

Apunts de teoria de la probabilitat

ALEIX TORRES I CAMPS

ANNA DE MIER (ANNA.DE.MIER@UPC.EDU), GUILLEM PEREARNAU I SONIA PEREZ

Índex

1	Espais de probabilitat	2
1.1	Motivació	2
1.2	Experiments i probabilitat	2
1.3	La probabilitat condicionada	3
1.4	Independència	4
1.5	Espais productes	5
1.6	Lemes de Borel-Cantelli	5
2	Variables aleatòries	6
2.1	Definició i distribució	6
2.2	Moments d'una v.a.	6
2.3	Vectors de v.a.	8
2.4	Independència de v.a.'s	8
2.5	Covariància i (una) llei dels grans nombres	8
3	V.a Discretes	9
3.1	Funció de probabilitat, esperança i independència	9
3.2	Funció generadora de probabilitat	10
4	V.a Contínues	10
5	Funcions característiques i famílies exponencials	10
6	Convergència de variables aleatòries	10

1 Espais de probabilitat

1.1 Motivació

L'objectiu de la teoria de la probabilitat és trobar models per a fenòmens que depenen de l'atzar (no deterministes), cada realització d'un fenomen en direm experiment, del qual n'obtidrem un resultat. A més, tindrem els successos (observables) que son totes les preguntes raonables que ens podem fer.

1.2 Experiments i probabilitat

Definició 1. Un experiment és un parell (Ω, \mathcal{A}) on Ω és un conjunt i $\mathcal{A} \subset \mathcal{P}(\Omega)$ tal que:

1. $\emptyset \in \mathcal{A}$
2. $A \in \mathcal{A} \implies A^c \in \mathcal{A}$
3. Si $\{A_n\}_{n \geq 1}$ és una col·lecció numerables d'elements de $\mathcal{A} \implies \bigcup_{n \geq 1} A_n \in \mathcal{A}$

Exemple 1. Uns quants exemples...

Volem una funció que assigni probabilitats de successos, és a dir, $P: \mathcal{A} \rightarrow \mathbb{R}$. Llavors definim:

Definició 2. Un espai de probabilitat és una terna (Ω, \mathcal{A}, P) on:

1. (Ω, \mathcal{A}) és un experiment.
2. $P: \mathcal{A} \rightarrow \mathbb{R}$ tal que: $P(\emptyset) = 0$, $P(A) \geq 0$, $\forall A \in \mathcal{A}$. Si $\{A_n\}_{n \geq 1}$ és una col·lecció de successos dos a dos dijunts $\implies P(\bigcup_{n \geq 1} A_n) = \sum_{n \geq 1} P(A_n)$.
3. $P(\Omega) = 1$.

Per tant, la probabilitat és una mesura a (Ω, \mathcal{A}) normalitzada a 1. A P se l'anomena funció de probabilitat.

Exemple 2. Espai discret, si Ω és numerable i $\mathcal{A} = \mathcal{P}(\Omega)$. Si $\Omega = \{w_1, w_2, w_3, \dots\}$ prenem $\sum_{i \geq 1} p_i = 1$ (amb $p_i \geq 0$) i definim $P(A) = \sum_{w_i \in A} P(\{w_i\})$, alleugerint la notació podem fer servir $P(\{w_i\}) = P(w_i) = p_i$.

Exemple 3. Espai clàssic, és un espai discret amb $|\Omega| = N$ i $p_i = 1/N$. "Çassos favorables entre cassos possibles": $P(A) = \frac{|A|}{N}$.

Exemple 4. Espais clàssics amb monedes o daus, tot ben repartit.

Exemple 5. Durada d'un mòbil? $\Omega = (0, \infty)$ o bé, $(0, L]$. Si $\mathcal{A} = \mathcal{P}(\Omega)$ sabem que no podem assignar-hi una mesura. Però com ens interessin els intervals com (a, b) , agafe, la σ -àlgebra que conté tots els intervals (oberts). Aquí apareixen els burelians $\mathcal{B} = \sigma(I)$ i podem agafar la mesura de Lebesgue a \mathbb{R} . En resum, $\Omega = (0, L)$, $\mathcal{B} = \sigma(I)$ i $P(B) = \frac{\mu(B)}{L}$. On $\mu(B)$ és la seva mesura de Lebesgue. Tot i així, no és realístic perquè és massa uniforme.

Proposició 3. Propietats d'espais de probabilitat. Sigui (Ω, \mathcal{A}, P) :

1. Per $r \geq 2$, si $A_1, \dots, A_r \in \mathcal{A}$, $A_i \cap A_j = \emptyset$ si $i \neq j$ llavors $P(\bigcup_{i=1}^r A_i) = \sum_{i=1}^r P(A_i)$
2. Si $A, B \in \mathcal{A}$ i $A \subset B \implies P(B \setminus A) = P(B) - P(A)$ i $P(A) \leq P(B)$.
3. $P(A^c) = 1 - P(A)$, $\forall A \in \mathcal{A}$
4. (Desigualtat de Boole) Si $A_1, \dots, A_r \in \mathcal{A} \implies P(\bigcup_{i=1}^r A_i) \leq P(A_1) + \dots + P(A_r)$

Demostració.

1. En els cassos finits, per $r < k$, cal agafar $A_k = \emptyset$, ja que així, com que $P(\emptyset) = 0$, $P(\bigcup_{1 \leq n \leq r} A_n) = P(\bigcup_{1 \leq n \leq k} A_n) = \sum_{1 \leq n \leq k} P(A_n) = \sum_{1 \leq n \leq r} P(A_n)$.

2. Primer de tot $B \setminus A \in \mathcal{A}$, ja que $B \setminus A = (B^c \cup A)^c \in \mathcal{A}$. Després, reordenant el fet que $P(A) + P(B \setminus A) = P(B)$, ens queda el que volíem. Com les propietats són positives, la desigualtat es demostra automàticament.
3. De $P(A) + P(A^c) = P(\Omega) = 1$ obtenim l'expressió de l'enunciat.
4. Ho anem a fer per inducció sobre r . Clarament per $r = 1$ és cert, suposem que ho és per $r - 1$, anem a veure-ho per a un r arbitrari. Sigui $B = (\bigcup_{i=1}^{r-1} A_i) \cap A_r$, llavors $P(\bigcup_{i=1}^r A_i) = P((\bigcup_{i=1}^{r-1} A_i) \cup A_r) = P((\bigcup_{i=1}^{r-1} A_i) \cup (A_r - B_r)) = P((\bigcup_{i=1}^{r-1} A_i)) + P(A_r - B_r) = [\text{per hipòtesi i per 2}] \leq P(A_1) + \dots + P(A_{r-1}) + P(A_r)$ que és la desigualtat de Boole.

□

Proposició 4. *Successions monòtones:*

Si $A \in \mathcal{A}$, $i \geq 1$ i $A_1 \subset A_2 \subset A_3 \subset \dots$, aleshores $P(\bigcup_{i \geq 1} A_i) = \lim_{i \rightarrow \infty} P(A_i)$.

Si $A \in \mathcal{A}$, $i \geq 1$ i $A_1 \supset A_2 \supset A_3 \supset \dots$, aleshores $P(\bigcap_{i \geq 1} A_i) = \lim_{i \rightarrow \infty} P(A_i)$.

Demostració. Fem $B_1 = A_1$, $B_2 = A_2 \setminus A_1$, $B_3 = A_3 \setminus A_2$, ... Aleshores, $B_i \cap B_j = \emptyset$ per $i \neq j$ i $A_i = \bigcup_{j=1}^i B_j$. Per tant:

$$P(\bigcup_{n \geq 1} A_n) = P(\bigcup_{n \geq 1} B_n) = \sum_{n \geq 1} P(B_i) = \lim_{N \rightarrow \infty} \sum_{n=1}^N P(B_n) = \lim_{N \rightarrow \infty} P(A_N)$$

L'altre és demostra passant al complementari.

□

Teorema 5. *Siguin $A_1, \dots, A_r \in \mathcal{A}$. Per $I \subset [r]$, posem: $A_I = \bigcap_{i \in I} A_i$ i $S_k = \sum_{I \subset [r] \mid |I|=k} P(A_I)$. Aleshores,*

$$P(\bigcup_{i=1}^r A_i) = \sum_{k=1}^r (-1)^{k+1} S_k$$

Demostració. Per inducció, el cassos $r = 1, 2$ són fàcils. Aleshores, pel cas inductiu fa falta el cas $r - 1$ i el cas 2.

□

Proposició 6. *Desigualtats de Bonferroni.* *Sigui $M_T = \sum_{k=1}^T (-1)^{k+1} S_k$. Aleshores, si:*

1. T és senar $\implies P(\bigcup_{i=1}^r A_i) \leq M_T$.

2. T és parell $\implies P(\bigcup_{i=1}^r A_i) \geq M_T$.

Demostració. La demostració per inducció és semblant a l'anterior.

□

1.3 La probabilitat condicionada

Sigui (Ω, \mathcal{A}, P) un espai de probabilitat. Prenem $B \in \mathcal{A}$ amb $P(B) > 0$. Volem recalculer la probabilitat P dels successos sabent que ha passat B .

Definició 7. Si $B \in \mathcal{A}$ amb $P(B) > 0$ i $A \in \mathcal{A}$, la probabilitat de A condicionada a B és

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

Observació 8.

1. $P(A|B)$, a priori, pot ser major o menor a $P(A)$.

2. Fixat B , $P_B : \mathcal{A} \rightarrow \mathcal{R}$ definida com $P_B(A) = \frac{P(A \cap B)}{P(B)}$, dona una funció de probabilitat en (Ω, \mathcal{A}) . (També ho és en $(\Omega, \mathcal{A}_B = \{A \cap B : A \in \mathcal{A}\})$).

Sigui ara B_1, \dots, B_n una partició de Ω (amb $B_i \in \mathcal{A}$ i $P(B) > 0$). Llavors, la llei de les probabilitats totals és

$$P(A) = P(A \cap \Omega) = P(A \cap (\bigcup B_i)) = P(\bigcup (A \cap B_i)) = \sum_{i=1}^n P(A \cap B_i) = \sum_{i=1}^n P(A|B_i)P(B_i)$$

Exemple 6. Ruïna del jugador (Huygens S.XVII). Sigui J un jugador que comença amb un capital de $k \geq 1$, el seu objectiu és arribar a $N \geq k$ i s'arruïna si arriba a 0. En cada torn guanya 1 amb probabilitat $1/2$ i perd 1 amb probabilitat $1/2$.

Sigui $B =$ "a la 1a jugada, +1" i $R_k =$ "s'ha arruïnat amb capital inicial k ". Aleshores,

$$P(R_k) = P(R_k|B)P(B) + P(R_k|B^c)P(B^c) = \frac{1}{2}P(R_{k+1}) + \frac{1}{2}P(R_{k-1})$$

Definint $p_k = P(R_k)$, aleshores, obtenim l'equació de recurrència: $p_k = \frac{1}{2}p_{k+1} + \frac{1}{2}p_{k-1}$ amb $p_0 = 1$ i $p_N = 0$.

Per resoldre'l, fem $\frac{1}{2}(p_k - p_{k-1}) = \frac{1}{2}(p_{k+1} - p_k)$, definim $b_k = p_k - p_{k-1}$ que per la pròpia recurrència es compleix que $p_k = p_{k-t} + tb_{k-t+1}$. I veiem que la solució és $p_k = 1 - \frac{k}{N}$ que es pot comprovar per inducció.

Pregunta: Qui era (Ω, \mathcal{A}, P) ?

Proposem $\Omega = \{w_1, w_2, w_3, \dots\} : w_i \in \{0, 1\}$, a cada successió li associem un real (no és injectiva, però només es repeteix en alguns racionals). $P(A) = \mu(\phi(A))$ on μ és la mesura de Lebesgue a $[0, 1]$ i ϕ passa de (w_1, w_2, \dots) a $0.w_1w_2\dots$, llavors \mathcal{A} és el conjunt de conjunts que van a parar a borelians per ϕ .

Teorema 9. Bayes. $P(B_i|A) = \frac{P(A|B_i)P(B_i)}{P(A)} = \frac{P(A|B_i)P(B_i)}{\sum P(A|B_i)P(B_i)}$

Demostració. Bé del fet que $P(A \cap B_i) = P(B_i \cap A)$ i escrivint la definició de probabilitat condicionada de dues maneres diferents. I després, utilitzant el lemma de les probabilitats totals. \square

1.4 Independència

R.Durrett "Aquí acaba la teoria de la mesura i comença la probabilitat".

Definició 10. En (Ω, \mathcal{A}, P) , dos successos $A, B \in \mathcal{A}$ són independents si $P(A \cap B) = P(A) \cdot P(B)$.

Observació 11. Si $P(B) > 0$, $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)P(B)}{P(B)} = P(A)$.

Exercici: Proveu que el conjunt buit i el total són independents amb qualsevol altre succés.

Demostració. La probabilitat del buit sempre és 0 i si intersequem quelcom amb el buit sempre el dona el buit, per tant, sempre es compleix que $0 = P(\emptyset)P(A) = P(\emptyset \cap A) = 0$.

Com que la probabilitat del total és 1, succeix el següent: $P(A)P(\Omega) = P(A) = P(A \cap \Omega)$. \square

Exercici: Proveu que si A i B són independents \implies que amb o sense complementaris també ho són.

Demostració. Només cal comprovar que A^c i B són independents, perquè veure que A i B^c són independents és el raonament simètric i per obtenir que A^c i B^c cal fer el mateix raonament dues vegades.

Simplement fem $P(A^c)P(B) = (1 - P(A))P(B) = P(B) - P(A)P(B) = P(B) - P(A \cap B) = P(A^c \cap B)$. \square

Definició 12. Si $\{A_i\}_{i \in I}$ és una col·lecció de successis, són independents si $\forall J \subseteq I, |J| < \infty, P(\bigcap_{j \in J} A_j) = \prod_{j \in J} P(A_j)$.

1.5 Espais productes

Tenim $(\Omega_1, \mathcal{A}_1, P_1)$ i $(\Omega_2, \mathcal{A}_2, P_2)$ dos espais de probabilitat. Volem un espai en $\Omega_1 \times \Omega_2$ tal que $P(A_1 \times A_2) = P_1(A_1)P_2(A_2)$ ($\forall A_i \in \mathcal{A}_i$ per $i = 1, 2$). La teoria de la mesura ens diu que es pot... I la σ -àlgebra. Com volem garantir que hi hagi $\mathcal{A}_1 \times \mathcal{A}_2 = \{A_1 \times A_2 : A_i \in \mathcal{A}_i, i = 1, 2\}$. Tot i així, això no tñe perquè ser un σ -àlgebra. Llavors agafem la més petita que ho contingui. Com a \mathcal{A} agafem $\sigma(\mathcal{A}_1 \times \mathcal{A}_2)$ és a dir, la generada. La probabilitat? Volem que $P(A_1 \times A_2) = P(A_1)P(A_2)$.

Definició 13. Una àlgebra i una premesura són una σ -àlgebra i una mesura (respectivament, però només garanteix la unió finita i la suma finita).

Teorema 14. Teorema d'extensió. Sigui p_0 una premesura en una àlgebra \mathcal{A}_0 . Aleshores existeixen una σ -àlgebra $\mathcal{A}^* \subset \sigma(\mathcal{A}_0)$ i una mesura p^* tal que p^* coincideix amb p_0 en \mathcal{A}_0 . A més, si p_0 és finita, p és única.

Com s'aplica?

$\mathcal{A}_0 = \{A \subset \Omega_1 \times \Omega_2 \mid \text{on } A \text{ és unió finita d'elements de } \mathcal{A}_1 \times \mathcal{A}_2\}$. (Caldria comprovar que és àlgebra i la unió es pot prendre disjunta i finita).

Exemple 7. Problema de l'agulla del comte Buffon (s. XVIII). I es pregunta, quina és la probabilitat que l'agulla talli alguna de les línies?

Sigui d la distància del punt mig a la recta més propera i sigui α l'angle amb la vertical. Aleshores, la longitud del catet $\frac{l}{2} \sin \alpha$. Tallarà si $d \leq \frac{l}{2} \sin \alpha$.

Així que com a espai mostral podem agafar $\Omega = [0, \pi) \times [0, \frac{l}{2}]$ la primera correspon a α i la segona a d . Amb la mesura de Lebesgue (volem uniformitat). Que finalment, calculem l'àrea sota la curva $l/2 \sin \alpha$ i dividim pel total, dona $\frac{2l}{L\pi}$.

1.6 Lemes de Borel-Cantelli

Sigui (Ω, \mathcal{A}, P) un espai de probabilitat i siguin $\{A_n\}_{n \geq 1}$ una col·lecció numerable de successos.

Definició 15. Sigui $\limsup A_n = \bigcap_{n \geq 1} \bigcup_{k \geq n} A_k$ i $\liminf A_n = \bigcup_{n \geq 1} \bigcap_{k \geq n} A_k$. Són d' \mathcal{A} perquè fem interseccions i unions numerables.

Nota 8. Si $w \in \Omega$ i $w \in \limsup A_n \iff \forall n \geq 1, w \in \bigcup_{k \geq n} A_k \iff \forall n \geq 1 \exists k \geq n, w \in A_k \iff w$ pertany a infinits dels A_n .

Nota 9. Fent el mateix amb \liminf , tenim $w \in \liminf A_n \iff w$ pertany a tots els A_n a partir d'algun en endavant.

Nota 10. Queda clar que $\liminf A_n \subseteq \limsup A_n$.

Lema 16. (Borel-Cantelli 1) Si $\sum_{n \geq 1} P(A_n) < \infty$, aleshores $P(\limsup A_n) = 0$.

Demostració. Volem fitar $P(\limsup A_n) = P(\bigcap_{n \geq 1} \bigcup_{k \geq n} A_k) = [\text{Successions monòtones}] = \lim_{n \rightarrow \infty} P(\bigcup_{k \geq n} A_k) \leq \lim_{n \rightarrow \infty} \sum_{k \geq n} P(A_k) = 0$, perquè la cua d'una sèrie convergent tendeix a 0. \square

Pregunta natural: Què passa si $\sum_{n \geq 1} P(A_n) = \infty$?

Lema 17. (Borel-Cantelli 2) Si $\sum_{n \geq 1} P(A_n) = \infty$ i els $\{A_n\}_{n \geq 1}$ són independents (dos a dos és suficient) $\implies P(\limsup A_n) = 1$.

Demostració. $P((\limsup A_n)^c) = P(\bigcap_{n \geq 1} \bigcap_{k \geq n} A_k^c)$. Fixem-nos que això és el mateix que $\liminf A_k^c$ i que els successos estan encaixats $\bigcap_{k \geq 1} A_k^n \subseteq \bigcap_{k \geq 2} A_k^c \subseteq \dots$. Llavors, per la proposició de successos encaixats tenim que la igualtat anterior és igual a $\lim_{n \rightarrow \infty} P(\bigcap_{k \geq n} A_k^c)$, fent servir de nou la proposició dels successos encaixats tenim que $P(\bigcap_{k \geq n} A_k^c) = \lim_{r \rightarrow \infty} P(\bigcap_{k=n}^r A_k^c) = [ind.] = \lim_{r \rightarrow \infty} \prod_{k=n}^r (1 - P(A_k)) \leq \lim_{r \rightarrow \infty} \prod_{k=n}^r e^{-P(A_k)} = \lim_{r \rightarrow \infty} e^{-\sum_{k=n}^r P(A_k)} = 0$. Per tant, com que cada un dels $P(\bigcup_{k \geq n} A_k^c) = 0$, $\lim_{n \rightarrow \infty} P(\bigcap_{k \geq n} A_k^c) = 0$. \square

2 Variables aleatòries

2.1 Definició i distribució

Definició 18. L'aplicació $X : \Omega \rightarrow \mathbb{R}$ és una variable aleatòria si $\forall B \in \mathcal{B}(\mathbb{R})$, es té que $X^{-1}(B) \in \mathcal{A}$ (on \mathcal{B} són els borelians).

Observació 19. X v.a. $\iff X$ és mesurable respecte $(\mathbb{R}, \mathcal{B})$.

Observació 20. X v.a. $\iff \forall a \in \mathbb{R}, X^{-1}((-\infty, a]) \in \mathcal{A}$.

Notació: Escriurem $P(X \in B) = P(X^{-1}(B)) = P(\{w \in \Omega : X(w) \in B\})$. Igualment $P(X \geq x), P(X = x), P(X > x), \dots$

Definició 21. La funció de distribució (acumulada) d'una v.a. X és $F_X : \mathbb{R} \rightarrow [0, 1]$ que envia $x \mapsto P(X \leq x) = P(\{w \in \Omega : X(w) \leq x\})$.

Exemple 11. La variable indicadora del succés $A \in \mathcal{A}$ és $\mathbb{I}_A : \Omega \rightarrow \mathbb{R}$ que envia $w \in \Omega$ a 1 si $w \in A$ i 0 altrament. Llavors, la funció indicadora $F_{\mathbb{I}_A}$ és igual a $P(x \leq x_0) = 0$ fins arribar a $x_0 = 0$. A partir de llavors, entre $[0, 1]$ la funció de distribució acumulada és $1 - P(A)$ i, a partir de A és igual a 1.

Proposició 22. Si X és v.a. F_x satisfà:

- i) F_x és creixent.
- ii) $\lim_{x \rightarrow +\infty} F(x) = 1, \lim_{x \rightarrow -\infty} F(x) = 0$.
- iii) F_X és contínua per la dreta ($\lim_{h \rightarrow 0^+} F_X(x+h) = F_X(x)$)

Demostració.

- i) Si $x_1 \leq x_2$ $F_X(x_1) = P(X \leq x_1) = P(\{w \in \Omega : X(w) \leq x_1\}) \leq P(\{w \in \Omega : X(w) \leq x_2\}) = F_X(x_2)$
- ii) Quan fem $\lim_{x \rightarrow +\infty} F(x) = \lim_{x \rightarrow +\infty} P(X \leq x) = \lim_{x \rightarrow +\infty} P(\{w \in \Omega : X(w) \leq x\}) = P(\Omega) = 1$ i el mateix quan x tendeix a menys infinit, el conjunt tendeix a \emptyset , per tant, té probabilitat 0.
- iii) Sigui $\{h_n\}_{n \geq 1}$ una successió decreixent tal que $h_n \rightarrow 0^+$, llavors: $\lim_{n \rightarrow \infty} F_X(x+h_n) = \lim_{n \rightarrow \infty} P(\{w \in \Omega : X(w) \leq x+h_n\}) = P(\bigcap_{n \geq 1} \{w \in \Omega : X(w) \leq x+h_n\}) = P(\{w \in \Omega : X(w) \leq x\}) = F_X(x)$.

□

Nota 12. Si $F : \mathbb{R} \rightarrow [0, 1]$ tal que satisfaci i), ii) i iii), aleshores existeix (Ω, \mathcal{A}, P) i $X : \Omega \rightarrow \mathbb{R}$ v.a. tal que $F_X = F$.

Definició 23. Si tenim una v.a. X induïx una funció de probabilitat P_X en $(\mathbb{R}, \mathcal{B})$, llavors $p_X(B) = P(X \in B) \forall B \in \mathcal{B}$, l'anomenarem llei de X o mesura de probabilitat induïda per X .

2.2 Moments d'una v.a.

Una de les característiques més notables d'una variable aleatòria és la mitjana aritmètica dels valors que pren, si X pren només valors $\{a_1, \dots, a_r\}$, la "mitjana" serà $\sum_{i=1}^r a_i P(X = a_i)$. Més generalment es considera l'esperança.

Definició 24. L'esperança de X es defineix com:

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

Sempre i quan existeixi i sigui finita o convergeix absolutament.

Proposició 25. Propietats de $\mathbb{E}[X]$ (heredades de l'integral):

- 1. Si $X = c$, llavors $\mathbb{E}[X] = c$.
- 2. Si $X = \mathbb{I}_A$, llavors $\mathbb{E}[X] = P(A)$

3. Si $X \leq Y$, llavors $\mathbb{E}[X] \leq \mathbb{E}[Y]$. (X, Y són v.a. en el mateix espai de probabilitat).

4. $\mathbb{E}[X]$ és lineal.

Observació 26. Si $f : \mathbb{R} \rightarrow \mathbb{R}$ i $f(X)$ és una v.a. (mesurable), aleshores $\mathbb{E}[f(X)]$ vol dir $\int_{\Omega} f(X)dp$ (si convergeix absolutament).

Una variable aleatòria interessant és quan tenim una v.a. X i considerem la v.a. $X - \mathbb{E}[X]$. Aquesta segona direm que està centrada en el 0, perquè quan calculem l'esperança $\mathbb{E}[X - \mathbb{E}[X]] = \mathbb{E}[X] - \mathbb{E}[\mathbb{E}[X]] = 0$, dona 0 (utilitzant linealitat i esperança d'una constant). Una altra que no sempre dona 0 i és una mesura de com d'allunyada està una variable aleatòria de la seva esperança és la següent:

Definició 27. La variancia de X és

$$\text{Var}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2]$$

si existeix. La seva arrel quadrada positiva és la desviació típica (o estàndar).

Proposició 28. Propietats de la variància:

1. $\text{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$
2. $\text{Var}[c] = 0$
3. $\text{Var}[c + X] = \text{Var}[X]$
4. $\text{Var}[cX] = c^2 \text{Var}[X]$

Demostració.

1. $\text{Var}[X] = \mathbb{E}[X^2 - \mathbb{E}[X]^2 - 2X\mathbb{E}[X]] = [X^2] + [\mathbb{E}[X]^2] - 2\mathbb{E}[X]\mathbb{E}[X] = [X^2] - [\mathbb{E}[X]^2]$.
2. Perquè $c - \mathbb{E}[c] = c - c = 0$.
3. Perquè $\mathbb{E}[X + c] = [X] + c$ i $X + c - \mathbb{E}[X + c] = X + c - \mathbb{E}[X] - c = X - \mathbb{E}[X]$.
4. Com que $\text{Var}[cX] = \mathbb{E}[(cX)^2] - \mathbb{E}[cX]^2$, per linealitat de l'esperança, $= c^2\mathbb{E}[X^2] - c^2\mathbb{E}[X]^2 = c^2 \text{Var}[X]$.

□

Definició 29. El moment k -èssim de X és $\mathbb{E}[X^k]$ i el moment k -èssim centrat és $\mathbb{E}[(X - \mathbb{E}[X])^k]$ si existeixen i son absolutament convergents..

Observació 30. $(\int_{\Omega} |X|^k)^{\frac{1}{k}}$ és una norma en els espais L_p .

Proposició 31. Algunes desigualtats heredades:

1. *Cauchy-Schwartz:* Si $\mathbb{E}[X^2], \mathbb{E}[Y^2] < \infty \implies \mathbb{E}[XY]^2 \leq \mathbb{E}[X^2]\mathbb{E}[Y^2]$
2. *Hölder:* Si $\mathbb{E}[|X|^p], \mathbb{E}[|Y|^q] < \infty$ amb $p > 1$ i $\frac{1}{p} + \frac{1}{q} = 1$. Llavors $\mathbb{E}[|XY|] \leq \mathbb{E}[|X|^p]^{\frac{1}{p}} \mathbb{E}[|Y|^q]^{\frac{1}{q}}$.
3. *Minkowski:* Si $\mathbb{E}[|X|^p], \mathbb{E}[|Y|^p] < \infty$, per $p \geq 1$. Llavors $\mathbb{E}[|X + Y|^p]^{\frac{1}{p}} \leq \mathbb{E}[|X|^p]^{\frac{1}{p}} + \mathbb{E}[|Y|^p]^{\frac{1}{p}}$.

Proposició 32. Desigualtat pròpies de la T.Prob:

1. *Markov:* X v.a. $X \geq 0$, $a \in \mathbb{R}$ amb $a > 0$. Llavors $P(X \geq a) \leq \frac{\mathbb{E}[X]}{a}$
2. *Txebixev:* Sigui X una v.a. amb $\mathbb{E}[X] = \mu < \infty$, $\text{Var}[X] = \sigma^2 < \infty$. Llavors $\forall b: P(|X - \mu| \geq b) \leq \frac{\sigma^2}{b^2}$.

Demostració.

1. Sigui $A = \{w \in \Omega : X(w) \geq a\}$, aleshores tenim que $a\mathbb{I}_A \leq X$, perquè per un mateix resultat pertanyent a A segur que $X(a) \geq a$, per altres resultats que no pertanyen a A , $X \geq 0$ que ens ho assegura l'enunciat. Llavors tenim $\mathbb{E}[X] \geq \mathbb{E}[a\mathbb{I}_A] = aP(A) = aP(X \geq a)$, passant dividint l' a tenim la desigualtat de Markov.
2. Tenim que $P(|X - \mu| \geq b) = P((X - \mu)^2 \geq b^2) \leq \frac{\mathbb{E}[(X - \mu)^2]}{b^2} = \frac{\sigma^2}{b^2}$, que és la fita de Txebixev.

□

2.3 Vectors de v.a.

En un mateix espai (Ω, \mathcal{A}, P) tenim diverses v.a. X_1, X_2, \dots, X_n que volem estudiar conjuntament.

Definició 33. En (Ω, \mathcal{A}, P) , un vector aleatori (o vector de v.a. o v.a. n -dimensional) és

$$\bar{X} = (X_1, \dots, X_n) : \Omega \rightarrow \mathbb{R}^n$$

tal que $\forall B \in \mathcal{B}(\mathbb{R}^n)$ (borelians de \mathbb{R}^n) es té $(\bar{X})^{-1}(B) \in \mathcal{A}$.

Nota 13. Perquè \bar{X} sigui vector aleatori, n'hi ha prou si

$$(\bar{X})^{-1}((-\infty, a_1] \times (-\infty, a_2] \times \dots \times (-\infty, a_n]) \in \mathcal{A}$$

Que és el mateix que $\bigcap_{i=1}^n X_i^{-1}((-\infty, a_i]) \in \mathcal{A}$. És a dir, que X_1, \dots, X_n siguin v.a.

Definició 34. La funció de distribució conjunta $\bar{F} = (X_1, \dots, X_n)$ és

$$\begin{aligned} F_{\bar{X}} : \mathbb{R}^n &\rightarrow [0, 1] \\ (x_1, \dots, x_n) &\mapsto P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n) \end{aligned}$$

Que és $\{w \in \Omega : X_1(w) \leq x_1, \dots, X_n(w) \leq x_n\}$.

Proposició 35. Propietats de les funcions de distribució $F_{(X,Y)}(x, y)$:

1. Si $x_1 \leq x_2$ i $y_1 \leq y_2$ llavors, $F_{(X,Y)}(x_1, y_1) \leq F_{(X,Y)}(x_2, y_2)$.
2. $\lim_{(x,y) \rightarrow (-\infty, -\infty)} F_{(X,Y)}(x, y) = 0$ i $\lim_{(x,y) \rightarrow (+\infty, +\infty)} F_{(X,Y)}(x, y) = 1$.
3. $\lim_{y \rightarrow \infty} F_{(X,Y)}(x, y) = P(X \leq x) = F_X(x)$, funció de distribució marginal.
4. És contínua per dalt: $\lim_{(h_1, h_2) \rightarrow (0,0)} F_{(X,Y)}(x + h_1, y + h_2) = F_{(X,Y)}(x, y)$.
5. Propietat dels rectangles. $P(a < X \leq b, c < Y \leq d) = P(X \leq b, Y \leq d) - P(X \leq a, Y \leq d) + P(X \leq a, Y \leq c)$.

Nota 14. Si $F : \mathbb{R} \rightarrow [0, 1]$ tal que satisfaci 1., 2., 4. i 5. (la 3. és de regal) caracteritzen les funcions de distribució bivariades.

2.4 Independència de v.a.'s

Tenim $F_{\bar{X}}(x_1, \dots, x_n) = P(X_1 \leq x_1, \dots, X_n \leq x_n)$ que son successos d' \mathcal{A} . Si són independents, llavors és igual a producte de probabilitats $= P(X_1 \leq x_1) \cdots P(X_n \leq x_n)$, en tal cas direm que les v.a. son independents.

Definició 36. Si $\forall x_1, \dots, x_n$ es té que:

$$F_{\bar{X}}(x_1, \dots, x_n) = F_{X_1}(x_1) \cdots F_{X_n}(x_n)$$

direm que X_1, \dots, X_n són independents.

Nota 15. Si tenim v.a. independents: $\forall B_1, \dots, B_n \in \mathcal{B}$ es té $P(X_1 \in B_1, \dots, X_n \in B_n) = P(X_1 \in B_1) \cdots P(X_n \in B_n)$.

2.5 Covariància i (una) llei dels grans nombres

Segui $\bar{X} = (X_1, \dots, X_n)$ vector aleatori, podem considerar el vector d'esperances $\mathbb{E}[\bar{X}] = (\mathbb{E}[X_1] \cdots \mathbb{E}[X_n])^t$.

Definició 37. La covariància de X i Y (en el mateix espai)

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

I el coeficient de correlació (de Pearson) $\rho_{X,Y} = \text{Var}(X, Y) / \sigma_X \sigma_Y$.

Observació 38. El coeficient de correlació està sempre -1 i 1. Es pot demostrar amb Cauchy-Schwartz:

$$|\text{Var}(X, Y)| \leq \sqrt{\mathbb{E}[(X - \mathbb{E}[X])^2] \mathbb{E}[(Y - \mathbb{E}[Y])^2]} = \sqrt{\text{Var}[X] \text{Var}[Y]} = \sigma_X \sigma_Y$$

Llavors $|\rho_{X,Y}| \leq 1$. X i Y són incorrelades si $\text{Var}(X, Y) = 0$.

Definició 39. La matriu de covariàncies és:

$$\text{Var}(\bar{X}) = \begin{pmatrix} \text{Var}[X_1] & \cdots & \text{Var}(X_1, X_n) \\ \vdots & \ddots & \vdots \\ \text{Var}(X_n, X_1) & \cdots & \text{Var}[X_n] \end{pmatrix}$$

Alguns càlculs:

$$\text{Var}(X, Y) = \mathbb{E}[XY - X\mathbb{E}[Y] - Y\mathbb{E}[X] + \mathbb{E}[X]\mathbb{E}[Y]] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

Nota 16. Si X i Y independents, llavors són incorrelades ($\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$). En general, surt del teorema de Fubini. El recíproc en general no és cert.

Proposició 40. Es té que:

1. $\text{Var}(c, X) = 0$
2. $\text{Var}(aX + bY, Z) = a \text{Var}(X, Z) + b \text{Var}(Y, Z)$
3. $\text{Var}(X + Y) = \text{Var}[X] + \text{Var}[Y] + 2 \text{Var}(X, Y)$
4. Si X_1, \dots, X_n són 2 a 2 incorrelades, llavors $\text{Var}[\sum_i X_i] = \sum_i \text{Var}[X_i]$.

Teorema 41. Llei (feble) dels grans nombres. Sigui X_1, \dots, X_n, \dots v.a.'s tal que totes tenen la mateixa esperança μ i variància σ^2 constants. A més, dues a dues incorrelades. Sigui $\bar{X}_n = \frac{X_1 + \dots + X_n}{n}$ la mitjana mostral, aleshores tenim, $\mathbb{E}[\bar{X}_n] = \mu$ i $\text{Var}[\bar{X}_n] = \frac{1}{n^2} \text{Var}[X_1 + \dots + X_n] = \frac{\sigma^2}{n}$. Ara, per Tchebishev tenim $P(|\bar{X}_n - \mu| \geq \epsilon) \leq \frac{\text{Var}[\bar{X}_n]}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2}$. Llavors

$$\lim_{n \rightarrow \infty} P(|\bar{X}_n - \mu| \geq \epsilon) = 0$$

3 V.a Discretes

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

Definició 42. La v.a. $X : \Omega \rightarrow \mathbb{R}$ és discreta si $\Im X$ és finita o numerable.

A partir d'ara les v.a. són totes discretes.

3.1 Funció de probabilitat, esperança i independència

Posem $\text{Im}(X) = \{a_1, a_2, a_3, \dots\}$ i $\Omega = \bigcup_{n \geq 1} \{w \in \Omega : X(w) = a_n\}$. Per tant:

$$1 = \sum_{n \geq 1} P(w \in \Omega : X(w) = a_n) = \sum_{n \geq 1} P(X = a_n)$$

Sigui $B \in \mathcal{B}$, aleshores escriurem $P(X^{-1}(B)) = P(X \in B) = \sum_{a_i \in B} P(X = a_i)$, o bé també, $P_X(B)$ que s'anomena la llei de X (és una probabilitat en $(\mathbb{R}, \mathcal{B})$), queda determinada per $\{a_1, a_2, \dots\}$ i $P(X = a_i)$, que és la funció de probabilitat de X (només per v.a. discretes).

Podem escriure la funció de distribució com:

$$F_X(x) = P(X \leq x) = \sum_{a_i \leq x} P(X = a_i)$$

I també podem reescriure l'esperança d'una variable discreta com:

$$\mathbb{E}[X] = \sum_{n \geq 1} a_n P(X = a_n) = \sum_{x \in \text{Im } X} x P(X = x)$$

Sempre que convergeixi absolutament. Si $g(X)$ és v.a., llavors

$$\mathbb{E}[g(X)] = \sum_{x \in \text{Im } X} g(x) P(X = x)$$

Independència, X, Y v.a. discretes:

$$X, Y \text{ són independents} \iff F_{X,Y}(x, y) = F_X(x)F_Y(y), \forall x, y \in \mathbb{R}$$

A més, podem caracteritzar la independència amb la funció de probabilitat:

$$X, Y \text{ són independents} \iff P(X = x, Y = y) = P(X = x)P(Y = y), \forall x, y \in \mathbb{R}$$

INDICACIÓ: Començar fixant només una variable $P(X = x, Y \leq y) = P(X = x)P(Y \leq y)$.

Proposició 43. Si X, Y són v.a. discretes, aleshores són incorrelades ($\text{Var}(X, Y) = 0$).

Demostració.

$$\begin{aligned} \mathbb{E}[XY] &= \sum_{u \in \text{Im } XY, u \neq 0} u P(XY = u) = \sum_{x \in \text{Im } X, x \neq 0} \sum_{\frac{u}{x} \in \text{Im } Y, u \neq 0} u P(X = x, Y = \frac{u}{x}) = \\ &= \sum_{x \in \text{Im } X, x \neq 0} \sum_{\frac{u}{x} \in \text{Im } Y, u \neq 0} u P(X = x) P(Y = \frac{u}{x}) = \sum_{x \in \text{Im } X, x \neq 0} x P(X = x) \sum_{\frac{u}{x} \in \text{Im } Y, u \neq 0} \frac{u}{x} P(Y = \frac{u}{x}) = \mathbb{E}[X] \mathbb{E}[Y] \end{aligned}$$

□

3.2 Funció generadora de probabilitat

Definició 44. Una v.a. és de recompte si $\text{Im } X \subseteq \{0, 1, 2, 3, \dots\}$.

4 V.a Contínues

5 Funcions característiques i famílies exponencials

6 Convergència de variables aleatòries