3 Variables aleatòries i funcions de distribució.

Sigui $(\Omega, \mathcal{F}, P)$ un espai de probabilitat associat a un cert experiment aleatori. S'anomena **variable aleatòria** a una funció X que assigna a cada element $\omega \in \Omega$ un nombre real $X(\omega) = x$, amb la condició que, per a cada $x \in \mathbb{R}$, l'esdeveniment $\{\omega : X(\omega) \leq x\}$ pertany a $\mathcal{F}$. És a dir, una variable aleatòria és una funció de la forma

$$\begin{array}{rcl} X : \Omega & \to & \mathbb{R} \\ \omega & \mapsto & x = X(\omega). \end{array}$$

Intuïtivament una variable aleatòria és una mesura o quantitat que varia en funció del resultat concret ω que s'observa en realitzar l'experiment aleatori.

3.1 Variables aleatòries discretes.

Si X pren valors sobre un conjunt $S \subseteq \mathbb{R}$ que és finit o infinit numerable, es diu que X és una **variable aleatòria discreta**. Per exemple, *el nombre de fills, el nombre de pacients que ingressen en un hospital, el nombre de cotxes que arriben a un peatge* són variables aleatòries discretes.

Funció de probabilitat. Si X és una variable aleatòria discreta que pren valors en $S = \{x_1, x_2, \ldots, x_n\}$, finit o infinit numerable, es defineix la **funció de probabilitat** de X com

$$f(x) = \begin{cases} P(X = x), & \text{si } x \in S, \\ 0, & \text{si } x \notin S. \end{cases}$$

Proposició 3.1 *Una funció qualsevol pot servir com una funció de probabilitat d'una variable aleatòria discreta X si i només si els seus valors $f(x)$ satisfan les condicions:*

1. $f(x) \geq 0$ *per a tot* $x \in \mathbb{R}$,
2. $\sum_x f(x) = 1$, *on la suma s'estén per a tots els valors x dins del domini de f.*

Si X és una variable aleatòria discreta que pren valors sobre un conjunt $S \subseteq \mathbb{R}$ i $A \subseteq S$, aleshores

$$P(X \in A) = \sum_{x \in A} f(x).$$

La representació gràfica més utilitzada de la funció de probabilitat d'una variable aleatòria discreta és el *diagrama de barres.*

Exemple 13 *La variable aleatòria X representa el nombre de cares menys el nombre de creus en 3 tirades d'una moneda que està trucada de manera que és dos cops més probable que surti cara que no pas creu.*

Indicarem "c" per cara i "+" per creu. L'espai mostral és $\Omega = \{\omega_1 = (c,c,c), \omega_2 = (c,c,+), \omega_3 = (c,+,c), \omega_4 = (+,c,c), \omega_5 = (c,+,+), \omega_6 = (+,c,+), \omega_7 = (+,+,c), \omega_8 = (+,+,+)\}$*. Assignem a cada punt mostral un valor de* X*:*
$X(\omega_1) = 3$,
$X(\omega_2) = X(\omega_3) = X(\omega_4) = 1$,
$X(\omega_5) = X(\omega_6) = X(\omega_7) = -1$,
$X(\omega_8) = -3$.

La funció de probabilitat és:

$$f(x) = \begin{cases} 1/27 & \text{si } x = -3, \\ 2/9 & \text{si } x = -1, \\ 4/9 & \text{si } x = 1, \\ 8/27 & \text{si } x = 3, \\ 0 & \text{altrament.} \end{cases}$$

Utilitzant la funció de probabilitat podem calcular

$$P(X < 0) = P(X = -3) + P(X = -1) = 1/27 + 2/9 = 7/27,$$

$$P(0 \leq X \leq 2) = P(X = 1) = 4/9.$$

Exercici 5 *Verifiqueu que la funció donada per*

$$f(x) = \frac{x+2}{25}, \text{ per a } x = 1, 2, 3, 4, 5,$$

pot servir com a funció de probabilitat d'una variable aleatòria discreta. (Solució: Només cal comprovar les condicions 1 i 2 que ha de complir una funció per a ser una funció de probabilitat).

La **funció de distribució** o **funció de probabilitat acumulada** d'una variable aleatòria X és una aplicació

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

En el cas de variables aleatòries discretes,

$$F(x) = P(X \leq x) = \sum_{t \leq x} f(t), \text{ per a } x \in \mathbb{R},$$

on $f(t)$ és el valor de funció de probabilitat de X en t. Si X només pren un nombre finit de valors $x_1 < x_2 < \ldots < x_n$, $(n < \infty)$, la funció de distribució de X en un punt x, $x_k < x < x_{k+1}$ és

$$F(x) = P(X \leq x) = \sum_{i=1}^{k} f(x_i).$$

Propietats de les funcions de distribució.

1. $\lim_{x\to-\infty} F(x) = 0$ i $\lim_{x\to+\infty} F(x) = 1$,
2. $F(x)$ és no decreixent: si $a < b \Rightarrow F(a) \leq F(b)$.
   En efecte, ja que si $a < b$, aleshores $(-\infty, a] \subset (-\infty, b]$ i, per tant, $P((-\infty, a]) \leq P((-\infty, b])$.

A més, la funció de distribució de les variables aleatòries discretes té les següents propietats:

1. Si X només pren un nombre finit de valors $x_1 < x_2 < \ldots < x_n$, aleshores $f(x_1) = F(x_1)$, $f(x_i) = F(x_i) - F(x_{i-1})$, per a $i = 2, 3, \ldots, n$.
2. Si X és una variable aleatòria discreta, F és una funció esglaonada que té salts només en els punts del conjunt S i l'alçada d'un salt en un punt $x_i \in S$ és igual a la probabilitat $P(X = x_i)$.

Exemple 13 (continuació) *Calculem la funció de distribució de X:*

$$F(x) = \begin{cases} 0 & si\ x < -3, \\ 1/27 & si\ -3 \leq x < -1, \\ 7/27 & si\ -1 \leq x < 1, \\ 19/27 & si\ 1 \leq x < 3, \\ 1 & si\ x \geq 3. \end{cases}$$

Exercici 6 *Si X té funció de distribució*

$$F(x) = \begin{cases} 0 & si\ x < -1, \\ 1/4 & si\ -1 \leq x < 1, \\ 1/2 & si\ 1 \leq x < 3, \\ 3/4 & si\ 3 \leq x < 5, \\ 1 & si\ x \geq 5. \end{cases}$$

trobeu:

$$\text{(a) } P(X \leq 3), \quad \text{(b) } P(X = 3), \quad \text{(c) } P(X < 3),$$
$$\text{(d) } P(X \geq 1), \quad \text{(e) } P(-0.4 < X < 4), \quad \text{(f) } P(X = 5).$$

3.2 Variables aleatòries contínues.

Ara considerem variables aleatòries que poden prendre qualsevol valor en un interval de la recta real o en una unió d'intervals. Aquest tipus de variables aleatòries s'anomenen **contínues**.

Funció de densitat de probabilitat. Una funció $f : \mathbb{R} \to [0, +\infty)$ s'anomena una **funció de densitat de probabilitat** o **funció de densitat** de X si i només si

$$P(X \in [a, b]) = P(a \leq X \leq b) = \int_a^b f(x)\, dx,$$

per a qualsevol a i b reals amb $a \leq b$.

Observació: El valor de $f(c)$, és a dir, la densitat de probabilitat de X en c, no dóna $P(X = c)$ com en el cas discret. En el cas continu, les probabilitats sempre estan associades a intervals i $P(X = c) = 0$ per a qualsevol constant real c. Una conseqüència d'aquesta propietat és que el valor d'una funció de densitat de probabilitat es pot canviar amb alguns dels valors d'una variable aleatòria sense canviar les probabilitats, i és per això que es diu que la funció $f(x)$ anterior és *una* densitat i no *la* densitat de X. També tenint en compte aquesta propietat, si X és una variable aleatòria contínua i $[a, b]$ és un interval de $\mathbb{R}$, aleshores

$$P(a \leq X \leq b) = P(a \leq X < b) = P(a < X \leq b) = P(a < X < b).$$

Proposició 3.2 *Una funció qualsevol pot servir com una funció de densitat d'una variable aleatòria contínua X si els seus valors $f(x)$ satisfan les condicions:*

1. $f(x) \geq 0$ *per a* $x \in \mathbb{R}$,
2. $\int_{-\infty}^{+\infty} f(x)\, dx = 1$.

Si X és una variable aleatòria contínua i el valor de la seva funció de densitat en t és $f(t)$, aleshores la funció donada per

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\, dt, \text{ per a } x \in \mathbb{R},$$

s'anomena **funció de distribució** o **funció de probabilitat acumulada** de X.

Les propietats de les funcions de distribució donades per al cas discret són també vàlides per al cas continu, és a dir, $\lim_{x \to -\infty} F(x) = 0$, $\lim_{x \to +\infty} F(x) = 1$ i $F(a) \leq F(b)$ quan $a < b$. A més, la funció de distribució de les variables aleatòries contínues té les següents propietats:

1. $P(a \leq X \leq b) = F(b) - F(a)$,
2. la funció de densitat pot obtenir-se derivant la funció de distribució, sempre que la derivada existeixi, és a dir,

$$f(x) = \frac{dF(x)}{dx}.$$

Exercici 7 *Trobeu la funció de densitat per a la variable aleatòria que té funció de distribució*

$$F(x) = \begin{cases} 0, & x \leq 0, \\ x, & 0 \leq x \leq 1, \\ 1, & x \geq 1 \end{cases}$$

i dibuixeu la seva gràfica.

Exercici 8 *(a) Demostreu que* $f(x) = 3x^2$*,* $x \in (0, +\infty)$*, representa una funció de densitat d'una variable aleatòria* X*.*
(b) Feu una gràfica d'aquesta funció i indiqueu l'àrea associada a la probabilitat $P(X > 1)$*.*
(c) Calculeu la probabilitat $P(X > 1)$*.*

3.3 Transformacions d'una variable aleatòria.

Moltes vegades ens interessarà obtenir la distribució d'una funció (o transformació) coneguda d'una variable aleatòria. Per exemple, si volem analitzar unes dades en logaritmes per a obtenir una distribució més simètrica o si volem canviar l'escala de mesura de la variable (metres per centímetres o dòlars per euros, per exemple). En general, donada una variable aleatòria X voldrem obtenir la distribució d'una altra variable $Y = h(X)$, on h és la funció o transformació coneguda.

Mètode basat en la funció de distribució. Anomenem $G(y)$ a la funció de distribució de la nova variable aleatòria $Y = h(X)$. Aleshores,

$$G(y) = P(Y \leq y) = P(h(X) \leq y) = P(X \in A),$$

on A representa el conjunt de valors de X en els que es verifica que $h(X) \leq y$. En el cas particular de variables contínues en que la funció h sigui contínua i monòtona creixent, la relació $h(X) \leq y$ equival a $X \leq h^{-1}(y)$ i, per tant,

$$G(y) = P(X \leq h^{-1}(y)) = F(h^{-1}(y)), \tag{3}$$

on F és la funció de distribució de la variable aleatòria X.
Si h és monòtona decreixent, la relació $h(X) \leq y$ equival a $X \geq h^{-1}(y)$ i aleshores,

$$G(y) = P(X \geq h^{-1}(y)) = 1 - P(X \leq h^{-1}(y)) = 1 - F(h^{-1}(y)). \tag{4}$$

Exemple 14 *Si la funció de densitat de probabilitat de* X *està donada per*

$$f(x) = \begin{cases} 6x(1-x), & x \in (0,1), \\ 0, & x \notin (0,1), \end{cases}$$

trobeu la funció de densitat de probabilitat de la variable aleatòria $Y = X^3$*.*
Sigui $G(y)$ *el valor de la funció de distribució de* Y *en el punt* y*,* $0 < y < 1$*, aleshores:*

$$G(y) = P(Y \leq y) = P(X^3 \leq y) = P(X \leq y^{1/3})$$
$$= \int_0^{y^{1/3}} 6x(1-x)\,dx = 3y^{2/3} - 2y.$$

Derivant $G(y)$ *obtenim la funció de densitat de* Y*:* $g(y) = 2(y^{-1/3} - 1)$*, si* $0 < y < 1$*, i* $g(y) = 0$ *en altre cas.*

Mètode basat en la funció de probabilitat. Per a variables aleatòries discretes la funció de probabilitat de $Y = h(X)$ serà

$$f(y) = P(Y = y) = \sum_{y=h(x_i)} f(x_i),$$

és a dir, per a calcular la probabilitat de y sumarem les probabilitats de tots els valors de la variable X que donen lloc al valor y.

Exemple 15 *Si X una variable aleatòria discreta amb funció de probabilitat*

x	*0*	*1*	*2*	*3*	*4*
$f(x)$	*1/16*	*4/16*	*6/16*	*4/16*	*1/16*

aleshores la funció de probabilitat de la variable aleatòria $Y = 1/(1+X)$ és

$y = 1/(1+x)$	*1*	*1/2*	*1/3*	*1/4*	*1/5*
$g(y)$	*1/16*	*4/16*	*6/16*	*4/16*	*1/16*

Observeu que en aquest exemple les probabilitats no han canviat, sinó que ara aquestes probabilitats estan relacionades amb els diversos valors que pren la variable Y enlloc de amb els corresponents valors de la variable X. Això es deu a que $g(y) = P(Y = y) = P(X = x) = f(x)$.