Teoria de la Probabilitat

Continguts

1 Espais de Probabilitat

1.1 Definició axiomàtica d'espai de probabilitat

Definició 1.1.1
Un **espai de probabilitat** és un espai de mesura $(\Omega, \mathcal{A}, p)$, tal que $p(\Omega) = 1$.

- Ω s'anomena **espai mostral**.
- $\mathcal{A}$ se l'anomena **conjunt d'esdeveniments** o **successos**.
- p se l'anomena **funció de probabilitat**.

Observació 1.1.2 $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ és una σ-àlgebra:
$\sigma 1)$ $\varnothing \in \mathcal{A}$
$\sigma 2)$ $A \in \mathcal{A} \iff \overline{A} \in \mathcal{A}$
$\sigma 3)$ Si $\{A_n\}_{n\geq 1}$ és una seqüència de successos en $\mathcal{A} \implies \bigcup\limits_{n\geq 1} A_n \in \mathcal{A}$

Observació 1.1.3 Recordem que p és una mesura i, per tant:
$p1)$ $p(\varnothing) = 0$
$p2)$ $\forall A \in \mathcal{A}$, $p(A) \geq 0$
$p3)$ Si $\{A_n\}_{n\geq 1}$ és una seqüència de successos en $\mathcal{A}$ disjunts 2 a 2 ($A_i \cap A_j = \varnothing \, si \, i \neq j$), aleshores

$$p\left(\bigcup_{n\geq 1} A_n\right) = \sum_{n\geq 1} p(A_n)$$

Vegem les primeres propietats dels espais de probabilitat:

Proposició 1.1.4
Per un espai de probabilitat $(\Omega, \mathcal{A}, p)$ es compleix que:
(i) $A \in \mathcal{A} \implies p(\overline{A}) = 1 - p(A)$
(ii) Si $A, B \in \mathcal{A}$, $A \subseteq B \implies p(A) \leq p(B)$
(iii) Si $A_1, \ldots, A_r \in \mathcal{A}$, i $A_i \cap A_j \neq \varnothing \, (i \neq j)$, aleshores $p\left(\bigcup\limits_{i=1}^{r} A_i\right) = \sum\limits_{i=1}^{r} p(A_i)$
(iv) Si $A, B \in \mathcal{A}$, $A \subseteq B \implies p(B - A) = p(B) - p(A)$
(v) Successions monòtones:

a) Si $A_1 \subseteq A_2 \subseteq A_3 \subseteq \ldots \subseteq A_i \in \mathcal{A} \implies p\left(\bigcup\limits_{n\geq 1} A_n\right) = \lim\limits_{n\to\infty} p(A_n)$

b) Si $A_1 \supseteq A_2 \supseteq A_3 \supseteq \ldots \supseteq A_i \in \mathcal{A} \implies p\left(\bigcap\limits_{n\geq 1} A_n\right) = \lim\limits_{n\to\infty} p(A_n)$

Si ara tenim un espai de probabilitat $(\Omega, \mathcal{A}, p)$, i successos $A_1, A_2, \ldots, A_i$ en general <u>no</u> disjunts, aleshores <u>no</u> és cert que $p\left(\bigcup_{i=1}^{r} A_r\right) = \sum_{i=1}^{r} p(A_i)$. En aquest cas, tenim la següent fita:

Lema 1.1.5 (Fita de la unió)
Siguin $A_1, A_2, \ldots, A_r$ successos en $(\Omega, \mathcal{A}, p)$, aleshores

$$p\left(\bigcup_{i=1}^{r} A_i\right) \leq \sum_{i=1}^{r} p(A_i)$$

Teorema (Desigualtats de Bonferroni) (1.1.6)
Siguin $A_1, \ldots, A_r$ successos en $(\Omega, \mathcal{A}, p)$. Denotem per $I \subseteq \{1, \ldots, r\} : = [r]$,

$$A_I = \bigcap_{i \in I} A_i$$

$$S_k = \sum_{\substack{I \subseteq [r] \\ |I| = k}} p(A_I)$$

Aleshores, si:
1) t és parell, $p\left(\bigcup_{i=1}^{r} A_i\right) \geq \sum_{i=1}^{t} (-1)^{i+1} \cdot S_i$
2) t és senar, $p\left(\bigcup_{i=1}^{r} A_i\right) \leq \sum_{i=1}^{t} (-1)^{i+1} \cdot S_i$

Exemple 1.1.7

1.2 Probabilitat condicionada

Definició 1.2.1
Sigui $(\Omega, \mathcal{A}, p)$ un espai de probabilitat , i $B \in \mathcal{A}$ amb $p(B) > 0$.
Per $A \in \mathcal{A}$, la **probabilitat condicionada** de A amb B és:

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

Observació 1.2.2 $p(A \mid B)$ mesura la probabilitat de que el succés A ocorri sabent que B ha succeït.

Observació 1.2.3 Si prenem

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

Aleshores P_B és una funció de probabilitat sobre $\Omega, \mathcal{A}$
De fet, si definim $\mathcal{A}_B = \{A \cap B \colon A \in \mathcal{A}\}$, aleshores $\mathcal{A}_B$ és una σ-àlgebra i P_B també defineix una probabilitat sobre $(\Omega, \mathcal{A}_B)$.

Proposició 1.2.4
Siguin $A_1, \dots, A_r \in \mathcal{A}$, tals que $p(A_i) > 0$, $A_i \cap A_j = \varnothing$ si $i \neq j$ i $\bigcup\limits_{i=1}^{r} A_i = \Omega$
($\{A_i\}_{i=1}^{r}$ és una partició de Ω)

(i) <u>Teorema de la probabilitat total</u>: $\forall A \in \mathcal{A}, p(A) = \sum\limits_{i=1}^{r} p(A \mid A_i) \cdot p(A_i)$.
(ii) <u>Fórmula de Bayes</u>: si $A \in \mathcal{A}$, $p(A) > 0$,

$$p(A_j \mid A) = \frac{p(A \mid A_j) \cdot p(A_j)}{\sum\limits_{i=1}^{r} p(A \mid A_i) \cdot p(A_i)}$$

Molts cops serà més senzill calcular probabilitats condicionades. Vegem un exemple:

Exemple 1.2.5 (La ruïna del jugador)
Partim d'un capital de k unitats ($k \geq 0$) i volem aconseguir un capital de N unitats ($N \geq k$) de la següent forma:
Llancem una moneda equilibrada, guanyant (surt cara) o perdent (surt creu) una unitat amb probabilitat $\frac{1}{2}$. El joc acaba si:

1. Ens quedem sense capital.
2. Assolim un capital igual a N

Calcularem la probabilitat de perdre.
No és bona idea intentar codificar tots els casos possibles i sumar les seves probabilitats (n^{o} tirades $\to \infty$).

Anem a resoldre el problema usant el teorema de la probabilitat total.
A_k = "el jugador, amb capital inicial igual a k, s'arruïna" (Volem calcular $p(A_k)$).
Escrivim $p(A_k) = p_k$. Aleshores $p_0 = 1$, $p_N = 0$.
Definim B = "la primera tirada de la moneda és cara". Aleshores B i $\overline{B}$ defineixen una partició de Ω.
Usem el teorema de probabilitat total:

$$\left.\begin{aligned} p(A_k) &= p(A_k \mid B)p(B) + p(A_k \mid \overline{B})p(\overline{B}) \\ &= p(A_k \mid B)\cdot\tfrac{1}{2} + p(A_k \mid \overline{B})\cdot\tfrac{1}{2} \\ &= \tfrac{1}{2}\cdot\Big(\underbrace{p(A_k \mid B)}_{p(A_{k+1})} + \underbrace{p(A_k \mid \overline{B})}_{p(A_{k-1})}\Big)\end{aligned}\right\} p_k = \frac{1}{2}(p_{k+1} + p_{k-1}); \quad p_0 = 1,\, p_N = 0$$

Resolent la recurrència tenim:

$$p_k = 1 - \frac{k}{N}$$

Si $k \in o(N)$, aleshores assimptòticament el jugador s'acabarà arruïnant.

1.3 Independència

Definició 1.3.1
Dos successos A i B (del mateix espai $(\Omega, \mathcal{A}, p)$) són **independents** si $p(A \cap B) = p(A) \cdot p(B)$.

Observació 1.3.2

- Si $p(B) > 0$, A i B independents $\iff p(A \mid B) = p(A)$
- El successos $\varnothing$ i Ω són independents amb qualsevol altre succés B
- Si A i B són independents, A i $\overline{B}$ també. (De fet tenim que A i B són independents $\iff \overline{A}$ i $\overline{B}$ són independents).

Definició 1.3.3
Donada una família de successos $\{A_i\}_{i \in I}$, es diu que és **independent** si:

$$\forall J \subseteq I \text{ amb } |J| < \infty, \text{ es té: } p\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} p(A_j)$$

Observació 1.3.4 Donats $\{A_i\}_{i=1,\ldots,r}$ successos, poden ser independents dos a dos però no com a conjunt.

1.4 Espai producte

Donats dos espais de probabilitat $(\Omega_1, \mathcal{A}_1, p_1)$ i $(\Omega_2, \mathcal{A}_2, p_2)$, volem combinar-los en un sol espai amb espai mostral $\Omega_1 \times \Omega_2$.
Això ja ho hem fet en el cas discret, però en general:

- Prenem com a espai mostral $\Omega_1 \times \Omega_2$
- Prenem com a σ-àlgebra $\sigma(\mathcal{A}_1 \times \mathcal{A}_2)$ (la més petita σ-àlgebra que conté $\mathcal{A}_1 \times \mathcal{A}_2$).
- Com definim p a $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \times \mathcal{A}_2)$?
  Voldríem que $p(A_1 \times A_2) = p(A_1) \cdot p(A_2) \quad \forall A_i \in \mathcal{A}_i$. Per definir p necessitarem el següent teorema:

Teorema (d'extensió o de Carathéodory) (1.4.1)
Sigui p_0 una funció de probabilitat en una àlgebra $\mathcal{A}_0$. Sigui $\mathcal{A} = \sigma(\mathcal{A}_0)$.
Aleshores p_0 es pot extendre a una funció de probabilitat p sobre $\mathcal{A}$ que coincideix amb p_0 en $\mathcal{A}_0$. A més, p és única.

En el nostre cas, quina àlgebra agafem?

$$\underbrace{(\mathcal{A}_1 \times \mathcal{A}_2)^*}_{\mathcal{A}_0} = \text{ totes les unions finites d'elements de } \mathcal{A}_1 \times \mathcal{A}_2.$$

Es comprova que $\sigma((\mathcal{A}_1 \times \mathcal{A}_2)^*) = \sigma(\mathcal{A}_1 \times \mathcal{A}_2)$ i apliquem el teorema.

1.5 Lema de Borel-Cantelli

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

Definició 1.5.1
Donada $\{A_n\}_{n\geq 1}$ un successió de successos,

$$\text{lim sup } A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$$

$$\text{lim inf } A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k$$

Observació 1.5.2 Tant *lim sup* A_n com *lim inf* A_n són successos.

Definició 1.5.3
Si $\{A_n\}_{n\geq 1}$ és una successió de successos amb $lim\,sup\, A_n = lim\,inf\, A_n$, aleshores $\{A_n\}_{n\geq 1}$ té **límit** i l'escriurem $limA_n = lim\,sup\, A_n$.

Propietats 1.5.4
Sigui $\{A_n\}_{n\geq 1}$ una successió de successos:

(i) $lim\,sup\,A_n = \{\omega \in \Omega\colon \omega$ pertany a un nombre infinit de $A_i's\}$

(ii) $lim\,inf\,A_n = \{\omega \in \Omega : \exists\, m = m(\omega)$ tal que $\omega \in A_r$ si $r \geq m(\omega)\}$

(iii) $lim\,inf\,A_n \subseteq lim\,sup\,A_n$

Si ara $lim\,sup\,A_n = lim\,inf\,A_n = lim A_n$, podem intercanviar lim amb p:

Proposició 1.5.5
Sigui $\{A_n\}_{n\geq 1}$ una successió de successos amb $lim A_n = A$. Aleshores, $p(A) = lim\, p(A_n)$.

El següent lema ens dóna condicions suficients molt senzilles per a calcular $p(lim\,sup\,A_n)$.

Lema 1.5.6 (Borel-Cantelli)
Sigui $(\Omega, \mathcal{A}, p)$ un espai de probabilitat i $\{A_n\}_{n\geq 1}$ una successió de successos. Aleshores:

(i) Si $\sum\limits_{n\geq 1} p(A_n) < \infty \implies p(lim\,sup\,A_n) = 0$

(ii) Si els $\{A_n\}_{n\geq 1}$ són successos independents i $\sum\limits_{n\geq 1} p(A_n) = \infty \implies p(lim\,sup\,A_n) = 1$

Observació 1.5.7 No podem treure la condició d'independència de (ii).
Si prenem $A_n = E$ succés amb $0 < p(E) < 1$, aleshores $\sum\limits_{n\geq 1} p(A_n) = \sum\limits_{n\geq 1} p(E)$, els A_n <u>no</u> són independents i $p(lim\,sup\,A_n) = p(E) < 1$.

2 Variables Aleatòries

2.1 Definició de variable aleatòria. Llei d'una v.a.

Sigui $(\Omega, \mathcal{A}, \beta)$ un espai de probabilitat. Volem estudiar funcions de Ω amb imatge en $\mathbb{R}$.

Definició 2.1.1
Una **variable aleatòria** és una funció $X : \Omega \to \mathbb{R}$ tal que per tot borelià B $\in \mathcal{B}$, $X^{-1}(B) \in \mathcal{A}$.

Per tant, una variable aleatòria és una funció mesurable entre els espais de mesura $(\Omega, \mathcal{A}, p)$ i $(\mathbb{R}, \mathcal{B}, \lambda)$.

Exemple 2.1.2
(1) Les funcions constants són variables aleatòries:

$$\begin{aligned} X : \Omega &\to \mathbb{R} \\ \omega &\mapsto c \end{aligned} \qquad \text{Si prenem } B \in \mathcal{B},\ X^{-1}(B) = \begin{cases} \varnothing & \text{si c} \notin B \\ \Omega & \text{si c} \in B \end{cases}$$

(2) **Variables aleatòries indicadores**:

$$\text{Sigui A} \in \mathcal{A}, \text{ definim } \mathbb{1}_A : \Omega \to \mathbb{R} \text{ on } \mathbb{1}_A(\omega) = \begin{cases} 0 & \text{si } \omega \notin A \\ 1 & \text{si } \omega \in A \end{cases}$$

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

(3) Si X i Y són v.a., aleshores $X + Y$, $X \cdot Y$, $|X|$, etc. són v.a.
En general, si $g : \mathbb{R}^2 \to \mathbb{R}$ és una funció mesurable, aleshores g(X,Y) és una v.a.

Estem dient que $\forall B \in \mathcal{B}$, $\{\omega \in \Omega : X(\omega) \in B\}$ és un succés i, per tant, podem calcular $P(\{\omega \in \Omega : X(\omega) \in B\}) \equiv P(X \in B)$.

Exemple 2.1.3
$P(X \leq 1) = P(\{\omega \in \Omega : X(\omega) \in (-\infty, 1)\})$

Les v.a. permeten traslladar l'estructura d'espai de probabilitat de $(\Omega, \mathcal{A}, p)$ en $(\mathbb{R}, \mathcal{B})$, donant lloc a mesures que no provenen de la mesura de Lebesgue.

Definició 2.1.4
Siguin $(\Omega, \mathcal{A}, p)$ un espai de probabilitat i X una v.a.
La **mesura de probabilitat induïda** per X és una mesura de probabilitat sobre $(\mathbb{R}, \mathcal{B})$ definida per

$$\begin{aligned} p_X \colon \mathcal{B} &\to \mathbb{R} \\ B &\mapsto p_X = P(\{\omega \in \Omega \colon X(\omega) \in B\}) \end{aligned}$$

Observació 2.1.5 $(\mathbb{R}, \mathcal{B}, p_X)$ és un espai de probabilitat.

De teoria de la mesura, és equivalent veure que $[\forall B \in \mathcal{B}, \quad X^{-1}(B)$ *és de* $\mathcal{A}]$ a veure que [*l'antiimatge de qualsevol interval* $\in \mathcal{A}$].

Per tant, per saber si una funció és una v.a. només cal veure si l'antiimatge dels intervals són de $\mathcal{A}$.

La següent definició dóna una funció en $\mathbb{R}$ que codifica molta informació de X:

Definició 2.1.6
Donada X v.a., la **funció de distribució de probabilitat** de X és:

$$\begin{aligned} F_X \colon \mathbb{R} &\to [0,1] \\ x &\mapsto P(X \leq x) \end{aligned}$$

Propietats 2.1.7
(i) Si $x_1 \leq x_2 \implies F_X(x_1) \leq F_X(x_2)$

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

(iii) $F_X(x)$ és contínua per la dreta: $\forall x, \lim_{h \to 0^+} F_X(x+h) = F_X(x)$

Observació 2.1.8

- $P(X > x) = 1 - P(X \leq x) = 1 - F_X(x)$
- $P(x_1 < X \leq x_2) = P(X \leq x_2) - P(X \leq x_1) = F_X(x_2) - F_X(x_1)$

Observació 2.1.9 Les propietats (i), (ii), (iii) de $F_X(x)$ són de fet suficients.
Si una funció $F(x)$ satisfà (i), (ii), (iii), aleshores és funció de probabilitat d'una variable aleatòria.

2.2 Moments d'una v.a. Desigualtats de Markov i Chebyshev

Siguin $(\Omega, \mathcal{A}, p)$ uns espai de probabilitat i X una v.a.

Definició 2.2.1
L'**esperança** de X és:

$$\mathbb{E}[X] = \int_\Omega X dp = \int_\mathbb{R} x\, dp_X$$

Més en general, si $f\colon \mathbb{R} \to \mathbb{R}$ és una funció mesurable,

$$\mathbb{E}[f(x)] = \int_\Omega f(x) dp = \int_\mathbb{R} f(x)\, dp_X$$

Observació 2.2.2 De teoria de la mesura, cal recordar que una funció g és integrable sii $|g|$ ho és (En general, $\mathbb{E}[f(x)]$ està definida sii $\mathbb{E}[|f(x)|] < +\infty$).

Si particularitzem f:

Definició 2.2.3
$f(x) = X^r \implies \mathbb{E}[X^r]$ és el **moment r-èssim**.

Definició 2.2.4
Si $\mathbb{E}[X] = p < +\infty$, $\mathbb{E}[(X-p)^r]$ és el **moment normalitzat r-èssim**.

En particular, si r = 2, $\mathbb{E}[(X-p)^2] = \mathbb{V}ar[X]$ és la **variància** de X.

Definició 2.2.5
Si $f(x) = x(x-1)\dots(x-r+1) \implies \mathbb{E}[f(x)] = \mathbb{E}[(X)_r]$ és el **moment factorial r-èssim**.

Proposició 2.2.6 (Propietats de l'esperança i la variància)

- Si c és la v.a. constant, $\mathbb{E}[c] = c$ i $\mathbb{V}ar[c] = 0$
- Linealitat: si $a, b \in \mathbb{R}$ i X, Y v.a., $\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$
- $A \in \mathcal{A}, X = \mathbb{1}_A$, $\mathbb{E}[\mathbb{1}_A] = P(A)$
- $|\mathbb{E}[X]| \leq \mathbb{E}[|X|]$
- $\mathbb{V}ar[c \cdot X] = c^2 \cdot \mathbb{V}ar[X]$
- $\mathbb{V}ar[c + X] = \mathbb{V}ar[X]$
- $\mathbb{V}ar[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$

Observació 2.2.7 Si $\mathbb{E}[|X|^p] < +\infty$, aleshores podem utilitzar tots els resultats de teoria dels espais L_p. Així doncs tenim les següents conseqüències:

- Hölder: $p, q > 0$, $\frac{1}{p} + \frac{1}{q} = 1$, $\mathbb{E}[|X|^p] < +\infty, \mathbb{E}[|Y|^q] < +\infty$
$\implies \mathbb{E}[|XY|] \leq \mathbb{E}[|X|^p]^{\frac{1}{p}} \cdot \mathbb{E}[|Y|^q]^{\frac{1}{q}} \quad (\mathbb{E}[|XY|]^{pq} \leq \mathbb{E}[|X|^p]^q \cdot \mathbb{E}[|Y|^q]^p)$
- Cauchy-Schwartz: si $\mathbb{E}[X^2], \mathbb{E}[Y^2] < +\infty$, aleshores $\mathbb{E}[XY]^2 \leq \mathbb{E}[X^2] \cdot \mathbb{E}[Y^2]$
- Minkowski: si $\mathbb{E}[|X|^p], \mathbb{E}[|Y|^p] < +\infty \implies \mathbb{E}[|X+Y|^p]^{\frac{1}{p}} \leq \mathbb{E}[|X|^p]^{\frac{1}{p}} + \mathbb{E}[|Y|^p]^{\frac{1}{p}}$

Teorema (Desigualtat de Markov) (2.2.8)
Sigui X un v.a. que pren valors positius i $a > 0$. Aleshores:

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

El següent resultat dóna estimacions quantitatives de quant es dispersa una v.a. en relació a la seva esperança:

Teorema (Desigualtat de Chebyshev) (2.2.9)
Sigui X una v.a. en $(\Omega, \mathcal{A}, p)$ amb $\mathbb{E}[X], \mathbb{V}ar[X] < +\infty$. Aleshores, $\forall k > 0$

$$P(|X - \mathbb{E}[X]| \geq k \cdot \mathbb{V}ar[X]^{\frac{1}{2}}) \leq \frac{1}{k^2}$$

També es pot escriure:

$$P(|X - \mathbb{E}[X]| \geq k) \leq \frac{\mathbb{V}ar[X]}{k^2}$$

2.3 Vectors de variables aleatòries. Independència de v.a.

Donat un espai de probabilitat $(\Omega, \mathcal{A}, p)$ considerem les v.a. $X_1, \ldots, X_n$. Cadascuna d'elles defineix una distribució de probabilitat sobre $\mathbb{R}$.
Aleshores podem considerar el vector $(X_1, \ldots, X_n) \colon \Omega \to \mathbb{R}^n$.

Definició 2.3.1
Un vector $(X_1, \ldots, X_n) \colon \Omega \to \mathbb{R}^n$ és un **vector de variables aleatòries** (o una v.a. multidimensional), si per tot $B \in \mathcal{B}_n$ (Borelians en $\mathbb{R}^n$), $(X_1, \ldots, X_n)^{-1}(B) \in \mathcal{A}$.

Com $\Pi_i \colon \mathbb{R}^n \to \mathbb{R}$ (projecció en la i-èssima component) és una funció mesurable, aleshores

$$\Pi_i(X_1, \ldots, X_n) \colon \Omega \xrightarrow{(X_1,\ldots,X_n)} \mathbb{R}^n \xrightarrow{\quad \Pi_i \quad} \mathbb{R}$$
$$\omega \longmapsto (X_1, \ldots, X_n)(\omega) \longmapsto X_i(\omega)$$

$\equiv X_i(\omega)$ és una v.a. (en el sentit unidimensional).

De la mateixa manera que vam fer per les v.a. unidimensionals, podem considerar les antiimatges només en intervals.

Definició 2.3.2
Donat un espai de probabilitat $(\Omega, \mathcal{A}, p)$, i un vector de v.a. $(X_1, \ldots, X_n) = \vec{X}$, aleshores la **funció de distribució de probabilitat** de $\vec{X}$ és $F_{\vec{X}}(x_1, \ldots, x_n)$ definida per:

$$\begin{aligned} F_{\vec{X}} \colon \mathbb{R}^n &\to [0,1] \subseteq \mathbb{R} \\ (x_1, \ldots, x_n) &\mapsto P\Big((X_1 \leq x_1) \cap (X_2 \leq x_2) \cap \ldots \cap (X_n \leq x_n)\Big) = P\left(\bigcap_{i=1}^{n} X_i \leq x_i\right) \end{aligned}$$

Vegem propietats de la funció de distribució pel cas n = 2 (Per $n > 2$, és idèntic):

Lema 2.3.3
(i) Si $x_1' \geq x_1$, $x_2' \geq x_2 \implies F_{\vec{X}}(x_1', x_2') \geq F_{\vec{X}}(x_1, x_2)$

(ii) $\lim\limits_{(x_1,x_2)\to(+\infty,+\infty)} F_{\vec{X}}(x_1, x_2) = 1 \qquad \lim\limits_{(x_1,x_2)\to(-\infty,-\infty)} F_{\vec{X}}(x_1, x_2) = 0$

(iii) $\lim\limits_{(h_1,h_2)\to(0^+,0^+)} F_{\vec{X}}(x_1 + h_1, x_2 + h_2) = F_{\vec{X}}(x_1, x_2)$ (contínua "per dalt")

Observació 2.3.4 Aquestes 3 condicions són necessàries i suficients per a definir una v.a. multidimensional.

Observació 2.3.5

- Si tenim $F_{\vec{X}}(x_1, x_2)$ associada a $\vec{X} = (X_1, X_2)$, aleshores

  $$\lim_{x_2\to+\infty} F_{\vec{X}}(x_1, x_2) = \lim_{x_2\to+\infty} P\Big((X_1 \leq x_1) \cap (X_2 \leq x_2)\Big) = P(X_1 \leq x_1) = F_{\vec{X}}(x_1)$$

  A aquesta funció $\Big(\lim_{x_1\to+\infty} F_{\vec{X}}(x_1, x_2)\Big)$ se l'anomena **funció de distribució marginal**.

- Prenem un rectangle en $\mathbb{R}^2$ i $\vec{X} = (X, Y)$ v.a. multidimensional:

  $$P(a < X \leq b,\, c < Y \leq d) = F_{\vec{X}}(b, d) - F_{\vec{X}}(a, d) - F_{\vec{X}}(b, c) + F_{\vec{X}}(a, c)$$

Definició 2.3.6
Sigui $(\Omega, \mathcal{A}, p)$ un espai de probabilitat i $\{x_i\}_{i\in I}$ un conjunt de v.a. Direm que $\{x_i\}_{i\in I}$ són **independents** si: $\forall k$, $\forall i_1, \ldots, i_k \subseteq I$, $\forall B_1, \ldots, B_k \in \mathcal{B}$

$$P(X_{i_1} \in B_1, X_{i_2} \in B_2, \ldots, X_{i_k} \in B_k) = \prod_{j=1}^{k} P(X_{i_j} \in B_j)$$

Si ara prenem $X_1, \ldots, X_k$ v.a., aleshores si són independents,

$$F_{X_1,\ldots,X_k}(x_1, \ldots, x_k) = P(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_k \leq x_k) = \prod_{j=1}^{k} P(X_j \leq x_j) = \prod_{j=1}^{k} F_{X_j}(x_j)$$

Observació 2.3.7 Si X i Y són v.a. independents i f, g funcions mesurables, aleshores $f(X)$ i $g(Y)$ són també independents.

Observació 2.3.8 Si $F_{X_1,\dots,X_k}(x_1,\dots,x_k) = \prod_{j=1}^{k} F_{X_j}(x_j)$, aleshores les v.a. $X_1,\dots,X_k$ són independents.

En quant al càlcul de moments, tenim el següent resultat:

Teorema (2.3.9)
Siguin $X_1,\dots,x_k$ v.a. independents. Aleshores, si $\mathbb{E}[X_i] < +\infty$, es compleix que

$$\mathbb{E}\left[\prod_{i=1}^{k} X_i\right] = \prod_{i=1}^{k} \mathbb{E}[X_i]$$

Vam veure que l'operador esperança és lineal, però això no és cert en general per la variància. De fet, en el segon cas obtenim un terme connector anomenat covariància.

Definició 2.3.10
Donades dues v.a. X, Y, la **covariància** de X i Y ($Cov(X,Y)$) és:

$$Cov(X,Y) = \mathbb{E}\Big[(X - \mathbb{E}[X]) \cdot (Y - \mathbb{E}[Y])\Big]$$

Si ara desenvolupem aquesta expressió, obtenim: $Cov(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$.

Observem que si X i Y són independents, aleshores $X - \mathbb{E}[X]$ i $Y - \mathbb{E}[Y]$ també ho són i, per tant:

$$Cov(X,Y) = \mathbb{E}\Big[(X-\mathbb{E}[X])\cdot(Y-\mathbb{E}[Y])\Big] = \mathbb{E}\Big[X-\mathbb{E}[X]\Big]\cdot\Big[Y-\mathbb{E}[Y]\Big] = (\mathbb{E}[X]-\mathbb{E}[X])\cdot(\mathbb{E}[Y]-\mathbb{E}[Y]) = 0$$

Observació 2.3.11 $Cov(X,Y) = 0 \not\Longrightarrow X$ i Y independents.

Si ara calculem la variància d'una suma, obtenim:

$$\mathbb{V}ar[X+Y] = \mathbb{V}ar[X] + \mathbb{V}ar[Y] + 2\cdot Cov(X,Y)$$

Observació 2.3.12 Si X i Y són independents, $Cov(X,Y) = 0$ i per tant, $\mathbb{V}ar[X+Y] = \mathbb{V}ar[X] + \mathbb{V}ar[Y]$.
En general, això és cert per n v.a. independents:

$$\mathbb{V}ar\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathbb{V}ar[X_i] \qquad \text{si } \{X_i\}_{i=1}^{n} \text{ són independents.}$$

Vegem ara propietats de la covariància:

Propietats 2.3.13
(i) $Cov(c, X) = 0$ (c és una constant)

(ii) Si c és una constant, $Cov(c + X, Y) = Cov(X, Y)$

(iii) $Cov(X, X) = \mathbb{V}ar[X]$

(iv) $Cov(X, Y) = Cov(Y, X)$

(v) $Cov(aX + bY, Z) = a \cdot Cov(X, Z) + b \cdot Cov(Y, Z)$

Definició 2.3.14
Anomenem **coeficient de correlació de Pearson** $\big(\rho(X, Y)\big)$ a:

$$\rho(X, Y) = \frac{Cov(X, Y)}{\mathbb{V}ar[X]^{\frac{1}{2}} \cdot \mathbb{V}ar[Y]^{\frac{1}{2}}} \in [-1, 1]$$

Observació 2.3.15
Sabem que la igualtat en *Cauchy-Schwartz* es dóna quan $\exists\, a$ tal que $a \cdot X = Y$.
A més, $Cov(aX + b, Y) = Cov(aX, Y)$. Per tant,

$$Y = aX + b \iff \rho(X, Y) \in \{\pm 1\}$$

Així doncs, com més proper sigui $\rho(X, Y)$ a ± 1, millor serà una aproximació lineal de Y usant X.