haig de tenir acotació.*

Observació 3.1.11. Per a qualsevol $x \in L^1(\Omega)$, existeix $E(X)$. Provem-ho per al cas discret:

$$X = \sum_{i=1}^{n} a_i \mathbb{1}_{A_i},\ E(X) = \sum_{i=1}^{n} a_i P(X = a_i).$$

Per a $x \geq 0$,

$$X = \sum_{i=1}^{\infty} a_i \mathbb{1}_{A_i}(\omega) \text{ i } \lim_{n\to\infty} \sum_{i=1}^{n} a_i \mathbb{1}_{A_i}(\omega).$$

Proposició 3.1.12 (Esperança d'una variable absolutament contínua). *Una variable aleatòria X amb densitat f pertany a $L^1(\Omega)$ si, i només si, $\int_{-\infty}^{+\infty} |x| \cdot f(x)\,dx < +\infty$. En aquest cas, $E(X) = \int_{-\infty}^{+\infty} x \cdot f(x)\,dx$.*

Demostració. Podem dividir l'interval $[a, b]$ de la manera següent:

$$[a, b] = [a, x_1] \cup (x_1, x_2] \cup \cdots \cup (x_{n-1}, x_n] \cup (x_n, b],\ \mathcal{P} = \{a = x_0, x_1, \dots, x_n, x_{n+1} = b\}$$

Sigui ara una variable aleatòria simple definida de la següent manera, amb la seva respectiva esperança:

$$D^n(\omega) = \sum_{i=0}^{n} x_i \mathbb{1}_{(x_i, x_[i+1]}(\omega),\ E(D^n) = \sum_{i=0}^{n} a_i \cdot P(x_i \leq X \leq x_{i+1}) = \sum_{i=0}^{n} a_i \int_{x_i}^{x_{i+1}} f(x)\,dx\,.$$

Ara, prenem m_i l'ínfim de f avaluada en $(x_i, x_{i+1}]$; és a dir, $m_i = \inf_{x\in(x_i,x_{i+1}]} f(x)$. Anàlogament, M_i el suprem de f avaluada en $(x_i, x_{i+1}]$; és a dir, $M_i = \sup_{x\in(x_i,x_{i+1}]} f(x)$ Aleshores:

$$S_n = \sum_{i=0}^{n} a_i m_i (x_{i+1} - x) \leq E(D^n) = \sum_{i=0}^{n} a_i M_i (x_{i+1} - x_i).$$

Ara anem a plantejar una desigualtat els extrems de la qual tendeixen simultàniament a zero, de manera que tindrem la igualtat que volem provar:

$$\begin{aligned}\sum_{i=0}^{n} x_i m_i (x_{i+1} - x_i) &\leq \sum_{i=0}^{n} \inf(xf(x))(x_{i+1} - x_i) \leq \int_a^b xf(x)\,dx \overset{1}{=} \sum_{i=0}^{n} \int_{x_i}^{x_{i+1}} xf(x)\,dx \\ &\leq \sum_{i=0}^{n} \sup(xf(x))(x_{i+1} - x_i) \leq \sum_{i=0}^{n} x_{i+1} M_i (x_{i+1} - x_i).\end{aligned}$$

I amb això ja hem acabat. ■

[1] Aquesta és la tesitura que plantejàvem.

3.2

Moments de variables aleatòries

Definició 3.2.1 (Moment no centrat). 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 *no centrat* d'ordre k (sempre que existeixi tal esperança).

1. El moment d'ordre $k = 1$ és l'*esperança* matemàtica.
2. El moment d'ordre $k = 2$ s'anomena *variància*, i és $\mathrm{Var}(X) = E(X - E(X))^2 = E(X)^2 - E^2(X)$. 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 $\sigma(X)$.

Procés 3.2.2 (Càlcul de moments). *Si existeix $\mu_1 = E(X)$, els altres moments es calculen tenint en compte la definició d'esperança d'una funció d'una variable aleatòria; així, existeix el moment no centrat d'ordre k, donat per:*

$$m_k = \begin{cases} \sum_{x \in E_X} x^k P(X = x), & \text{si } X \text{ és discreta.} \\ \int_{\mathbb{R}} x^k f_X(x)\, dx, & \text{si } X \text{ és contínua.} \end{cases}$$

Proposició 3.2.3.

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. *Desigualtat de Minkowskii. Si X, Y són dues variables aleatòries amb moment d'ordre k finit, la variable aleatòria $X + Y$ té també moment d'ordre k finit.*
3. *Desigualtat de Hölder. Sigui X una variable aleatòria amb moment d'ordre p finit, sigui Y una variable aleatòria amb moment d'ordre q finit, tal que $\frac{1}{p} + \frac{1}{q} = 1$. Aleshores:*

$$E(|XY|) \leq (E|X|^p)^{\frac{1}{p}} + (E|Y|^q)^{\frac{1}{q}}$$

Demostració.

1. 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 $|x|^r \leq |x|^k$; així:

$$|x|^r = |x|^r \cdot \mathbb{1}_{|x| \leq 1} + |x|^r \cdot \mathbb{1}_{|x| > 1} \leq 1 + |x|^k.$$

En conseqüència, suposant que $\{a_1, a_2, \dots\}$ és el conjunt de valors presos per la variable X i denotant per $p(a_i)$ les probabilitats $P(\{X = a_i\})$, tenim:

$$\sum_i |a_i|^r p(a_i) \leq \sum_i (1 + |a_i|^k) p(a_i) \leq 1 + \sum_i |a_i|^k p(a_i).$$

Per hipòtesi, el darrer membre d'aquesta expressió és finit, i amb això ja en tenim prou.

2. Conseqüència de *Minkowskii.* Per a tot nombre natural k i per a tots $x, y \in \mathbb{R}$ es compleix que:

$$|x+y|^k \leq (|x|+|y|)^k \leq c(k)\cdot(|x|^k+|y|^k).$$

Aquesta $c(k)$ és una constant que depèn de k. A [San99], tenim $c(k) = 2^{k-1}$. Denotem per $(a_i)_i$ i $(b_j)_j$ els valors presos per les variables aleatòries X, Y i per $p(a_i, b_j) = P(\{X = a_i, Y = b_j\})$. Aleshores, tenim les desigualtats següents, aplicant l'esperança (*fet a classe*):

$$E(|X+Y|^n) \leq c(k)\left(E(|X|^k)+E(|Y|^k)\right) < \infty.$$

3. Conseqüència de *Hölder.* Siguin $c, d > 0$ i $\alpha, \beta > 1$, $\frac{1}{\alpha}+\frac{1}{\beta} = 1$. Podem escriure $cd = (c^\alpha)^{\frac{1}{\alpha}} \cdot (d^\beta)^{\frac{1}{\beta}}$. Aplicant logaritmes tenim:

$$\log(cd) = \frac{1}{\alpha}c^\alpha + \frac{1}{\beta}d^\beta \overset{2}{\leq} \log\left(\frac{1}{\alpha}c^\alpha+\frac{1}{\beta}d^\beta\right) \implies cd \leq \frac{1}{\alpha}c^\alpha + \frac{1}{\beta}d^\beta.$$

Aplicant-ho al nostre cas, $\|x\|_p = (E(|x|^p))^{\frac{1}{p}}$ i $\|y\|_q = (E(|y|^q))^{\frac{1}{q}}$, tals que $\alpha = p$ i $\beta = q$. Aleshores, $c = \frac{|x|}{\|x\|_p}$ i $d = \frac{|y|}{\|y\|_q}$. Per tant:

$$\frac{|X-Y|}{\|x\|_p\cdot\|y\|_q} \leq \frac{1}{p}\frac{|x|^p}{\|x\|_p^p} + \frac{1}{q}\frac{|y|^q}{\|y\|_q^q}$$

Així doncs, com volíem veure:

$$E\left(\frac{|X-Y|}{\|x\|_p\cdot\|y\|_q}\right) \leq \frac{1}{p}E\left(\frac{|x|^p}{\|x\|_p^p}\right) + \frac{1}{q}E\left(\frac{|y|^q}{\|y\|_q^q}\right) = \frac{1}{p}+\frac{1}{q} = 1.$$ ■

Definició 3.2.4 (Moment central d'ordre k). El moment central d'ordre k d'una variable aleatòria X és $\alpha_k(X) = E((X-E(X))^k)$, el segon moment central és la variància. En general, si $c \in \mathbb{R}$ es defineix el moment *centrat* en c d'ordre k com $E((X-c)^k)$ i els moments centrats en mitjana s'anomenen simplement *moments centrats.* Els moments centrats es calculen, com els no centrats, tenint en compte la definició d'esperança d'una funció d'una variable aleatòria:

$$\alpha_k(X) = \begin{cases} \displaystyle\sum_{x\in E_X}(x-E[X])^k P(X=x), & \text{si } X \text{ és discreta;} \\ \displaystyle\int_{\mathbb{R}}(x-E[X])^k f_X(x)\,dx, & \text{si } X \text{ és contínua.}\end{cases}$$

Proposició 3.2.5 (Desigualtat de Txebitxev). *Sigui X una variable aleatòria no negativa i $f : \mathbb{R}_+ \longrightarrow \mathbb{R}_+$ una funció creixent tal que $E(f(X)) < \infty$. Per a tot nombre real a es compleix que $f(a)P(X \geq a) \leq E(f(X))$. Equivalentment, $P(X \geq a) \leq \frac{1}{f(a)}E(f(X))$.*

Demostració. Considerem la desigualtat $f(a)\mathbb{1}_{X\geq A} \leq f(X)$.

$$f(a)\mathbb{1}_{X\geq A} = \begin{cases} 0, & \text{si } X(\omega) < a \\ f(a), & \text{si } X(\omega) \geq a\end{cases}$$

[2] Fem servir que el logaritme és una funció còncava.

Prenent esperances s'obté que $0 \leq E(f(a)\mathbb{1}_{\{X\geq A\}} \leq E(f(X))$, i com $E(f(a)\mathbb{1}_{\{X\geq a\}} = f(a)E\mathbb{1}_{\{X\geq a\}}$ i això és $f(a)P(X \geq a)$, que és la desigualtat buscada. ■

Observació 3.2.6. En particular, si X té moment d'ordre k finit i $a > 0$, tindrem que $P(\{X \geq a\}) \leq \frac{E(|X|^k)}{a^k}$. En efecte, ho obtenim d'aplicar la desigualtat de Txevitxev a la funció $f(x) = x^k$ i a la variable $|X|$.

Proposició 3.2.7 (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, $g(E(X)) \leq E(g(X))$.*

Demostració. Com que g és convexa, existeix $m_{E(X)} \in \mathbb{R}$ tal que per a tot $x \in \mathbb{R}$,

$$g(x) \geq g(E(X)) + m_{E(X)}(x - E(X)) \xrightarrow{x \to X(\omega)} g(X(\omega)) \geq g(E(X)) + m_{E(X)}(X(\omega) - E(X)),$$

$\forall \omega \in \Omega$. La propietat de monotonia de l'esperança matemàtica permet concloure la demostració:

$$E(g(X(\omega))) \geq E(g(E(X))) + m_{E(X)}E(X(\omega) - E(X)) \xrightarrow{X(\omega) - E(X) \to 0} E(g(X(\omega))) \geq E(g(E(X))).$$

I, per tant, $E(g(X(\omega))) \geq E(g(E(X)))$, com volíem. ■

Exemple 3.2.8. Fem el cas particular $g(x) = |x|$; clarament, g és convexa i tenim $|E(X)| \leq E(|X|)$.

3.3 Independència i esperança

Sigui $(X_1, \ldots, X_n)$ un vector aleatori discret que pren valors a I. Sigui $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ mesurable. Recordem que aquesta propietat implica que $g(X)$ és una variable aleatòria.

Proposició 3.3.1 (Esperança d'una funció amb v.a. absolutament contínua). *Si X és un vector aleatori discret que pren els seus valors en un conjunt mesurable I, la variable aleatòria $g(X)$ té esperança finita si, i només si,*

$$\sum_{(x_1,\ldots,x_n)\in I} |g(x_1, \ldots, x_n)| \cdot P(\{X_1 = x_1, \ldots, X_n = x_n\}) \ll \infty.$$

En tal cas,

$$E(g(X)) = \sum_{(x_1,\ldots,x_n)\in I} g(x_1, \ldots, x_n) \cdot f(x_1, \ldots, x_n)\, dx_1 \cdots dx_n\,.$$

Si, a més, X és un vector aleatori absolutament continu, $Y = g(X)$ té esperança finita si:

$$\int_{-\infty}^{+\infty} |g(x)| \cdot f(x)\, dx < \infty \implies E(Y) = \int_{-\infty}^{+\infty} g(x)f(x)\, dx$$

O fent un canvi de variable, ens quedaria l'equivalent $E(Y) = \int_{-\infty}^{+\infty} y f_Y(y)\, dy$.

Demostració. En el cas particular de $g : I \longrightarrow J$, bijectiva i de classe $\mathcal{C}^1$ de l'interval I en un altre interval obert J, en què sabem calcular la pdf de Y:

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

Existeix $E(Y)$ si, i només si, $\int_J |y| f_Y(y)\, dy \ll \infty$. Fent el canvi de variable $y = f(x)$ a la integral, surt la condició de l'enunciat. ■

Proposició 3.3.2. *Siguin X, 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 del cas discret. Siguin $\{x_1, x_2, \dots\}$ i $\{y_1, y_2, \dots\}$ els conjunts on prenen els seus valors les variables aleatòries X i Y, respectivament. Per provar que XY té una esperança finita, cal provar 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 es té que:

$$\sum_{i,j} |x_i y_j| P(X = x_i, Y = y_j) = \sum_{i,j} |x_i| P(X = x_i) \cdot \sum_{i,j} |y_j| P(Y = y_j) = E(|X|)E(|Y|) \ll \infty.$$

D'altra banda, els càlculs fets anteriorment sense prendre els valors absoluts demostren la igualtat $E(XY) = E(X)E(Y)$. ■

Demostració del cas continu. Suposem ara que X i Y tenen llei absolutament contínua amb densitats f_X i f_Y, respectivament. El vector (X, Y) té, doncs, llei absolutament contínua i la seva densitat és $f(x, y) = f_X(x)f_Y(y)$. En conseqüència:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |xy| f(x, y)\, dx\, dy = \left(\int_{-\infty}^{\infty} |x| f_X(x)\, dx\right)\left(\int_{-\infty}^{\infty} |y| f_Y(y)\, dy\right) = E(XY) \ll \infty.$$

D'altra banda, sense prendre els valors absoluts demostrem la igualtat $E(XY) = E(X)E(Y)$. ■

Proposició 3.3.3. *Siguin X, 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 variància:

$$\begin{aligned}
\mathrm{Var}(X + Y) &= E\left[((X + Y) - E(X + Y))^2\right] \\
&= E\left[((X - E(X)) + (Y - E(Y)))^2\right] \\
&= E\left[(X - E(X))^2\right] + E\left[(Y - E(Y))^2\right] + 2E[(X - E(X))(Y - E(Y))]
\end{aligned}$$

La propietat d'independència i la linealitat de l'esperança implica de manera que el resultat queda demostrat.

$$E[(X - E(X))(Y - E(Y))] = 0.$$ ■

La proposició anterior s'estén sense dificultat a un nombre finit de variables aleatòries independents.

Definició 3.3.4 (Covariància). 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

$$\operatorname{Cov}(X,Y) = E[(X - E(X))(Y - E(Y))]$$

Si $X = Y$, aleshores $\operatorname{Cov}(X) = \operatorname{Var}(X)$. Si les variables X, Y són independent, aleshores $\operatorname{Cov}(X,Y) = 0$. Les variables aleatòries que tenen covariància nul·la s'anomenen *incorrelacionades*. Ara bé, dues variables no nul·les poden tenir convariància nul·la i no ser independents.

Definició 3.3.5 (Coeficient de correlació). Es defineix el coeficient de correlació entre dues variables aleatòries X, 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, Y són independents, aleshores $\rho_{X,Y} = 0$.

Proposició 3.3.6. *Siguin X, Y variables aleatòries amb moments de segon ordre finits. Aleshores, es compleix que $(\operatorname{Cov}(X,Y))^2 \leq \operatorname{Var}(X)\operatorname{Var}(Y)$. Dit d'una altra manera, $\rho_{X,Y} \in [-1,1]$.*

Demostració. Sigui $\lambda \in \mathbb{R}$. Se 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,

$$4(\operatorname{Cov}(X,Y))^2 - 4\operatorname{Var}(X)\operatorname{Var}(Y) \leq 0 \iff (\operatorname{Cov}(X,Y))^2 - \operatorname{Var}(X)\operatorname{Var}(Y) \leq 0,$$

que és el que es volia provar. ∎

IV

Successions de variables aleatòries

4.1 Lemes de Börel-Cantelli

Definició 4.1.1. Considerem una successió $\{A_n\}_n$ de conjunts d'$\mathcal{A}$, una σ-àlgebra de parts d'un conjunt Ω. Definim:

$$\limsup_n A_n = \bigcap_{n\geq 1}\bigcup_{k\geq n} A_k, \text{ i } \liminf_n A_n = \bigcup_{n\geq 1}\bigcap_{k\geq n} A_k.$$

A causa de l'estructura de σ-àlgebra d'$\mathcal{A}$, clarament $\limsup_n A_n$ i $\liminf_n A_n$ són elements d'$\mathcal{A}$.

Proposició 4.1.2.

1. *Un element ω pertany a $\limsup_n A_n$ si, i només si, hi ha infinits n tals que ω pertany a A_n.*
2. *Un element ω pertany a $\liminf_n A_n$ si, i només si, existeix un $n_0 \in \mathbb{N}$, que depèn de ω, tal que ω pertany a tots els A_n per a $n \geq n_0$.*

Demostració.

1. 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:

$$\omega \in \bigcap_{j\in\mathbb{N}} A_{k_j} \subset \bigcap_{n\geq 1}\bigcup_{k\geq n} A_k.$$

2. Suposem ara que ω pertany a $\liminf_n A_n$. Aleshores, existeix $n_0(\omega) \geq 1$ tal que 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 $\liminf_n A_n$. ■

Observació 4.1.3. Es dedueix fàcilment de 4.1.2 que $\liminf_n A_n \subset \limsup_n A_n$. D'altra banda, de les regles de les operacions amb conjunts s'obté que:

$$\left(\limsup_n A_n\right)^c = \liminf_n A_n^c \text{ i } \left(\liminf_n A_n\right)^c = \limsup_n A_n^c$$

Lema 4.1.4. *Donada una successió $\{A_n\}_n$ d'esdeveniments d'A es compleix que $P(\limsup_n A_n) \geq \limsup_n P(A_n)$ i, anàlogament, $P(\liminf_n A_n) \leq \liminf_n P(A_n)$.*

Demostració. Podem escriure:

$$P\left(\limsup_n A_n\right) = \lim_{n\to\infty} P\left(\bigcup_{m\geq n} A_m\right) \geq \lim_{n\to\infty}\left(\sup_{m\geq n} P(A_m)\right) = \limsup_n P(A_n).$$

L'altre resultat es dedueix per pas al complementari, en efecte:

$$P\left(\liminf_n A_n\right) = 1-P\left(\left(\liminf_n A_n\right)^c\right) = 1-P\left(\limsup_n A_n^c\right) \leq 1-\limsup_n P(A_n^c) = \liminf_n P(A_n).$$

■

Lema 4.1.5 (Primer lema de Börel-Cantelli). *Sigui $\{A_n\}_n$ una successió de conjunts d'$\mathcal{A}$. Si $\sum_n P(A_n) < \infty$, es compleix que $P(\limsup_n A_n) = 0$.*

Demostració. Es té que:

$$P\left(\limsup_n A_n\right) = P\left(\bigcap_{n\geq 1}\bigcup_{m\geq n} A_m\right) = \lim_{n\to\infty} P\left(\bigcup_{m\geq n} A_m\right) \leq \lim_n \sum_{m\geq n} P(A_m) = 0.$$ ■

Lema 4.1.6 (Segon lema de Börel-Cantelli). *Sigui $\{A_n\}_n$ una successió d'esdeveniments independents d'$\mathcal{A}$. Si $\sum_n P(A_n) = \infty$, es compleix que $P(\limsup_n A_n) = 1$.*

Demostració. Veurem que $P((\limsup_n A_n)^c) = P(\bigcup_n \bigcap_{m\geq n} A_m^c) = 0$. I per 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:

$$P\left(\bigcap_{m=n}^{n+j} A_m^c\right) = \prod_{m=n}^{n+j}(1-P(A_m)) \leq \prod_{m=n}^{n+j} e^{-P(A_m)} = e^{-\sum_{m=n}^{n+j} P(A_m)} \to 0.$$ ■

Exemple 4.1.7. 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 Börel-Cantelli, 4.1.6, $P\{\limsup_n A_n\} = 1$, 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 4.1.5 i 4.1.6 demostrats abans, tenim el resultat següent.

Lema 4.1.8 (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 $\sum_{n=1}^{\infty} P(A_n) = \infty$.*

Teorema 4.1.9 (Infinite monkey theorem, [Wik23]). *The infinite monkey theorem states that a monkey hitting keys at random on a typewriter keyboard for an infinite amount of time will almost surely type any given text, including the complete works of William Shakespeare. In fact, the monkey would almost surely type every possible finite text an infinite number of times. The theorem can be generalized to state that any sequence of events which has a non-zero probability of happening will almost certainly eventually occur, given unlimited time. Més generalment, se sol considerar una cadena C de longitud s dels k símbols: si el procés segueix indefinidament, eventualment apareixerà la cadena C en la successió resultat de l'experiment.*

Demostració. Recall that if two events are statistically independent, then the probability of both happening equals the product of the probabilities of each one happening independently. Suppose that the keys are pressed randomly and independently, meaning that each key has an equal chance of being pressed regardless of what keys had been pressed previously. Subdividim la successió resultant en segments de longitud s. La probabilitat que un d'aquests segments coincideixi amb C és $\frac{1}{k^s}$. La probabilitat que en $n \geq 1$ dels segments *no* s'hagi obtingut C és de $(1 - \frac{1}{k^s})^n$. Com que $(1 - \frac{1}{k^s}) \in (0, 1)$, quan n creix indefinidament:

$$\lim_{n \to \infty} \left(1 - \frac{1}{k^s}\right)^n = 0.$$

En efecte, l'esdeveniment complementari (la cadena C eventualment apareix) té una probabilitat 1. ∎

4.2 Convergència quasi segura

Definició 4.2.1 (Convergència quasi segura). Direm que una successió de variables aleatòries $\{X_n\}_n$ convergeix *quasi segurament* cap a una variable aleatòria X si existeix un conjunt $N \in \mathcal{A}$ de probabilitat zero tal que $\lim_n X_n(\omega) = X(\omega)$, per a tot $\omega \notin N$.

Si X_n convergeix quasi segurament cap a X quan $n \to \infty$, escriurem $X_n \to X$ QS o bé $\lim_n X_n = X$, QS.

1. Es tracta, doncs, d'una convergència puntual de la successió numèrica $X_n(\omega)$, en tots els punts $\omega \in \Omega \setminus N$.
2. La variable límit és única llevat el conjunts de probabilitat zero. Naturalment, que la successió $\{X_n\}_n$ convergeix QS si, i només si, és de Cauchy amb probabilitat 1.

Exemple 4.2.2. 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 $X_n \longrightarrow 0$, QS. 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, si $x \in (0,1]$. Si $x = 0$, $\lim_n X_n(x) = +\infty$. 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, $\lim_n X_n(x) = 0$, QS.

Proposició 4.2.3. *La successió $\{X_n\}$ convergeix QS a la variable aleatòria X si, i només si, per a tot $\varepsilon > 0$ la probabilitat $P(\liminf_n\{|X_n - X| \leq \varepsilon\}) = 1$.*

Demostració.

$\boxed{\Rightarrow}$ Suposem primer que $\lim_n X_n = X$, QS, i sigui N el subconjunt d'$\mathcal{A}$ de probabilitat zero on falla la convergència puntual. Designem per Ω_0 el conjunt $\Omega \setminus N$. Es compleix que:

$$\left\{\sup_{n\geq m} |X_n - X| \leq \varepsilon\right\} = \bigcap_{n\geq m} \{|X_n - X_m| \leq \varepsilon\} := A_m^\varepsilon.$$

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, que:

$$1 = P(\Omega_0) \leq P\left(\bigcup_{m=1}^\infty A_m^\varepsilon\right) = \lim_{m\to\infty} P(A_m^\varepsilon) = P\left(\left\{\liminf_n |X_n - X| \leq \varepsilon\right\}\right) = 1.$$

$\boxed{\Leftarrow}$ 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 $m \geq m_0$ tenim $|X_m(\omega) - X(\omega)| \leq \varepsilon$. Sigui $N = \bigcup_{\varepsilon\in\mathbb{Q}^+} N_\varepsilon$. Observem que $P(N) = 0$, ja que N és una reunió numerable de conjunts de probabilitat zero. Aleshores, per a tot $\omega \notin N$ tenim $\lim_n X_n(\omega) = X(\omega)$; és a dir, la convergència QS de la successió $\{X_n\}_n$ cap a X. ■

Proposició 4.2.4. *Sigui $\{\varepsilon_n\}_n$ una successió de nombres reals positius tal que $\sum_n \varepsilon_n < \infty$. Sigui $\{X_n\}_n$ una successió de variables aleatòries i suposem que $\sum_n P(|X_{n+1} - X_n| \geq \varepsilon_n) < \infty$, aleshores $\{X_n\}_n$ convergeix QS.*

Demostració (no entra). 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 Börel-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 4.2.3 i 4.2.4 és similar. En efecte, la condició que apareix en la proposició 4.2.3 és equivalent a

$$P\left\{\limsup_n |X_n - X| \geq \varepsilon\right\} = 0.$$

Si $\sum_{n=1}^{\infty} P\{|X_n - X| \geq \varepsilon\} < \infty$, el primer lema de Börel-Cantelli assegura la seva validesa. Observem també que, en la proposició 4.2.3 apareix un candidat a variable aleatòria límit, mentre que en la proposició 4.2.4 es demostra únicament l'existència del límit.

Exemple 4.2.5. Considerem una successió de variables aleatòries independents, $\{X_n\}_n$ amb $P(X_n = 0) = 1 - \frac{1}{n}$, $P(X_n = 1) = \frac{1}{n}$. Fixem $\varepsilon > 0$. La sèrie $\sum_n P(|X_{n^2}| > \varepsilon)$ convergeix. En efecte,

$$\sum_{n\geq 1} P(|X_{n^2}| > \varepsilon) = \sum_{n\geq 1} P(X_{n^2} = 1) = \sum_{n\geq 1} \frac{1}{n^2} < \infty.$$

Pel primer lema de Börel-Cantelli, $P(\limsup_n\{|x_{n^2}| \geq \varepsilon\}) = 0$, i per la proposició 4.2.3, deduïm que $\lim_n X_{n^2} = 0$, QS. En canvi, la successió $\{X_n\}_n$ no convergeix QS a 0. En efecte, si això fos cert, per 4.2.3 un altre cop es tindria que $P(\limsup_n\{|x_n| \geq \varepsilon\}) = 0$ per a tot $\varepsilon > 0$. Ara bé:

$$\sum_{n=1}^{\infty} P(|X_n| > \varepsilon) = \sum_{n\geq 1} P(|X_n| = 1) = \sum_{n=1}^{\infty} \frac{1}{n} = +\infty.$$

Com que els esdeveniments $\{|X_n| > \varepsilon\}$ són independents, ja que les variables aleatòries X_n ho són, el segon lema de Börel-Cantelli implica que $P(\{\limsup_n\{|X_n| > \varepsilon\}\}) = 1$.

4.3 Convergència en probabilitat

Podríem dir, en paraules planeres, que en la convergència en probabilitat la probabilitat juga un paper més important: serveix per mesurar el grau d'apropament (o distanciament) dels termes de la successió al límit hipotètic.

Definició 4.3.1 (Convergeix en probabilitat). Direm que una successió de variables aleatòries $(X_n)_n$ convergeix en probabilitat *cap a una variable aleatòria* X si per a tot $\varepsilon > 0$ tenim $\lim_n P(\{|X_n - X| > \varepsilon\}) = 0$. Escriurem $P - \lim_n X_n = X$. En augmentar n és cada vegada menys probable que X_n i X difereixin en més de ε, per a tot $\varepsilon > 0$, fixat.

Exemple 4.3.2. Sigui $\Omega = [0,1]$, $\mathcal{A}$ la σ-àlgebra de Börel 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 $\{X_{n,m} \bigm| n \geq 1, 1 \leq m \leq n\}$ convergeix en probabilitat cap a zero. En efecte:

$$P(\{|X_{n,m}| > \varepsilon\}) = \begin{cases} 0, & \text{si } \varepsilon > 1, \\ P(\{x \bigm| X_{n,m}(x) = 1\}), & \text{si } \varepsilon \leq 1. \end{cases}$$

La convergència cap a zero es dedueix del fet que $P(\{x \bigm| X_{n,m}(x) = 1\}) = \frac{1}{n}$.

Proposició 4.3.3. *Siguin $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ una funció contínua i $(X_n)_n, (Y_n)_n$ successions de variables aleatòries que convergeixen en probabilitat cap a les variables aleatòries X, Y, respectivament. Aleshores, la successió $(f(X_n, Y_n))_n$ convergeix en probabilitat cap a $f(X, Y)$.*

Demostració. Fixem $\varepsilon > 0$ i $\eta > 0$. 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 $\lim_n P(|Z| > n) = 0$, 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 $|x'|, |x|, |y'|, |y| \leq k$, $|x - x'| \leq \delta$ i $|y - y'| \leq \delta$, aleshores $|f(x, y) - f(x', y')| \leq \varepsilon$. Considerem els conjunts:

$$\begin{aligned} A &= \{|f(X_n, Y_n) - f(X, Y)| > \varepsilon\}, \\ B &= \left\{|X| \leq \frac{k}{2}, |Y| \leq \frac{k}{2}, |X_n - X| \leq \frac{k}{2}, |Y_n - Y| \leq \frac{k}{2}\right\}. \end{aligned}$$

Escrivint $P(A) = P(A \cap B) + P(A \cap B^c)$, i donat que $B \subset \{|X|, |Y|, |X_n|, |Y_n| \leq k\}$ (conseqüència de la desigualtat triangular), tenim:

$$\begin{aligned} P(\{|f(X_n, Y_n) - f(X, Y)| > \varepsilon\}) &\leq P\left(\left\{|X| > \frac{k}{2}\right\}\right) + P\left(\left\{|Y| > \frac{k}{2}\right\}\right) \\ &+ P\left(\left\{|X_n - X| > \frac{k}{2}\right\}\right) + P\left(\left\{|Y_n - Y| > \frac{k}{2}\right\}\right) \\ &+ P(|X| \leq k, |Y| \leq k, |X_n| \leq k, |Y_n| \leq k, |f(X_n, Y_n) - f(X, Y)| \geq \varepsilon) \\ &\leq 2\eta + P\left(|X - X_n| > \frac{k}{2}\right) + P\left(|Y - Y_n| > \frac{k}{2}\right) \\ &+ P\left(|X - X_n| > \delta\right) + P\left(|Y - Y_n| > \delta\right) \end{aligned}$$

Fent $n \to \infty$, i tenint en compte que η és arbitrari, obtenim la convergència desitjada. ■

Corol·lari 4.3.4. *EL límit d'una suma (producte) de successions convergents en probabilitat és la suma (producte) dels respectius límits.*

Definició 4.3.5. 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 $X = Y$ QS. Denotarem per $L^0(\Omega, \mathcal{A}, P)$ el conjunt de les classes d'equivalència de variables aleatòries mòdul aquesta relació. L'espai $L^0(\Omega, \mathcal{A}, P)$ amb la convergència en probabilitat és un espai mètric complet.

Proposició 4.3.6. *Definida l'aplicació $d : L^0(\Omega, \mathcal{A}, P) \times L^0(\Omega, \mathcal{A}, P) \longrightarrow \mathbb{R}$ per $E(\frac{|X-Y|}{1+|X-Y|})$, és una distància sobre $L^0(\Omega, \mathcal{A}, P)$ i metritza la convergència en probabilitat. El límit en probabilitat d'una successió de variables aleatòries, si existeix, és únic.*

Definició 4.3.7 (de Cauchy en probabilitat). Una successió $(X_n)_n$ de variables aleatòries és de Cauchy en probabilitat si $P - \lim_{n,m}(X_m - X_n) = 0$.

Proposició 4.3.8. *Si una successió de variables aleatòries convergeix en probabilitat, aleshores és de Cauchy en probabilitat.*

Demostració fàcil, però no entra. De la desigualtat triangular es dedueix la inclusió:

$$\{|X_n - X_m| > \varepsilon\} \subset \left\{|X_n - X| > \frac{\varepsilon}{2}\right\} \cup \left\{|X_m - X| > \frac{\varepsilon}{2}\right\}. \qquad \blacksquare$$

4.4 Convergència en mitjana

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; això és, $E(|X|^p) < \infty$. Per a successions de variables aleatòries, en aquest conjunt podem estudiar la convergència en *mitjana*.

Definició 4.4.1 (Convergeix en mitjana d'ordre p). Sigui $(X_n)_n$ 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 també té moment d'ordre p finit:

$$\lim_{n\to\infty} E(|X_n - X|^p) = 0.$$

Escriurem $L^p - \lim_n X_n = X$.

Exemple 4.4.2. Sigui $(X_n)_n$ una successió de variables aleatòries amb $P(X_n = n) = \frac{1}{n^2}$ i $P(X_n = \frac{1}{n}) = 1 - \frac{1}{n^2}$. 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,

$$P(|X_n| > \varepsilon) = P(X_n = n) = \frac{1}{n^2} \to 0,\ n \to \infty.$$

En canvi, no hi ha convergència en mitjana quadràtica cap a zero, ja que:

$$E(|X|^2) = 1 + \frac{1}{n^2} - \frac{1}{n^4} \to 1, \ n \to \infty.$$

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 4.5.1 (Convergència en llei). Sigui $(X_n)_n$ una successió de variables aleatòries i $(F_n)_n$ la successió de les seves funcions de distribució respectives. Direm que la successió $(X_n)_n$ convergeix en llei cap a una variable aleatòria X que té funció de distribució F si es compleix que $\lim_n F_n(x) = F(x)$, per a tot $x \in \mathbb{R}$, on la funció F sigui contínua. Denotarem aquest tipus de convergència mitjançant $\mathcal{L} - \lim_n X_n = X$.

Observació 4.5.2. 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.

Proposició 4.5.3. *Sigui $(X_n)_n$ 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 $\lim_n F_n(x) = F(x)$ és que per a tot $k \in \mathbb{Z}_+$ es compleixi que $\lim_n P(X_n = k) = P(X = k)$.*

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:

$$\lim_{n\to\infty} P(X_n = k) = \lim_{n\to\infty} \left(F_n\left(k + \frac{1}{2}\right) - F_n\left(k + \frac{1}{2}\right)\right) = F\left(k + \frac{1}{2}\right) - F\left(k - \frac{1}{2}\right) = P(X = k).$$

Recíprocament, suposem $\lim_n P(X_n = k) = P(X = k)$. Sigui $x \in \mathbb{R}$, aleshores:

$$F_n(x) = P(X_n \leq x) = \sum_{k=0}^{x} P(X_n = k) \to \sum_{k=0}^{x} P(X = k) = P(X \leq x) = F(x), \ n \to \infty. \quad \blacksquare$$

Exemple 4.5.4. Considerem una successió de variables aleatòries $(X_n)_n$ tals que $P(X_n = a_n) = 1$. Suposem, a més, que $(a_n)_n \to a$, $a \in \mathbb{R}$, quan $n \to \infty$. Aleshores, (X_n) 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

$n_0 \in \mathbb{N}$ 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, que $\lim_n F_n(x) = F(x)$, per a tot punt $x \in \mathbb{R}$ on F és contínua.

Teorema 4.5.5. *Una successió $(X_n)_n$ de variables aleatòries convergeix en llei cap a una variable aleatòria X si, i només si, per a tota funció $f \in \mathcal{C}_b$ tenim $\lim_n E(f(X_n)) = E(f(X))$.*

4.6 Relacions entre convergències

Teorema 4.6.1. *Sigui $(X_n)_n$ una successió de variables aleatòries i X una altra variable aleatòria.*

1. *Si $X_n \to x$ QS, quan $n \to \infty$, aleshores $P - \lim_n X_n = X$.*
2. *Si $P - \lim_n X_n = x$, existeix una subsuccessió $(X_{n_k})_k$ que convergeix QS cap a la variable X quan $k \to \infty$.*

Demostració.

$\boxed{\Rightarrow}$ Fixem $\varepsilon > 0$, de manera que tenim $P(\liminf_n |X_n - X| \leq \varepsilon) = 1$. Aleshores:

$$\limsup_n P(|X_n - X| > \varepsilon) \leq P\left(\limsup_n |X_n - X| > \varepsilon\right) = 0 \implies P - \lim_n X_n = X.$$

$\boxed{\Leftarrow}$ Suposem que $P - \lim_n X_n = X$. Es compleix la condició de Cauchy, $\lim_{m,n} P(|X_n - X_m| > \varepsilon) = 0$, per a tot $\varepsilon > 0$. Considerem dues sèries de termes positius $\sum_n \varepsilon_n < \infty$ i $\sum_n \delta_n < \infty$. Definirem ara una successió de nombres naturals $(n_k)_k$ estrictament creixent de nombres naturals de la manera següent. Posem $n_0 = 0$ i per a tot k 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 4.2.4, resulta que la successió parcial $(X_{n_k})_k$ convergeix QS. El límit haurà de ser necessàriament la variable X, ja que **la convergència quasi segura implica la convergència en probabilitat** i sabem que $P - \lim_k X_{n_k} = X$. Fixem-nos que solament hem usat que $(X_n)_n$ és de Cauchy en probabilitat (que és equivalent a què sigui convergent en probabilitat). ■

Corol·lari 4.6.2. *Sigui $(X_n)_n$ una successió de variables aleatòries i X una altra variable aleatòria. Si la successió de variables és monòtona (creixent o decreixent), aleshores la convergència en probabilitat implica la quasi segura.*

Demostració. Suposem que $(X_n)_n$ convergeix en probabilitat cap a X i és creixent QS; 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,

$$\limsup_n(|X_n - X| > \varepsilon) = \bigcap_n \{|X_n - X| > \varepsilon\} \implies P\left(\limsup_n \{|X_n - X| > \varepsilon\}\right) = \lim_n P(|X_n - X| > \varepsilon) = 0.$$

Aquesta última implicació per la propietat de continuïtat de la probabilitat. En aquesta situació, $(X_n)_n$ convergeix QS cap a X. ∎

Proposició 4.6.3. *Sigui $(X_n)_n$ 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 també és 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 d'$\mathcal{A}$ de probabilitat zero i $k \in \mathbb{R}$.*

Demostració. Per a tot $\varepsilon > 0$ i $x \in \mathbb{R}$ es tenen les desigualtats:

$$\begin{aligned} &P(X_n \leq X) \leq P(|X_n - X| \geq \varepsilon) + P(X \leq x + \varepsilon), \\ &\qquad P(X \leq x - \varepsilon) \leq P(|X_n - X| \geq \varepsilon) + P(X_n \leq x). \end{aligned} \tag{4.1}$$

En efecte, per a demostrar la primera considerem les inclusions:

$$\{X_n \leq x\} \subset \{X_n \leq x\} \cap \{X \cap x + \varepsilon\} \cup \{X \leq x + \varepsilon\} \subset \{|X_n - X| \geq \varepsilon\} \cup \{X \leq x + \varepsilon\}.$$

Usant primer la segona desigualtat de (4.1), 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 $(X_n)_n$ convergeix en probabilitat cap a X. Tenim, doncs, que:

$$F(x) - \delta \leq F(x - \varepsilon) - \frac{\delta}{2} \leq F(x - \varepsilon) - P(|X_n - X| \geq \varepsilon) \leq F_n(x) \leq P(|X_n - X| \geq \varepsilon) + F(x + \varepsilon) \leq \delta + F(x).$$

Per tant, $\lim_n F_n(x) = F(x)$, per a tot $x \in \mathbb{R}$ on la funció F és contínua. Així doncs, 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:

$$P(|X_n - k| > \varepsilon) = P(X_n > k + \varepsilon) + P(X_n < k - \varepsilon) = 1 - F_n(k + \varepsilon) + F_n((k - \varepsilon)^-).$$

D'aquesta manera,

$$\lim_{n \to \infty} P(|X_k - X| \geq \varepsilon) \leq 1 - F(k + \varepsilon) + F(k - \varepsilon) = 0.$$ ∎

Teorema 4.6.4. *Sigui* $(Y_n)_n$ *una successió de variables aleatòries que convergeixen en llei cap a una variable aleatòria* Y*. Existeixen una successió de variables aleatòries* $(X_n)_n$ *i una variable aleatòria* X *que compleixen les següents propietats:*

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* $(X_n)_n$ *convergeix* QS *cap a la variable aleatòria* X*.*

V

Lleis dels grans nombres

Considerem una successió $(X_n)_n$ de variables aleatòries definides en un mateix espai de probabilitat $(\Omega, \mathcal{A}, P)$. Les lleis dels grans nombres estudien en el comportament asimptòtic de la successió de les mitjanes aritmètiques $\{\frac{S_n}{n}\}_n$, on $S_n = X_1 + \cdots + X_n$. S'analitzen dos tipus diferents de convergències: en probabilitat i quasi segura.

5.1 Lleis febles dels grans nombres

Teorema 5.1.1. *Sigui $(X_n)_n$ una successió de variables aleatòries incorrelacionades* ($\mathrm{Cov}(X_i, X_j) = 0$ *per a tot* $i \neq j$) *i amb un moment de segon ordre fitat per una constant independent de n. Aleshores,*

$$E\left(\left|\frac{S_n - E(S_n)}{n}\right|^2\right) \leq \frac{C}{n} \implies \lim_{n\to\infty} \frac{S_n - E(S_n)}{n} = 0,$$

en el sentit de la convergència en mitjana quadràtica.

Demostració. Com que la successió està formada per variables aleatòries incorrelacionades, es compleix que:

$$\begin{aligned} E\left(\left|\frac{S_n - E(S_n)}{n}\right|^2\right) &= \frac{1}{n^2} E\left(\left(\sum_{i=1}^n X_i - \sum_{i=1}^n E(X_i)\right)^2\right) = \frac{1}{n^2} E\left(\left(\sum_{i=1}^n X_i - E(X_i)\right)^2\right) \\ &= \frac{1}{n^2}\left(E\left(\sum_{i=1}^n (x_i - E(X_i))^2\right) + E\left(\sum_{i\neq j} (X_i - E(X_i))(X_j - E(X_j))\right)\right) \\ &= \frac{1}{n^2} \sum_{i=1}^n E\left((X_i - E(X_i))^2\right). \end{aligned}$$

D'altra banda, per hipòtesi tenim que $E(X_i - E(X_i))^2 = \mathrm{Var}(X_i) \leq C$. Per tant, l'últim terme està fitat per $\frac{C}{n}$. En més detall:

$$\frac{1}{n^2}\sum_{i=1}^n \mathrm{Var}(X_i) = \frac{1}{n^2}\sum_{i=1}^n (E(X_i^2) - E(X_i)^2) \leq \frac{1}{n^2}\sum_{i=1}^n E(X_i^2) \leq \frac{1}{n^2}\sum_{i=1}^n C = \frac{nC}{n^2} = \frac{C}{n}. \qquad \blacksquare$$

Notació 5.1.2. Moltes vegades ens trobem amb successions de variables aleatòries independents i amb la mateixa distribució (independents i idènticament distribuïdes, i.i.d).

Corol·lari 5.1.3. *Sigui $(X_n)_n$ una successió de v.a.i.i.d. amb moment de segon ordre finit. Sigui m el valor de $E(X_1)$. Es compleix que $\lim_n \frac{S_n}{n} = m$, en mitjana quadràtica.*

Teorema 5.1.4 (Llei feble de Bernoulli). *Sigui $(X_n)_n$ una successió de variables aleatòries independents, totes amb la distribució $B(1,p)$, $p \in (0,1)$. La successió $(\frac{S_n}{n})_n$ convergeix cap a p quan $n \to \infty$ en el sentit de la convergència en probabilitat.*

El teorema anterior, 5.1.4, 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 dona el nombre de vegades que ha aparegut el resultat A en les n repeticions de l'experiència. La llei de S_n és $B(n,p)$. Per tant, la freqüència relativa de l'esdeveniment A, S_n, s'acosta, en el sentit de la convergència en probabilitat (i també en mitjana quadràtica) a la probabilitat de A.

5.2 Lleis fortes dels grans nombres

Teorema 5.2.1 (Llei forta dels grans nombres per a variables incorrelacionades). *Sigui $(X_n)_n$ una successió de variables aleatòries incorrelacionades amb moments de segon ordre fitats uniformement per una constant C. Aleshores, en la convergència* QS *tenim:*

$$\lim_{n\to\infty} \frac{S_n - E(S_n)}{n} = 0.$$

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 $\lim_n \frac{S_n}{n} = 0$. La propietat d'incorrelació ($E(X_iX_j) = 0$ amb $i \neq j$) implica que:

$$E(S_n^2) = E((X_1 + \cdots + X_n)^2) = E(X_1^2) + \cdots + E(X_n^2) \leq nC.$$

Tenim, per la desigualtat de Txebitxev, per a tot $\varepsilon > 0$ i $n \geq 1$,

$$P\left(\left|\frac{S_{n^2}}{n^2}\right|\right) = P(|S_{n^2}| > n^2\varepsilon) \leq \frac{1}{n^4\varepsilon^2}E(|S_{n^2}|^2) \leq \frac{Cn^2}{n^4\varepsilon^2} = \frac{C}{n^2\varepsilon^2}.$$

Per tant, $\sum_n P(|S_{n^2}| > n^2\varepsilon) \leq \sum_n \frac{C}{n^2\varepsilon^2} < \infty$. Aleshores, pel primer lema de Börel-Cantelli tenim $P(\overline{\lim}_n\{|S_{n^2}| > n^2\varepsilon\}) = 0$ o, equivalentment $P(\underline{\lim}_n\{|S_{n^2}| \leq n^2\varepsilon\}) = 1$. Per tant, tenim $P(\overline{\lim}_n\{\frac{|S_{n^2}|}{n^2} > \varepsilon\}) = 0$, de manera que $(\frac{S_{n^2}}{n^2})_n \to 0$ quan $n \to \infty$, QS.

Es tracta ara de *controlar* els termes de la successió que corresponen a índexs k entre n^2 i $(n+1)^2$. Així doncs, definim $D_n = \max_k |S_k - S_{n^2}|$, tal que $k \in [n^2, (n+1)^2]$. Es compleix que:

$$\begin{aligned}
E(D_n^2) &= \sum_{k=n^2}^{(n+1)^2} E|S_k - S_{n^2}|^2 = \sum_{k=n^2}^{(n+1)^2} E(X_{n^2+1} + X_{n^2+2} + \cdots + X_k)^2 \\
&= \sum_{k=n^2}^{(n+1)^2} E(X_{n^2+1}^2 + X_{n^2+2}^2 + \cdots + X_k^2) \leq \sum_{k=n^2}^{(n+1)^2} E(X_{n^2+1}^2 + X_{n^2+2}^2 + \cdots + X_k^2 + \cdots + X_{(n+1)^2}^2) \\
&= (2n+2)E(S_{(n+1)^2} - S_{n^2})^2 \leq (2n+2)^2C.
\end{aligned}$$

Per tant, tenim que $P(D_n > n^2\varepsilon) \leq \frac{1}{n^4\varepsilon^2}E(|D_k|^2) \leq \frac{(2n+2)C}{n^4\varepsilon^2}$. Els mateixos arguments que hem usat abans ens permeten dir que $\lim_n \frac{D_n}{n^2} = 0$. Sigui, doncs, $n^2 \leq k < (n+1)^2$. Aleshores:

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

Com que $\frac{|S_{n^2}|+D_n}{n^2} \to 0$ quan $n \to \infty$, ja hem acabat. ∎

Demostració alternativa, versió 2018. Siguin $(X_n)_n$ variables incorrelacionades per ales quals existeix $C > 0$ tal que $E(X_n^2) \leq C$, per a tot $n \geq 1$. Sigui $S_n = X_1 + \cdots + X_n$. Aleshores, $\lim_{n\to 0} \frac{S_n - E(S_n)}{n} = 0$ en mitjana quadràtica; és a dir,

$$\lim_{n\to\infty} E\left(\left|\frac{S_n - E(S_n)}{n}\right|^2\right) = \lim_{n\to\infty} E\left(\left|\frac{(X_1 + \cdots + X_n) - (E(X_1) + \cdots + E(X_n))}{n}\right|^2\right) = 0.$$

Per a provar això utilitzem que les correlacions són zero; és a dir, per a $j \neq k$:

$$\mathrm{Cov}(X_j, X_k) = E\left((X_j - E(X_j)) \cdot (X_k - E(X_k))\right) = 0.$$

Tenim, doncs:

$$E\left(\left|\frac{(X_1 + \cdots + X_n) - (E(X_1) + \cdots + E(X_n))}{n}\right|^2\right) = \frac{1}{n^2}E\left(\left|\sum_{j=1}^n (X_j - E(X_j))\right|^2\right)$$
$$= \frac{1}{n^2}\sum_{j=1}^n E\left((X_j - E(X_j))^2\right) = \frac{1}{n^2}\sum_{j=1}^n E(X_j^2) - (E(X_j))^2 \leq \frac{1}{n^2}nC = \frac{C}{n}.$$

I d'aquí, deduïm el resultat. ∎

Teorema 5.2.2. *Sigui $(X_n)_n$ una successió de variables aleatòries i.i.d.:*

1. *Suposem que $E(|X_1|) < \infty$. Aleshores, $\lim_n \frac{S_n}{n} = E(X_1)$, QS.*
2. *Recíprocament, si $E(|X_1|) = \infty$. Aleshores, $\limsup_n \frac{S_n}{n} = \infty$, QS.*

Proposició 5.2.3 (Desigualtat de Kolmogorov). *Siguin $X_1, \ldots, X_n$ variables aleatòries independents, de quadrat integrable i centrades. Posem $S_k = X_1 + \cdots + X_k$, $1 \leq k \leq n$. Aleshores, per a tot $\varepsilon > 0$ tenim:*

$$P\left(\max_{k\in[1,n]} |S_k| > \varepsilon\right) \leq \frac{\sigma^2(S_n)}{\varepsilon^2}.$$

Proposició 5.2.4. *Sigui $(X_n)_n$ una successió de variables aleatòries independents, de quadrat integrable i centrades. Si la sèrie $\sum_n \sigma^2(X_n)$ convergeix, aleshores la sèrie $\sum_n X_n$ convergeix QS.*

Teorema 5.2.5. *Sigui $(X_n)_n$ una successió de variables aleatòries independents, de quadrat integrable i centrades. Considerem una successió $(a_n)_n$ de nombres reals tal que $0 < a^n \nearrow \infty$. Aleshores,*

$$\sum_n \frac{\sigma^2(X_n)}{a_n^2} < \infty \implies \frac{S_n}{a_n} \to 0,\ \text{QS},\ n \to \infty.$$

Observació 5.2.6. El cas particular de la successió definida per $a_n = n$, 5.2.5 estableix que la condició $\sum_n \frac{\sigma^2(X_n)}{n^2} < \infty$ és suficient per assegurar la convergència quasi segura $\frac{S_n}{n} \to 0$, quan $n \to \infty$. En particular, si $(X_n)_n$ 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 $\lim_{n\to\infty} \frac{S_n}{n} = m$, QS.

Lema 5.2.7. *Per a tota variable aleatòria Y es compleix que:*

$$\sum_{n=1}^{\infty} P(|Y| \geq n) \leq E(|Y|) \leq 1 + \sum_{n=1}^{\infty} P(|Y| \geq n).$$

En altres paraules, la variable Y té esperança finita si, i només si, la sèrie $\sum_n P(|Y| \geq n)$ és convergent.

VI

El teorema del límit central

El teorema del límit central es refereix al comportament asimptòtic, en la convergència *en distribució*, de la successió $(S_n)_n$ de les sumes parcials d'una família $(X_n)_n$ de variables aleatòries independents, convenientment normalitzades. Seguirem, doncs, estudiant el comportament en el límit de sumes de variables aleatòries independents. Barroerament, 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, estandaritzades pel factor $\sqrt{n}$.

6.1 Convergència de la llei binomial

Sigui $p \in (0,1)$, $q = 1-p \in (0,1)$; definim, per a $k = 0, 1, \ldots, n$:

$$P_n(k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Considerem una successió de variables aleatòries $(X_n)_n$ amb distribució $B(n, p_n)$ on $(p_n)_n$ és una successió en $[0,1]$ tal que $\lim_n np_n = \lambda$, per a $\lambda > 0$. Aleshores:

$$\lim_{n\to\infty} P(X_n = k) = \frac{\lambda^k}{k!} e^{-\lambda}.$$

Aquest resultat significa que una variable aleatòria amb llei Poiss(λ), s'obté com el límit en llei d'una successió de distribucions binomials com les descrites anteriorment, i s'anomena convergència de la binomial cap a la llei de Poisson. Aquí estem interessats en la convergència de $(P_n)_n$ en altres circumstàncies.

Teorema 6.1.1. *Considerem una successió $(X_n)_n$ 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 que:*

$$\lim_{n\to\infty} P\left(a < \frac{S_n - np}{\sqrt{np(1-p)}} \leq b\right) = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx\,.$$

El resultat expressa una convergència en llei o en distribució.

6.2 El teorema del límit central de Lévy-Lindeberg

L'objectiu final d'aquesta secció és demostrar el teorema del límit central de Lévy-Lindeberg.

Definició 6.2.1 (Funció generatriu de moments). Sigui $X : \Omega \longrightarrow \mathbb{R}$ una variable aleatòria. Definim una aplicació φ de la següent manera:

$$\begin{aligned} \varphi : \quad \mathbb{R} &\longrightarrow \mathbb{C} \\ t &\longmapsto \varphi_X(t) = E(e^{itX}) = E(\cos tX) + iE(\sin(tX)) = \int_{\mathbb{R}} e^{itX} f(X)\, dx\,. \end{aligned}$$

- Si X és discret: $\varphi(t) = \sum_{k\in I} e^{itk} P(X = k)$.
- Si X és absolutament contínua: $\varphi(t) = \int_{\mathbb{R}} e^{itX} f(X)\, dx$.

Podem generalitzar aquest concepte a vectors aleatoris: sigui $X : \Omega \longrightarrow \mathbb{R}^n$ i $\varphi : \mathbb{R}^n \longrightarrow \mathbb{C}$ definida de manera anàloga a l'anterior:

$$\begin{aligned} \varphi : \quad \mathbb{R}^n &\longrightarrow \mathbb{C} \\ t &\longmapsto \varphi_X(t) = E(e^{i\langle t,X\rangle}) \end{aligned}$$

Propietat 6.2.2. *Algunes propietats d'aquesta φ són les següents:*

1. $\varphi(0) = E(1) = 1$ *per a tot X variable aleatòria.*
2. $|\varphi(t)| \leq E(e^{itX}) \leq 1$ *i* $\varphi(-t) = \overline{\varphi(t)}$.
3. P_X *és simètrica si, i només si,* φ_X *és real.*
4. φ_X *és uniformement contínua.*
5. *Sigui* $X : \Omega \longrightarrow \mathbb{R}^n$, $A_{n\times n}$, $b \in \mathbb{R}^n$. *Aleshores,* $\varphi_{AX+b}(t) = e^{i\langle t,b\rangle}\varphi_X(A^*t)$.
6. ***(Propietat d'injectivitat)*** $P_X = P_Y$ *si, i només si,* $\varphi_X = \varphi_Y$.
7. *Siguin* $X_1, \dots, X_n$ *variables aleatòries independents i idènticament distribuïdes. Aleshores,* $\varphi_{X_1+\cdots+X_n}(t) = \prod_{i=1}^n \varphi_{X_i}(t)$, $t \in \mathbb{R}$.

Demostració. Copiades *a grosso modo* de les infames classes de teoria.

4. Veurem que $|\varphi(t) - \varphi(s)|$ quan $t - s \to 0$:

$$\begin{aligned} |\varphi(t) - \varphi(s)| &= |E(e^{itX}) - E(e^{isX})| = |E(e^{itX} - e^{isX})| = |E(e^{isX}(e^{i(t-s)X} - 1)| \\ &\leq E(|e^{isX}(e^{i(t-s)X} - 1)|) \leq E(|e^{i(t-s)X} - 1| \to 0,\ t - s \to 0. \end{aligned}$$

Això és perquè $e^{i(t-s)X} - 1 \to 0$ quan $t - s \to 0$.

5. Pel que sembla, en tenim prou amb la següent cadena d'igualtats:

$$\varphi_{AX+B}(t) = E(e^{i\langle t,AX+B\rangle} = e^{i\langle t,b\rangle}E(e^{i\langle t,AX\rangle}) = e^{i\langle t,b\rangle}.$$

7. Per provar la igualtat ens cal demostrar necessari i suficient.

$\boxed{\Rightarrow}$ Una cadena d'igualtats:

$$\begin{aligned} P_X(t) &= \int_{\mathbb{R}^n} e^{i\langle t,X\rangle} f(x)\, dx = \int_{\mathbb{R}^n} e^{i\sum_j t_j X_j} f_{X_i}(x_i) \cdot f_{X_n}(x_n)\, dx \\ &= \int_{\mathbb{R}} e^{it_1x_1} f_{X_1}(x_1)\, dx_1 \cdots \int_{\mathbb{R}} e^{it_nx_n} f_{X_n}(x_n)\, dx_n = \prod_{i=1}^n \varphi_{X_i}(t_i). \end{aligned}$$

$\boxed{\Leftarrow}$ Una (altra) cadena d'igualtats:

$$\int_{\mathbb{R}^n} e^{i\langle t,X\rangle} f(x)\, dx = \int_{\mathbb{R}^n} \prod_{j=1}^{n} e^{it_jX_j} f_{X_i}(x_i)\, dx_i = \int_{\mathbb{R}^n} e^{i\langle t,X\rangle} f_{X_1}(x_1) \cdots f_{X_n}(x_n)\, dx_1 \cdots dx_n\,.$$

D'aquesta manera, queda vist que són independents.

I amb això ja hem acabat. ■

Teorema 6.2.3 (del límit central de Lévy-Lindeberg)**.** *Sigui (Z_n) 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:*

$$\mathcal{L} - \lim_{n\to\infty} \frac{S_n - nm}{\sigma\sqrt{n}} = Y,$$

on Y és una variable aleatòria amb distribució $\mathcal{N}(0,1)$.

Exemple 6.2.4. 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 dona 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.

Demostració. 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:

$$\Phi\left(0.04\sqrt{\frac{n}{p(1-p)}}\right) \geq \frac{1+0.95}{2}. \tag{6.1}$$

Com que $p(1-p) \leq \frac{1}{4}$, si n compleix que $\Phi(0.08\sqrt{n}) \geq 0.975$ també tindrem (6.1). Aquesta darrera condició dona n de l'ordre de 600. En la deducció de (6.1) hem usat l'expressió $\Phi(x) - \Phi(-x) = 2\Phi(x) - 1$, que s'obté per la simetria de la llei normal estàndard. ■

Bibliografia

[San99] Marta Sanz-Solé. *Probabilitats*. cat. UB; 28. Barcelona: Edicions Universitat de Barcelona, 1999. isbn: 8483380919.

[Wik23] Wikipedia contributors. *Infinite monkey theorem — Wikipedia, The Free Encyclopedia*. [Online; accessed 23-December-2023]. 2023. url: `https://en.wikipedia.org/w/index.php?title=Infinite_monkey_theorem`.