Fonaments de Probabilitat i Estadística. Notes de classe

Pedro Delicado

7 de març de 2018

Índex

Tema 2

Variables aleatòries

2.1 Variables aleatòries

Els resultats d'un experiment aleatori són elements de l'espai mostral Ω. Segons quin sigui l'experiment, aquests elements podrien ser poc tractables.

Exemple 2.1.1.
Considerem l'experiment de llençar un dau:

$$\Omega = \{\text{⚀}, \text{⚁}, \text{⚂}, \text{⚃}, \text{⚄}, \text{⚅}\}$$

Quant val la suma de dos resultats, ⚁ i ⚂ per exemple?
Els sabem sumar perquè associem un nombre real a cada resultat: el 2 a ⚁ i el 3 a ⚂.

□

Donar una representació numèrica als resultats d'un experiment aleatori és en general una pràctica útil. Quan fem això direm que estem definint una *variable aleatòria* associada a l'experiment.

2.1.1 Definició de variable aleatòria

Definició 2.1.1. *Sigui* $(\Omega, \mathcal{A}, \Pr)$ *un espai de probabilitat. Una* VARIABLE ALEATÒRIA X *és una funció que associa un nombre real a cada resultat de l'experiment aleatori*

$$\begin{array}{rccc} X: & \Omega & \longrightarrow & \mathbb{R} \\ & w & \longmapsto & X(w) \end{array}$$

tal que $\{w \in \Omega : X(w) \leq x\} \in \mathcal{A}$ *per a tot* $x \in \mathbb{R}$.

Dit d'una altra manera, $X : \Omega \longrightarrow \mathbb{R}$ és una VARIABLE ALEATÒRIA si per a qualsevol $x \in \mathbb{R}$ es pot calcular la probabilitat de l'esdeveniment $\{w \in \Omega : X(w) \leq x\}$.

Aquesta condició tècnica es verifica, en particular, quan l'espai mostral Ω és finit i treballem amb $\mathcal{A} = \mathcal{P}(\Omega)$.

Exemple 2.1.2.

Considerem l'experiment de llençar un dau, amb

$$\Omega = \{⚀, ⚁, ⚂, ⚃, ⚄, ⚅\}$$

Definim la variable aleatoria $X =$"Punts obtinguts":

$$\begin{array}{rccl} X: & \Omega & \longrightarrow & \mathbb{R} \\ & ⚀ & \longmapsto & X(⚀) = 1 \\ & ⚁ & \longmapsto & X(⚁) = 2 \\ & ⚂ & \longmapsto & X(⚂) = 3 \\ & ⚃ & \longmapsto & X(⚃) = 4 \\ & ⚄ & \longmapsto & X(⚄) = 5 \\ & ⚅ & \longmapsto & X(⚅) = 6 \end{array}$$

També es podrien definir altres variables aleatòries, com ara $Y =$"Indicadora 0-1 de que surt un nombre parell":

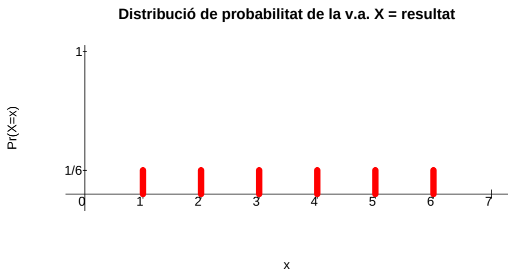
$$\begin{array}{rccl} Y: & \Omega & \longrightarrow & \mathbb{R} \\ & ⚀ & \longmapsto & 0 \\ & ⚁ & \longmapsto & 1 \\ & ⚂ & \longmapsto & 0 \\ & ⚃ & \longmapsto & 1 \\ & ⚄ & \longmapsto & 0 \\ & ⚅ & \longmapsto & 1 \end{array}$$

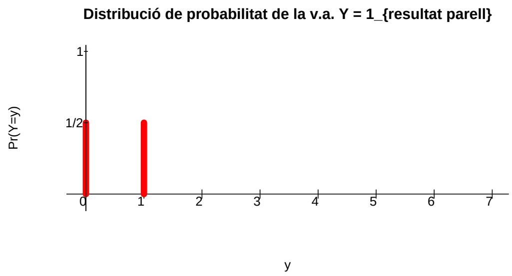
□

A l'exemple anterior les variables aleatòries X i Y han traslladat la probabilitat total de l'espai $\Omega = \{⚀, ⚁, ⚂, ⚃, ⚄, ⚅\}$ a la recta real de dues formes diferents, que podem representar en forma de taula y també gràficament:

x	1	2	3	4	5	6
$\Pr(X = x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

y	0	1
$\Pr(Y = y)$	$\frac{1}{2}$	$\frac{1}{2}$





2.1.2 Funció de distribució

Distribució d'una variable aleatòria

Una variable aleatòria $X : \Omega \longrightarrow \mathbb{R}$ determina una mesura de probabilitat P_X sobre la recta real $\mathbb{R}$ que per $A \subseteq \mathbb{R}$ es defineix com

$$P_X(A) = \Pr(X \in A) = \Pr\left(\{w \in \Omega : X(w) \in A\}\right) = \Pr(X^{-1}(A)).$$

És a dir, cada variable aleatòria X distribueix la probabilitat total d'Ω sobre $\mathbb{R}$ d'una determinada manera, que ve indicada per la mesura de probabilitat P_X que X indueix a $\mathbb{R}$.

En realitat, no podrem assignar probabilitats a qualsevol subconjunt A de $\mathbb{R}$, només als subconjunts A de la σ-ÀLGEBRA DE BOREL, $\mathcal{B}$, que és la σ-àlgebra més petita que conté els semi-intervals $(-\infty, x]$, $x \in \mathbb{R}$.

Definició 2.1.2 (Distribució de probabilitat d'una variable aleatòria)**.** *Sigui* $(\Omega, \mathcal{A}, \Pr)$ *un espai de probabilitat i* $X : \Omega \longrightarrow \mathbb{R}$ *una variable aleatòria. La mesura de probabilitat* P_X *sobre* $(\mathbb{R}, \mathcal{B})$ *determinada per la relació*

$$P_X\{(-\infty, x]\} = \Pr\left(\{w \in \Omega : X(w) \le x\}\right) \ \textit{per a tot} \ \ x \in \mathbb{R},$$

s'anomena DISTRIBUCIÓ DE PROBABILITAT *de la variable aleatòria* X*, o* LLEI *de* X*.*

Per determinar la distribució de probabilitat P_X de X n'hi ha prou amb conèixer la probabilitat dels semi-intervals $(-\infty, x]$,

$$P_X\{(-\infty, x]\} = \Pr(w \in \Omega : X(w) \le x) = \Pr(X \le x), \ x \in \mathbb{R}.$$

Definició 2.1.3 (Funció de distribució d'una variable aleatòria)**.** *Sigui* X *una variable aleatòria. Les probabilitats dels semi-intervals* $(-\infty, x]$*,* $x \in \mathbb{R}$*, defineixen la funció*

$$\begin{array}{rccl} F_X : & \mathbb{R} & \longrightarrow & [0,1] \\ & x & \longmapsto & \Pr(X \le x) \end{array}$$

que s'anomena FUNCIÓ DE DISTRIBUCIÓ *de la variable aleatòria* X*.*

Teorema 2.1.4. *La distribució de probabilitat P_X d'una variable aleatòria X queda caracteritzada completament per la seva funció de distribució F_X.*

Teorema 2.1.5.

1. *La funció de distribució és no decreixent, és a dir: si $x_1 < x_2$ llavors $F(x_1) \leq F(x_2)$.*
2. $\lim_{x \longrightarrow -\infty} F(x) = 0$ *i* $\lim_{x \longrightarrow \infty} F(x) = 1$.
3. *La funció de distribució és contínua per la dreta, és a dir:*
$$F(x) = F(x^+) = \lim_{y \longrightarrow x, y > x} F(y).$$
4. *Determinació de probabilitats a partir de la funció de distribució:*
   - $\Pr(X > x) = 1 - \Pr(X \leq x) = 1 - F(x)$.
   - $\Pr(x_1 < X \leq x_2) = F(x_2) - F(x_1)$.
   - $\Pr(X < x) = F(x^-)$.
   - $\Pr(X = x) = F(x) - F(x^-)$.

Demostració:

1. Si $x_1 < x_2$ llavors $[X \leq x_1] \subseteq [X \leq x_2]$. Per tant $\Pr(X \leq x_1) \leq \Pr(X \leq x_2)$.
2. La proposició és intuïtiva, però la seva demostració no és immediata. Es pot provar usant el *Teorema de la probabilitat de la unió infinita de conjunts encaixats*, que diu que
$$\text{Prob}\{\bigcup_{i=1}^{\infty} A_i\} = \lim_{n \longrightarrow \infty} \Pr\{A_n\}, \text{si} A_1 \subset A_2 \subset \dots, \text{i}$$
$$\text{Prob}\{\bigcap_{i=1}^{\infty} A_i\} = \lim_{n \longrightarrow \infty} \Pr\{A_n\}, \text{si} A_1 \supset A_2 \supset \dots$$
Anem a la demostració. Volem veure que $\lim_{x \longrightarrow \infty} \Pr(X \leq x) = 1$. Considerem $A_n = \{X \leq n\}$. És clar que $A_1 \subset A_2 \subset \dots$ Tenim que $\{\bigcup_{n=1}^{\infty} A_n\} = A_\infty = \Omega$ i per tant $\Pr\{\bigcup_{n=1}^{\infty} A_n\} = 1$. D'altra banda
$$1 = \Pr\{\bigcup_{n=1}^{\infty} A_n\} = \lim_{n \longrightarrow \infty} \Pr\{A_n\} = \lim_{n \longrightarrow \infty} \Pr\{X \leq n\} = \lim_{n \longrightarrow \infty} F(n).$$
L'altra igualtat es veu anàlogament.

3. La demostració no és immediata. Sigui $y_1 > y_2 > \ldots$ una successió decreixent amb $\lim_n y_n = x$. Aleshores $A_n = \{X \leq y_n\}$ és una successió de conjunts decreixent ($A_n \supset A_{n+1}$ per tot n). A més a més, $\{X \leq x\} = \cap_{n=1}^{\infty}\{X \leq y_n\}$. Per tant

$$F(x) = \Pr(X \leq x) = \Pr(\cap_{n=1}^{\infty}\{X \leq y_n\}) =$$

$$\lim_n \Pr(X \leq y_n) = \lim_n F(y_n) = F(x^+).$$

De nou hem fet servir el *Teorema de la probabilitat de la unió infinita de conjunts encaixats.*

4. La prova es immediata.

$\square$

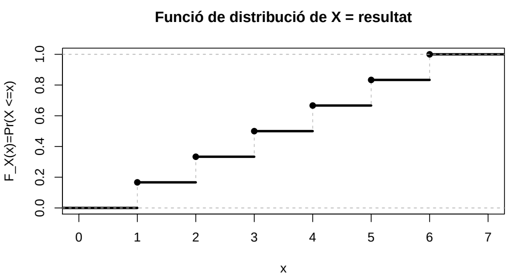
Observem que una funció de distribució no té perquè ser contínua en general, el que ens diu el tercer apartat del Teorema és que ho és per la dreta. Pel que fa al quart apartat, recordem que $F(x^-) = \lim_{y \longrightarrow x, y<x} F(y)$.

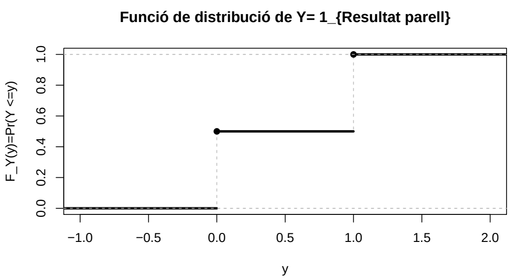
Exemple 2.1.3.

En el llançament d'un dau, hem definit les variables aleatòries
$X =$"Punts obtinguts",
$Y =$"Indicadora 0-1 de que surt un nombre parell".

Les seves funcions de distribució són aquestes:

$$F_X(x) = \begin{cases} 0 & \text{si} & & x & < 1 \\ 1/6 & \text{si} & 1 \leq & x & < 2 \\ 2/6 & \text{si} & 2 \leq & x & < 3 \\ 3/6 & \text{si} & 3 \leq & x & < 4 \\ 4/6 & \text{si} & 4 \leq & x & < 5 \\ 5/6 & \text{si} & 5 \leq & x & < 6 \\ 1 & \text{si} & 6 \leq & x & \end{cases} \qquad F_Y(y) = \begin{cases} 0 & \text{si} & & y & < 0 \\ 1/2 & \text{si} & 0 \leq & y & < 1 \\ 1 & \text{si} & 1 \leq & y & \end{cases}$$





$\square$

Tots els exemples de variables aleatòries que hem vist fins ara tenen una característica en comú: aquestes variables aleatòries només poden prendre un nombre finit de valors. Però no totes les variables aleatòries compleixen això. Vegem dos exemples:

1. Llencem una moneda a l'aire fins que surti la primera cara. Definim la v.a. X com *el nombre de llançaments que hem hagut de fer*. En aquest cas els valors que pot prendre X son tots els nombres naturals. Per tant la v.a. X pot prendre una quantitat numerable de valors. Tot i que ara la quantitat de valors possibles per a X ja no és finita, aquesta v.a. comparteix un tret fonamental amb les v.a. anteriors: els possibles valors de X són punts aïllats de la recta real, i d'aquests pot haver-ne una quantitat finita o, com a màxim, una quantitat infinita numerable. A aquest tipus de variables aleatòries les anomenarem **variables aleatòries discretes**.

2. Hem escrit un programa que ordena n nombres reals que llegeix d'un fitxer ASCII. El temps d'execució depèn de n i també de l'ordenació inicial dels nombres al fitxer. Hem comprovat que quan n és més gran de 10000, el temps d'execució és superior a 3 segons. Definim la v.a. Y com la part decimal del temps d'execució del programa (expressat en segons) quan el fitxer conté n nombres aleatoris, $n \geq 10000$. En aquest cas Y pot prendre qualsevol valor a l'interval $[0, 1)$, és a dir, hi ha una quantitat no numerable de possibles valors. Sembla raonable suposar que la probabilitat es reparteix uniformement a tot l'interval, és a dir, que no hi ha zones privilegiades de l'interval $[0, 1)$ que tendeixin a contenir els valors de Y més vegades que altres zones. Per exemple,

$$\Pr(Y \in [a - \epsilon, a + \epsilon]) = \Pr(Y \in [b - \epsilon, b + \epsilon])$$

per a tot $\epsilon \in (0, 1/2)$, i tots a i b en $(\epsilon, 1 - \epsilon)$. Aleshores la funció de distribució de Y hauria de ser

$$F_Y(y) = \Pr(Y \leq y) = \begin{cases} 0 & \text{si} \quad y < 0 \\ y & \text{si} \quad 0 \leq y < 1 \\ 1 & \text{si} \quad 1 \leq y \end{cases}$$

A les variables aleatòries del tipus de Y les anomenarem **variables aleatòries contínues**.

2.2 Variable aleatòria discreta. Funció de probabilitat

Definició 2.2.1. *Una variable aleatòria X és* VARIABLE ALEATÒRIA DISCRETA *si només pot prendre un nombre finit o numerable de valors. És a dir, el conjunt imatge de l'aplicació $X : \Omega \longrightarrow \mathbb{R}$ és un conjunt finit o numerable de $\mathbb{R}$: $\{x_1, \ldots, x_k, \ldots\}$.*

El SUPORT d'una variable aleatòria discreta X és

$$\mathcal{X} = \{x \in \mathbb{R} : \Pr(X = x) > 0\}.$$

Definició 2.2.2. *Donada una variable aleatòria discreta X es defineix la* FUNCIÓ DE PROBABILITAT DE X *com*

$$p(x_j) = \Pr\{X = x_j\} = \Pr\{w \in \Omega : X(w) = x_j\}, \text{ per a } x_j \in \mathcal{X}.$$

Observem que per a tot $A \subseteq \mathbb{R}$

$$\Pr(X \in A) = \Pr(X \in \mathcal{X} \cap A) = \Pr(X \in \cup_{x_j \in \mathcal{X} \cap A} \{x_j\}) = \sum_{x_j \in \mathcal{X} \cap A} p(x_j).$$

En particular,

$$\Pr\{X \in \mathbb{R}\} = \Pr(\Omega) = 1 = \sum_{x_j \in \mathcal{X}} p(x_j).$$

Això prova la següent proposició.

Proposició 2.2.3. *La distribució de probabilitat d'una v.a. discreta queda caracteritzada completament per la seva funció de probabilitat.*

Definició 2.2.4. *Una funció $p : \mathbb{R} \longrightarrow \mathbb{R}$ és una* FUNCIÓ DE PROBABILITAT *si verifica*

1. *$p(x) \geq 0$ per tot $x \in \mathbb{R}$.*

2. *Existeix un conjunt $\mathcal{X} \subseteq \mathbb{R}$ finit o numerable tal que $\sum_{x_j \in \mathcal{X}} p(x_j) = 1$.*

Proposició 2.2.5 (Funció de distribució d'una variable aleatòria discreta)**.** *Sigui X una variable aleatòria que pren valors $x_1 < x_2 < \ldots$ amb probabilitats $p_1, p_2, \ldots$ La* FUNCIÓ DE DISTRIBUCIÓ F_X *és una funció esglaonada que salta* *en cada x_i, amb un salt d'alçada p_i, i que és constant entre dos valors de x_i consecutius. Podem escriure*

$$F_X(x) = \begin{cases} 0 & si\ x < x_1, \\ p_1 & si\ x_1 \leq x < x_2, \\ p_1 + p_2 & si\ x_2 \leq x < x_3, \\ \vdots & \vdots \\ p_1 + \ldots + p_j & si\ x_j \leq x < x_{j+1}, j > 1, \\ \vdots & \vdots \end{cases}$$

Exemple 2.2.1.

Llencem una moneda fins que obtenim cara. Denotem per X el nombre de llançaments necessaris. Llavors $\Pr(X = x) = (1-p)^{x-1}p$ per a $x = 1, 2, \ldots$. (Més endavant direm que X té distribució geomètrica i ho escriurem així: $X \sim \mathrm{Geom}(p)$).

Calculem la funció de distribució de X:

$$F(x) = \Pr(X \leq x) = \sum_{i=1}^{x} \Pr(X = i) = \sum_{i=1}^{x} (1-p)^{i-1} p = \frac{1-(1-p)^x}{1-(1-p)} p = 1-(1-p)^x,$$

on hem fet servir que la suma dels n primers termes d'una progressió geomètrica de raó $r \neq 1$ i primer terme a_1 és

$$S_n = \frac{a_1(1-r^n)}{1-r}.$$

Fixem-nos que $F(x)$ és plana entre dos enters consecutius.

Nota: Una altra manera d'arribar a l'expressió de $F(x)$ és aquesta:

$$F(x) = \Pr(X \leq x) = 1 - \Pr(X > x) =$$

$$1 - \Pr(\text{“Surt creu les primeres } x \text{ vegades”}) = 1 - (1-p)^x.$$

$\square$

2.3 Variable aleatòria contínua. Funció de densitat de probabilitat

Definició 2.3.1. *Una variable aleatòria $X : \Omega \longrightarrow \mathbb{R}$ per a la qual existeix una funció no negativa f tal que:*

$$\Pr(a \leq X \leq b) = \Pr(\{w : a \leq X(w) \leq b\}) = \int_a^b f(u)\, du,$$

per a tots a i b reals amb $a \leq b$, és una VARIABLE ALEATÒRIA CONTÍNUA *amb* FUNCIÓ DE DENSITAT *f. En aquest cas*

$$\int_{-\infty}^{\infty} f(u)\, du = 1.$$

Definició 2.3.2. *Direm que una funció $f : \mathbb{R} \longrightarrow \mathbb{R}$ és una* FUNCIÓ DE DENSITAT *si*

1. $f(x) \geq 0$ *per tot* $x \in \mathbb{R}$.
2. $\int_{\mathbb{R}} f(x)dx = 1$.

Propietats:

1. Si X és v.a. contínua amb funció de densitat f llavors

   $$\Pr(X = x) = \int_x^x f(u)du = 0$$

   per tot x.

2. Per tot A subconjunt (mesurable) de $\mathbb{R}$ tenim que

   $$\Pr(X \in A) = \int_A f(u)\, du.$$

3. Una interpretació de la funció de densitat. Per a f contínua en x i δ petit

   $$\Pr(x - \frac{\delta}{2} \leq X \leq x + \frac{\delta}{2}) = \int_{x-\frac{\delta}{2}}^{x+\frac{\delta}{2}} f(u)du \approx \delta f(x),$$

   llavors

   $$f(x) \approx \frac{\Pr(x - \frac{\delta}{2} \leq X \leq x + \frac{\delta}{2})}{\delta} = \frac{\text{probabilitat}}{\text{longitud}},$$

   que són justament les unitats de mesura d'una *densitat de probabilitat per unitat de longitud.*

   Una altra forma de escriure aquesta propietat és aquesta:

   $$f(x)dx \approx \Pr(X \in [x, x + dx]),$$

   on dx és una longitud infinitesimal.

4. Per tot a, b tenim

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

Això és conseqüència del fet que $\Pr(X = x) = 0$ per a tot $x \in \mathbb{R}$

Proposició 2.3.3 (Funció de distribució d'una variable aleatòria contínua)**.** *Una variable aleatòria X contínua en l'interval $[a, b]$ (amb possibilitat de tenir $-\infty$ o ∞ en els extrems) té* FUNCIÓ DE DISTRIBUCIÓ

$$F_X(x) = \Pr(X \leq x) = \int_{-\infty}^{x} f_X(u)du$$

que és una funció contínua i diferenciable, amb derivada

$$\frac{d}{dx} F_X(x) = f_X(x).$$

Observem que les darreres afirmacions són conseqüència del TEOREMA FONAMENTAL DEL CÀLCUL.

Exemple 2.3.1.

Direm que la variable aleatoria X té distribució uniforme a l'interval $[a, b] \subset \mathbb{R}$ (amb longitud $b - a < \infty$) si és contínua i la seva funció de densitat és constant a l'interval $[a, b]$ i val 0 fora d'aquest interval:

$$f(x) = \begin{cases} k & \text{si } x \in [a, b] \\ 0 & \text{altrament} \end{cases}$$

Per a calcular el valor de k, fem servir que la integral de f sobre tot $\mathbb{R}$ ha de ser igual a 1:

$$1 = \int_{-\infty}^{\infty} f(x)dx = \int_{a}^{b} kdx = k\, x\big|_a^b = k(b - a) \Longrightarrow k = \frac{1}{b - a}.$$

La densitat $f(x)$ també s'escriu

$$f(x) = \frac{1}{b - a} I_{[a,b]}(x),$$

on la funció $I_A(x)$ és la *funció indicadora* del conjunt A, que val 1 si $x \in A$ i val 0 en cas contrari.

□

No totes les variables aleatorias són contínues o discretes:
No totes les variables aleatorias són o bé contínues o bé discretes. N'hi ha que són variables aleatorias SINGULARS: la seva funció de distribució és contínua en tot $\mathbb{R}$ però no és derivable en cap punt. De fet hi ha un resultat que diu que qualsevol funció de distribució F és la MIXTURA (o BARREJA) de tres funcions de distribució: una discreta F_d, una (absolutament) contínua F_c i l'altra singular F_s:

$$F(x) = \alpha_d F_d(x) + \alpha_c F_c(x) + \alpha_s F_s(x),$$

on $\alpha_d + \alpha_c + \alpha_s = 1$.

2.4 Característiques d'una variable aleatòria: esperança i moments

La distribució de probabilitat d'una variable aleatòria X (és a dir, la forma en que X reparteix la probabilitat sobre la recta real) es pot descriure (o resumir) per un petit nombre de característiques numèriques, també anomenades PARÀMETRES o INDICADORS de la v.a., que tenen una interpretació pràctica. Per exemple, aquests paràmetres poden indicar al voltant de quin punt ha distribuït X la probabilitat (paràmetres de localització), o si la v.a. X ha concentrat molt o poc la probabilitat al voltant d'aquest punt (paràmetres de dispersió).

2.4.1 El valor esperat d'una variable aleatòria

Cas discret

Sigui una variable aleatòria discreta que pren valors al conjunt $\mathcal{X} = \{x_1, x_2, \ldots\}$ i sigui $p(x_i)$ la seva funció de probabilitat.

Definició 2.4.1. *Definim l'*ESPERANÇA MATEMÀTICA *o* MITJANA *de X com el següent valor μ:*

$$E(X) = \mu = \sum_{x_i \in \mathcal{X}} x_i p(x_i) = \sum_{x_i \in \mathcal{X}} x_i \ \Pr(X = x_i),$$

sempre que $\sum_{x_i \in \mathcal{X}} |x_i| p(x_i) < \infty$.

El valor $E(X)$ així definit és el centre de gravetat del gràfic de la funció de probabilitat de X (veieu la Figura 2.1). És per això que $E(X)$, l'esperança matemàtica de X, és un bon paràmetre de localització de la distribució de

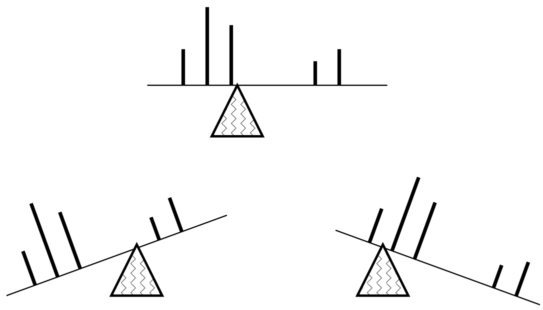


Figura 2.1: L'esperança és el punt que aguanta en equilibri el diagrama de barres de la funció de probabilitat; és el seu centre de gravetat.

probabilitat de X, és a dir, és un bon candidat com a **centre de la distribució** de X.

Per altra banda, hi ha un resultat teòric (la *Llei Forta dels Grans Nombres*, que estudiarem més endavant) que diu que si s'observen n realitzacions independents d'una variable aleatòria X, les mitjanes mostrals $\bar{X}_n = (1/n)\sum_{i=1}^{n} X_i$ tendeixen cap a $E(X)$ quan n tendeix cap a infinit (amb probabilitat 1).

Proposició 2.4.2. *Sigui X una variable aleatoria discreta que pren valors al conjunt $\mathcal{X}$. Sigui $g(x)$ una funció real definida sobre $\mathcal{X}$. Definim la variable aleatoria $Y = g(X)$. Llavors Y és una variable aleatoria discreta i podem calcular la seva esperança com*

$$E\left(g(X)\right) \sum_{x_i \in \mathcal{X}} g(x_i)p(x_i),$$

sempre que $\sum_{x_i \in \mathcal{X}} |g(x_i)|p(x_i) < \infty$.

Exemple 2.4.1.

Participem en un joc que consisteix en llençar dos daus equilibrats i guanyar un nombre de monedes igual a la suma de puntuacions obtingudes. Ens preguntem quin és el guany esperat en aquest joc.

Considerarem la variable aleatòria $X =$"Guany en una jugada". Construïm una taula amb els valors i les respectives probabilitats de X:

resultat	probabilitat
$2, 12$	$\frac{1}{36}$
$3, 11$	$\frac{2}{36}$
$4, 10$	$\frac{3}{36}$
$5, 9$	$\frac{4}{36}$
$6, 8$	$\frac{5}{36}$
7	$\frac{6}{36}$

Per tant $E(X) = \sum_{x=2}^{12} x\ p(x) = \frac{1}{36}(2 \cdot 1 + 3 \cdot 2 + 4 \cdot 3 + 5 \cdot 4 + 6 \cdot 5 + 7 \cdot 6 + 8 \cdot 5 + 9 \cdot 4 + 10 \cdot 3 + 11 \cdot 2 + 12 \cdot 1) = 7$. □

Exemple 2.4.2.
En el joc de la ruleta americana apostem 10 euros als senars. Ens preguntem quin és el guany net esperat en cada aposta.

Tenim que $R = \{1, 2, \ldots, 36, 0, 00\}$ i

$$X = \begin{cases} +10 & \text{si } r \text{ senar} \\ -10 & \text{si } r \text{ parell, 0 o 00} \end{cases}$$

Per tant $E(X) = 10 \cdot \frac{18}{38} - 10 \cdot \frac{20}{38} = \frac{-10}{19} = -0.5263$.

Hem obtingut doncs que per cada 10 euros apostats en perdrem, en terme mig, 0.5263. □

Cas continu

Sigui X una variable aleatòria contínua amb funció de densitat $f_X(x)$.

Definició 2.4.3. *Si $\int_{-\infty}^{\infty} |x| f_X(x)\, dx < \infty$ es defineix l'*ESPERANÇA *de X com*

$$E(X) = \int_{-\infty}^{\infty} x\ f_X(x)\, dx.$$

Si la variable aleatòria X està fitada, és a dir $\Pr(a \le X \le b) = 1$ per alguns a i b reals, aleshores $E(X)$ existeix sempre.

Exemple 2.4.3.
Sigui X variable aleatòria uniforme a l'interval $[0, 1]$, amb funció de densitat, per tant,

$$f(x) = \begin{cases} 1 & \text{si } 0 < x < 1 \\ 0 & \text{altrament} \end{cases}$$

Calcularem l'esperança de X:

$$E(X) = \int_0^1 x \cdot 1\, dx = \left.\frac{x^2}{2}\right|_0^1 = \frac{1}{2}.$$

$\square$

Exemple 2.4.4.
Sigui X variable aleatòria amb funció de densitat

$$f(x) = \begin{cases} 2x & \text{si } 0 < x < 1 \\ 0 & \text{altrament} \end{cases}$$

Calcularem l'esperança de X:

$$E(X) = \int_0^1 x\, 2x\, dx = \int_0^1 2\, x^2\, dx = 2\, \left.\frac{x^3}{3}\right|_0^1 = \frac{2}{3}.$$

$\square$

Sigui X una variable aleatòria contínua amb funció de densitat $f_X(x)$. Sigui g una funció de $\mathbb{R}$ en $\mathbb{R}$ *prou regular* (tècnicament es demana que g sigui MESURABLE: les funcions que fem servir habitualment ho són). Aleshores $Y = g(X)$ també és una variable aleatoria i té sentit preguntar-se pel valor esperat de Y.

Proposició 2.4.4. *Sigui $Y = g(X)$. Suposem que $\int_{-\infty}^{\infty} |g(x)| f_X(x)\, dx < \infty$, llavors l'esperança de $Y = g(X)$ es pot calcular com*

$$E(g(X)) = \int_{-\infty}^{\infty} g(x) f_X(x)\, dx.$$

Nota: Observeu que la variable aleatoria $Y = g(X)$ podria ser tant contínua com discreta.

Exemple 2.4.5.
Sigui X variable aleatòria amb funció de densitat

$$f(x) = \begin{cases} 2x & \text{si } 0 < x < 1 \\ 0 & \text{altrament} \end{cases}$$

Calcularem l'esperança de $\sqrt{X}$:

$$E(\sqrt{X}) = \int_0^1 \sqrt{x}\, 2x\, dx = \int_0^1 2\, x^{\frac{3}{2}}\, dx = 2\left.\frac{x^{\frac{3}{2}+1}}{\frac{3}{2}+1}\right|_0^1 = \frac{4}{5}.$$

□

Propietats de l'esperança

Teorema 2.4.5. *Sigui X variable aleatòria amb $E(X)$ ben definida.*

1. *Si $Y = aX + b$ aleshores $E(Y)$ existeix i $E(Y) = aE(X) + b$.*
2. *Si existeixen $E(g(X))$ i $E(h(X))$ aleshores $E(g(X)+h(X)) = E(g(X))+E(h(X))$.*
3. *Si existeix a tal que $\Pr(X \geq a) = 1$ aleshores $E(X) \geq a$.*
4. *Si existeix b tal que $\Pr(X \leq b) = 1$ aleshores $E(X) \leq b$.*

Demostració: Totes les parts del teorema es demostren usant propietats elementals de la integral o del sumatori. Aquí demostrarem la tercera afirmació per variables aleatòries contínues. Sabem que existeix a tal que $\Pr(X \geq a) = 1$. Llavors $f(x) = 0$ per $x < a$ i per tant

$$\int_{-\infty}^{\infty} x\, f(x)\, dx = \int_a^{\infty} x\, f(x)\, dx \geq \int_a^{\infty} a\, f(x)\, dx = a \cdot \Pr(X \geq a) = a$$

□

El següent resultat afirma que si volem resumir tota una variable aleatòria per només un nombre real, l'esperança és la millor tria (en el sentit que tindrem l'error quadràtic esperat més petit possible). Això reforça la interpretació de l'esperança com a mesura de localització de la distribució de probabilitat d'una variable aleatòria.

Teorema 2.4.6. *Sigui X variable aleatòria amb $E(X^2)$ ben definida. Aleshores $\mu = E(X)$ verifica que*

$$E((X-\mu)^2) \leq E((X-a)^2), \text{ per tot } a \in \mathbb{R}.$$

Demostració:

$$E((X-a)^2) = E(((X-\mu)+(\mu-a))^2) = E((X-\mu)^2+(\mu-a)^2+2(\mu-a)(X-\mu))$$
$$= E((X-\mu)^2)+(\mu-a)^2+2(\mu-a)E(X-\mu) = E((X-\mu)^2)+(\mu-a)^2 \geq E((X-\mu)^2).$$

□

2.4.2 Variància i desviació estàndard

Ara definirem mesures de dispersió (al voltant del valor esperat) de la distribució de probabilitat d'una variable aleatòria.

Definició 2.4.7. *Sigui X variable aleatòria amb $\mu = E(X) < \infty$. Definim la* VARIÀNCIA *de X com*

$$Var(X) = E[(X - \mu)^2].$$

Aquesta quantitat és finita si i només si $E(X^2) < \infty$. Si $E[(X - \mu)^2]$ és infinita, direm que $Var(X)$ no existeix.

Si existeix $Var(X)$, es defineix la DESVIACIÓ TIPUS *(o estàndard, o típica) de X com*

$$DT(X) = \sqrt{Var(X)}.$$

Observem que com que $(X - \mu)^2 \geq 0$ llavors $Var(X) \geq 0$.

Si X és variable aleatòria discreta, $V(X) = \sum_i \Pr(X = x_i)(x_i - \mu)^2$.

Si X és variable aleatòria contínua amb funció de densitat $f_X(x)$ llavors

$$Var(X) = E[(X - \mu)^2] = \int_{-\infty}^{\infty} (x - \mu)^2 f(x)\, dx.$$

Observeu que $Var(X)$ és el valor esperat de les desviacions (al quadrat) dels valors que pren X respecte de la seva esperança. Per tant $Var(X)$ és una mesura de dispersió de la distribució de X.

Com que les unitats en les quals està expressada $Var(X)$ són les unitats de X al quadrat, val la pena prendre l'arrel quadrada de $Var(X)$, definint així DT(X), per tenir una mesura de dispersió expressada en las mateixes unitats que la variable X.

Exemple 2.4.6.

Sigui X variable aleatòria uniforme a l'interval $[0, 1]$, amb densitat

$$f(x) = \begin{cases} 1 & \text{si } 0 < x < 1 \\ 0 & \text{altrament} \end{cases}$$

Recordem que $E(X) = 1/2$. Calcularem la variància X:

$$Var(X) = \int_0^1 \left(x - \frac{1}{2}\right)^2 1\, dx = \int_0^1 \left(x^2 - x + -\frac{1}{4}\right) dx$$

$$= \left(\frac{x^3}{3} - \frac{x^2}{2} + \frac{x}{4}\right)\Bigg|_0^1 = \frac{1}{3} - \frac{1}{2} + \frac{1}{4} = \frac{1}{12}.$$

La desviació tipus de X serà doncs

$$DT(X) = \sqrt{V(X)} = \sqrt{\frac{1}{12}} = 0.2887.$$

□

Teorema 2.4.8.

1. *$Var(X) = 0$ si i només si existeix c tal que $\Pr(X = c) = 1$.*
2. *Per qualsevol a i b tenim: $Var(aX + b) = a^2 Var(X)$.*
3. *$Var(X) = E(X^2) - (E(X))^2$, o també $E(X^2) = Var(X) + (E(X))^2$.*

Demostració:

1. Si existeix c tal que $\Pr(X = c) = 1$ llavors $E(X) = c$ i $\Pr((X - c)^2 = 0) = 1$. Per tant $Var(X) = E((X - c)^2) = 0$.

   Recíprocament, si $Var(X) = 0$ tenim que: $E((X - \mu)^2) = 0$. D'altra banda $\Pr((X - \mu)^2 \geq 0) = 1$, per tant

   $$\left.\begin{array}{l} E((X-\mu)^2) = Var(X) = 0 \\ \text{Però } \Pr((X-\mu)^2 \geq 0) = 1 \end{array}\right\} \Rightarrow \Pr((X-\mu)^2 = 0) = 1 \Rightarrow \Pr(X = \mu) = 1.$$

2. $Var(aX + b) = E((aX + b - a\mu - b)^2) = E(a^2(X - \mu)^2) = a^2 Var(X)$.
3. $Var(X) = E((X - \mu)^2) = E(X^2 + \mu^2 - 2X\mu) = E(X^2) + \mu^2 - 2\mu^2 = E(X^2) - (E(X))^2$.

□

Exemple 2.4.7.

Sigui X variable aleatòria amb funció de densitat

$$f(x) = \begin{cases} 2x & \text{si } 0 < x < 1 \\ 0 & \text{altrament} \end{cases}$$

Recordem que $E(X) = 2/3$. Calcularem ara la variància de X. Farem servir la fórmula

$$Var(X) = E(X^2) - (E(X))^2.$$

Necessitem calcular

$$E(X^2) = \int_0^1 x^2\, 2x\, dx = \int_0^1 2\, x^3\, dx = 2\, \left. \frac{x^4}{4} \right|_0^1 = \frac{1}{2}.$$

Per tant

$$Var(X) = E(X^2) - (E(X))^2 = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{2} - \frac{4}{9} = \frac{1}{18}$$

La desviació tipus de X serà doncs

$$DT(X) = \sqrt{V(X)} = \sqrt{\frac{1}{18}} = 0.2357.$$

$\square$

2.4.3 Quantils, mediana i quartils

Hi ha d'altres característiques que descriuen la distribució de probabilitat d'una v.a. X que es defineixen a partir de la seva funció de distribució i no involucren càlculs d'esperances.

Definició 2.4.9 (Quantils d'una distribució). *Sigui X v.a. amb funció de distribució F i sigui $p \in (0,1)$. Es defineix el* quantil d'ordre p de X *com*

$$x_p = \inf\{x \in \mathbb{R} : F(x) \geq p\}.$$

Si X és absolutament contínua i F és invertible, aleshores

$$x_p = F^{-1}(p).$$

En general es fa servir la notació

$$x_p = F^{-1}(p)$$

tot i que la funció F no sigui invertible.

A vegades també està definit x_p per $p = 0$ o $p = 1$.

Definició 2.4.10. Mediana: *$Med(X) = x_p$ per $p = 1/2$.*

Quartils: *$Q_1 = x_p$ per $p = 1/4$, $Q_2 = x_p$ per $p = 2/4 = 1/2$ (aleshores $Q_2 = Med(X)$), $Q_3 = x_p$ per $p = 3/4$.*

Deciles: x_p *per* $p = k/10$, $k = 1, \ldots, 9$.

Percentils: x_p *per* $p = k/100$, $k = 1, \ldots, 99$.

La mediana divideix la recta real en dues parts connexes que tenen el 50% de la probabilitat cadascuna. La mediana (que sempre existeix) és un paràmetre de localització alternatiu al valor esperat (que per algunes variables aleatorias no està ben definit).

Els quartils divideixen la recta real en quatre parts connexes que tenen el 25% de la probabilitat cadascuna. El RANG INTERQUARTÍLIC es defineix com la distància entre el primer i el tercer quartil:

$$\mathrm{RIQ} = Q_3 - Q1.$$

És una mesura de dispersió de la distribució de probabilitat.

2.4.4 Moments i moments centrats d'una variable aleatoria

Definició 2.4.11. *Sigui X una variable aleatòria i sigui k un enter positiu qualsevol, llavors definim el* MOMENT D'ORDRE *k de X com $E(X^k)$.*

Observem que el moment d'ordre k d'una variable aleatòria X existeix si, i només si, $E(|X|^k) < \infty$. Es pot demostrar que si $E(|X|^k) < \infty$ per un enter k, aleshores $E(|X|^j) < \infty$ per tot $j < k$.

Definició 2.4.12. *Donada X variable aleatòria i k enter positiu, definim el* MOMENT CENTRAT D'ORDRE *k com $E((X - \mu)^k)$ on $\mu = E(X)$.*

Observacions

1. L'esperança d'una variable aleatòria és el seu moment no centrat de primer ordre.

2. La variància d'una variable aleatòria és el seu moment centrat de segon ordre.

3. El primer moment centrat és 0: $E(X - \mu) = 0$.

4. Si X és simètrica al voltant de $\mu = E(X)$ (és a dir, $f(\mu - x) = f(\mu + x)$ per tot $x \in \mathbb{R}$) i $E((X - \mu)^k)$ existeix per un k senar, aleshores $E((X - \mu)^k) = 0$. Això és degut a que els termes positius i els negatius es cancel·len entre ells.

Definició 2.4.13 (Coeficients d'asimetria i d'apuntament)**.** *Sigui X una v.a. amb $E(X) = \mu$ i $V(X) = \sigma^2$ El* COEFICIENT D'ASIMETRIA *de X és el moment d'ordre 3 de la variable estandarditzada:*

$$CAs(X) = E\left(\left(\frac{X-\mu}{\sigma}\right)^3\right) = \frac{E((X-\mu)^3)}{\sigma^3}.$$

El COEFICIENT D'APUNTAMENT*, o de* CURTOSI*, de X és el moment d'ordre 4 de la variable estandarditzada:*

$$CAp(X) = E\left(\left(\frac{X-\mu}{\sigma}\right)^4\right) = \frac{E((X-\mu)^4)}{\sigma^4}.$$

El coeficient d'asimetria mesura l'asimetria de la distribució al voltant del valor esperat. Serà zero quan la variable sigui simètrica al voltant del valor esperat, positiva si hi ha asimetria a la dreta (la cua de la dreta és més pesada que la de l'esquerra) i negativa si hi ha asimetria a l'esquerra. El coeficient d'apuntament mesura com de pesades són les cues de la distribució.

2.5 Models de distribucions usuals

Hi ha variables aleatorias que, tot i haver estat definides en espais mostrals diferents, responen a un mateix model matemàtic. Revisarem alguns d'aquests models de variables aleatorias. També veurem que, sota certes condicions, les distribucions de probabilitat de certs models de variables aleatorias tendeixen a semblar-se a altres distribucions de probabilitat també conegudes.

2.5.1 Models de variables aleatòries discretes

Distribució uniforme discreta

Definició 2.5.1. *Direm que la variable aleatoria discreta X és* UNIFORME DISCRETA *entre* 1 *i n si*

$$\Pr(X = i) = \frac{1}{n}, \quad \textit{per a tot } i \in \{1, \dots, n\}.$$

Aquesta distribució correspon a l'esquema d'escollir un element a l'atzar entre n elements, quan aquests són igualment probables.

Proposició 2.5.2. *Si X és variable aleatoria uniforme discreta entre* 1 *i n aleshores*

$$E(X) = \frac{n+1}{2}, \quad Var(X) = \frac{n^2-1}{12}.$$

Demostració:

$$E(X) = \sum_{i=1}^{n} i\frac{1}{n} = \frac{1}{n}\frac{n(n+1)}{2} = \frac{n+1}{2}.$$

$$E(X^2) = \sum_{i=1}^{n} i^2\frac{1}{n} = \frac{1}{n}\frac{n(n+1)(2n+1)}{6} = \frac{(n+1)(2n+1)}{6}.$$

$$\begin{aligned}\mathrm{Var}(X) = E(X^2) - (E(X))^2 &= \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4} \\ = \frac{2(n+1)(2n+1) - 3(n+1)^2}{12} &= \frac{(n+1)(4n+2-3n-3)}{12} \\ = \frac{(n+1)(n-1)}{12} &= \frac{n^2-1}{12}.\end{aligned}$$

□

Distribució de Bernoulli

Definició 2.5.3. *Un experiment aleatori que té només dos possibles resultats, als quals anomenarem* èxit *(que apareix amb probabilitat p) i* fracàs *(amb probabilitat $q = (1-p)$), es diu un* EXPERIMENT DE BERNOULLI *amb paràmetre p. Escriurem $X \sim Bern(p)$.*

Definició 2.5.4. *Es diu que la variable aleatòria*

$$X = \begin{cases} 1 & si\ \grave{e}xit \\ 0 & si\ frac\grave{a}s \end{cases}$$

té distribució de Bernoulli amb paràmetre p: $X \sim B(p)$.

Proposició 2.5.5. *Si X és variable aleatoria de Bernoulli amb paràmetre p aleshores*

$$E(X) = p, \ \ Var(X) = p(1-p).$$

Demostració:

$$E(X) = 0 \cdot (1-p) + 1 \cdot p = p.$$

Observeu que sempre es compleix que $X^2 = X$, tant si X val 0 com si val 1. Aleshores $E(X^2) = E(X) = p$. Per tant

$$\mathrm{Var}(X) = E(X^2) - (E(X))^2 = p - p^2 = p(1-p).$$

□

Distribució Binomial

Definició 2.5.6. *La variable aleatòria X té distribució* BINOMIAL *amb paràmetres n i p si X és el nombre d'èxits en n assajos independents d'un experiment de Bernoulli amb paràmetre p. Escriurem $X \sim B(n,p)$.*

Proposició 2.5.7. *Si $X \sim B(n,p)$, aleshores pren valors $0,1,\dots,n$ i la seva funció de probabilitat és*

$$p(k) = \Pr(X=k) = \binom{n}{k} p^k (1-p)^{n-k}, \ \textit{per a tot } k \in \{0,\dots,n\}.$$

A més a més,

$$E(X) = np, \ \ Var(X) = np(1-p).$$

Demostració: Per tal que $X = k$, a la seqüència d'n experiments de Bernoulli s'han d'haver produït k èxits i $(n-k)$ fracassos, en qualsevol ordre. Hi ha un total de $\binom{n}{k}$ possibles ordenacions dels k èxits i els $(n-k)$ fracassos (tantes com a formes tenim de triar, entre els n experiments, el subconjunt de mida k on apareixeran els èxits). Per la independència dels n experiments de Bernoulli, cada seqüència ordenada d'experiments amb k èxits i $(n-k)$ fracassos té probabilitat $p^k(1-p)^k$. Aleshores

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

Observeu que la fórmula del binomi de Newton garanteix que $\sum_{k=0}^n \Pr(X = k) = 1$:

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

Sigui X_k, $k = 0,\dots,n$, la variable aleatoria de Bernoulli associada al k-èssim experiment. Aleshores

$$X = \sum_{i=1}^{n} X_i,$$

on $X_1,\dots,X_n$ són n variables aleatòries independents, totes $\mathrm{Bern}(p)$. Aplicant propietats de l'esperança i la variància de sumes de variables aleatorias tenim que

$$E(X) = E(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} E(X_i) = np,$$

i, tenint en compte la independència (en el Tema 3 veurem que si X i Y són variables aleatorias independents, aleshores $\text{Var}(X+Y) = \text{Var}(X)+\text{Var}(Y)$),

$$\text{Var}(X) = \text{Var}(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} \text{Var}(X_i) = np(1-p).$$

$\square$

Distribucions geomètrica i binomial negativa

Definició 2.5.8. *Es repeteix successivament de forma independent un experiment de Bernoulli. Sigui X el nombre d'assaigs fins a obtenir el primer èxit. Direm que la variable aleatòria X segueix una distribució* GEOMÈTRICA *amb paràmetre p. Escriurem $X \sim Geom(p)$.*

Proposició 2.5.9. *Si $X \sim Geom(n,p)$, aleshores pren valors $k \in \{1,\dots,n,\dots\}$ i la seva funció de probabilitat és*

$$p(k) = \Pr(X = k) = (1-p)^{k-1}p, \ \textit{per a tot } k \geq 1.$$

A més a més,

$$E(X) = \frac{1}{p}, \ Var(X) = \frac{1-p}{p^2}.$$

Demostració: La variable aleatoria X val k si i només si es produeixen primer $(k-1)$ fracassos i després un èxit. Això té probabilitat $(1-p)^(k-1)p$.

Fent servir que la suma dels infinits termes d'una progressió geomètrica de raó $r \in (-1,1)$ i primer terme a_1 és

$$S = \frac{a_1}{1-r},$$

es prova que $p(k)$ és una funció de probabilitat:

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

Calculem ara $E(X) = \sum_{k=1}^{\infty} k(1-p)^{k-1}p$. Observeu primer que

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

Aleshores podem calcular $g'(p)$ de dues maneres: derivant la sèrie terme a terme,

$$g'(p) = -\sum_{k=1}^{\infty} k(1-p)^{k-1},$$

o bé derivant la seva suma:

$$g'(p) = \frac{-p-(1-p)}{p^2} = -\frac{1}{p^2}.$$

Però

$$E(X) = \sum_{k=1}^{\infty} k \Pr(X=k) = \sum_{k=1}^{\infty} k(1-p)^{k-1}p = -pg'(p) = \frac{p}{p^2} = \frac{1}{p}.$$

Per a calcular $\mathrm{Var}(X)$ farem la segona derivada de la funció g:

$$g''(p) = \sum_{k=1}^{\infty} k(k-1)(1-p)^{k-2} = \frac{1}{p(1-p)} \sum_{k=1}^{\infty} k(k-1)(1-p)^{k-1}p$$

$$= \frac{1}{p(1-p)} E(X(X-1)) = \frac{1}{p(1-p)} \left(E(X^2) - E(X)\right).$$

Observeu però que

$$g''(p) = \frac{2}{p^3}.$$

Per tant,

$$E(X^2) - E(X) = \frac{2p(1-p)}{p^3} = \frac{2(1-p)}{p^2}.$$

Aleshores

$$\mathrm{Var}(X) = E(X^2) - (E(X))^2 = E(X^2) - E(X) + E(X) - (E(X))^2$$

$$= \frac{2(1-p)}{p^2} + \frac{1}{p} - \frac{1}{p^2} = \frac{2-2p+p-1}{p^2} = \frac{1-p}{p^2}.$$

□

Definició 2.5.10. *Es repeteix successivament de forma independent un experiment de Bernoulli. Sigui X el nombre d'assaigs fins a obtenir el r-èssim èxit. Direm que la variable aleatòria X segueix una distribució* BINOMIAL NEGATIVA *amb paràmetres r i p. Escriurem $X \sim BinNeg(r,p)$.*

Proposició 2.5.11. *Si $X \sim BinNeg(r,p)$, aleshores pren valors k naturals amb $k \geq r$, i la seva funció de probabilitat és*

$$p(k) = \Pr(X = k) = \binom{k-1}{r-1}(1-p)^{k-r}p^r, \text{ per a tot } k \geq r.$$

A més a més,

$$E(X) = \frac{r}{p}, \ Var(X) = \frac{r(1-p)}{p^2}.$$

Demostració: Donat que han d'aparèixer r èxits com a mínim haurem de fer r experiments de Bernoulli. Per tant els valors $k \in \mathbb{N}$ que pren X són $k \geq r$. Ara, cada seqüència de k experiments on apareixen r èxits i $(k-r)$ fracassos té probabilitat $(1-p)^{k-r}pr$. D'aquestes només ens interessen aquelles que tenen èxit a l'últim experiment (el k-èssim): n'hi ha $\binom{k-1}{r-1}$. Per tant

$$\Pr(X = k) = \binom{k-1}{r-1}(1-p)^{k-r}p^r.$$

Per a calcular l'esperança i la variància de X, escriurem aquesta variable aleatoria com a suma de r variables aleatorias geomètriques independents. Efectivament, definim X_1 com el nombre d'experiment fins a obtenir el primer èxit. Observeu que $X_1 \sim \text{Geom}(p)$.

Després d'haver tingut el primer èxit continuem repetint l'experiment de Bernoulli fins que apareix el 2n èxit. Definim X_2 com el nombre d'experiments que cal fer fins a obtenir aquest segon èxit, sense comptar els primers X_1 experiments. Aleshores $X_2 \sim \text{Geom}(p)$ i X_2 és independent de X_1 perquè els experiments que fem després d'haver obtingut el primer èxit són independents dels que havíem realitzat abans.

De la mateixa manera es defineixen $X_3, \ldots, X_r$: per a $i = 2, \ldots, r$, X_i és el nombre d'experiments que, després haver obtingut l'èxit $(i-1)$-èssim, necessitem fins a que apareix el següent èxit. Aleshores $X_1, \ldots, X_r$ són variables aleatorias independents amb distribució Geom(p). A més a més

$$X = \sum_{i=1}^{r} X_i.$$

Per tant

$$E(X) = E\left(\sum_{i=1}^{r} X_i\right) = \sum_{i=1}^{r} E(X_i) = \frac{r}{p},$$

$$\text{Var}(X) = \text{Var}\left(\sum_{i=1}^{r} X_i\right) = \sum_{i=1}^{r} \text{Var}(X_i) = \frac{r(1-p)}{p^2}.$$

$\square$

Distribució de Poisson

Definició 2.5.12. *La variable aleatòria discreta X té distribució de* POISSON *amb paràmetre $\lambda > 0$ si pren valors $0, 1, \ldots$ i la seva funció de probabilitat és*

$$p(k) = \Pr(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}, k = 0, 1, 2, \ldots.$$

Escriurem $X \sim Poisson(\lambda)$.

La expansió en sèrie de Taylor de la funció e^{λ} al voltant del 0 garanteix que $\sum_{k=1}^{\infty} \Pr(X = k) = 1$:

$$g(\lambda) = e^{\lambda} = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}.$$

La distribució de probabilitat de Poisson serveix com a model per a variables aleatòries que compten el nombre d'esdeveniments *rars* (poc freqüents i que apareixen de forma imprevisible) apareguts durant un interval de temps fixe:

- nombre de cotxes que arriben a un peatge durant un minut,
- nombre de clients que arriben a una botiga durant una hora,
- nombre d'estels fugaços observats durant una hora.

En altres contextos *l'interval de temps fixe* durant el qual s'observa el fenomen aleatori és substituït per altres tipus de *finestres d'observació*, com per exemple:

- nombre de defectes en una fulla de vidre d'un metre quadrat d'àrea,
- nombre d'estels fugaços observats pel requadre d'una finestra (durant una nit).
- nombre d'arbres d'una determinada espècie que trobem a una parcel·la de bosc de 50×50 m^2, definida abans d'anar a fer el treball de camp de comptar els arbres.

Proposició 2.5.13. *Si $X \sim Poisson(\lambda)$, aleshores*

$$E(X) = \lambda, \ \ Var(X) = \lambda.$$

Demostració:

$$E(X) = \sum_{k=0}^{\infty} k \Pr(X = k) = \sum_{k=1}^{\infty} k e^{-\lambda} \frac{\lambda^k}{k!}$$

$$= \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda} \sum_{h=0}^{\infty} \frac{\lambda^h}{h!} = \lambda e^{-\lambda} e^{\lambda} = \lambda.$$

Hem anomenat h a $(k-1)$. De forma similar

$$E(X^2) - E(X) = E(X(X-1)) = \sum_{k=0}^{\infty} k(k-1) \Pr(X = k)$$

$$= \sum_{k=2}^{\infty} k(k-1) e^{-\lambda} \frac{\lambda^k}{k!} = \lambda^2 e^{-\lambda} \sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!}$$

$$= \lambda^2 e^{-\lambda} \sum_{h=0}^{\infty} \frac{\lambda^h}{h!} = \lambda^2 e^{-\lambda} e^{\lambda} = \lambda^2.$$

Ara hem anomenat h a $(k-2)$. Tenim doncs que $E(X^2) = \lambda^2 + E(X) = \lambda^2 + \lambda$. Aleshores,

$$\mathrm{Var}(X) = E(X^2) - (E(X))^2 = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

□

Proposició 2.5.14 (Aproximació de la Binomial a la Poisson)**.** *Per a $n \geq 1$, sigui $Y_n \sim B(n, p_n)$ amb $np_n = \lambda$. Sigui $Y \sim Poisson(\lambda)$. Aleshores, per a tot $k \geq 0$,*

$$\lim_{n\to\infty} \Pr(Y_n = k) = \Pr(Y = k).$$

Demostració:

$$\lim_{n\longrightarrow\infty} P(Y_n = k) = \lim_{n\longrightarrow\infty} \binom{n}{k} p^k (1-p)^{n-k} =$$

$$\lim_{n\longrightarrow\infty} \frac{n!}{k!(n-k)!} \left(\frac{\lambda}{n}\right)^k \left(\frac{n-\lambda}{n}\right)^{n-k} = \frac{\lambda^k}{k!} \lim_{n\longrightarrow\infty} \frac{n!}{(n-k)!} \frac{1}{n^k} \frac{(n-\lambda)^{n-k}}{n^{n-k}}$$

$$= \frac{\lambda^k}{k!} \lim_{n\longrightarrow\infty} \frac{n!}{(n-k)!} \frac{1}{n^n} (n-\lambda)^{n-k} = \frac{\lambda^k}{k!} \lim_{n\longrightarrow\infty} \frac{n!}{(n-k)!} \frac{(n-\lambda)^n}{n^n} \frac{1}{(n-\lambda)^k}$$

$$= \frac{\lambda^k}{k!} \lim_{n\longrightarrow\infty} \left(1 - \frac{\lambda}{n}\right)^n \lim_{n\longrightarrow\infty} \frac{n(n-1)(n-2)\ldots 1}{(n-k)(n-k-1)\ldots 1} \frac{1}{(n-\lambda)^k}$$

(el primer límit és $e^{-\lambda}$)

$$= \frac{\lambda^k}{k!} e^{-\lambda} \lim_{n \longrightarrow \infty} n(n-1)(n-2)\dots(n-k+1) \frac{1}{(n-\lambda)^k}$$

(observem que els primers factors dins del límit es comporten com n^k)

$$= \frac{\lambda^k}{k!} e^{-\lambda} \lim_{n \longrightarrow \infty} n^k \frac{1}{(n-\lambda)^k} = \frac{\lambda^k}{k!} e^{-\lambda} \left(\lim_{n \longrightarrow \infty} \frac{n}{n-\lambda} \right)^k = \frac{\lambda^k}{k!} e^{-\lambda} = \Pr(Y = k).$$

$\square$

L'aplicació pràctica d'aquest resultat és la següent. Sigui $X \sim B(n,p)$, $\lambda = np$ i $Y \sim \text{Poisson}(\lambda)$. Si np és gran (diguem-ne $np > 1$) i p és petit (diguem-ne $p < 0.1$), aleshores

$$\Pr(X = k) \approx \Pr(Y = k), \text{ per tot } k \in \mathbb{N} \cup \{0\}.$$

La llei hipergeomètrica

Una caixa conté A boles vermelles i B blaves. Es seleccionen sense reemplaçament n boles de la caixa i es denota per X el nombre de boles vermelles obtingudes.

Observem que d'una banda $X \leq \min\{n, A\}$ i de l'altra $n - X \leq \min\{n, B\}$. Llavors:

$$X \geq n - \min\{n, B\} = \max\{0, n-B\} = \begin{cases} n - n = 0 & \text{si } n \leq B (\text{és a dir: } n - B \leq 0) \\ n - B & \text{si } n \geq B (\text{és a dir: } n - B \geq 0) \end{cases}$$

Per tant X és una variable aleatòria que pren valors entre $\max\{0, n-B\}$ i $\min\{n, A\}$.

Denotem per $p(x|A,B,n)$ la probabilitat d'obtenir exactament x boles vermelles, i direm que és la funció de probabilitat d'una variable aleatòria que segueix la Llei Hipergeomètrica. Llavors:

$$P(X = x|A,B,n) = \begin{cases} \frac{\binom{A}{x}\binom{B}{n-x}}{\binom{A+B}{n}} & \text{si } \max\{0, n-B\} \leq x \leq \min\{n, A\} \\ 0 & \text{altrament} \end{cases}$$

Plantegem ara un exemple per mostrar la relació entre la distribució hipergeomètrica i la binomial.

Exemple 2.5.1.

Suposem que tenim una urna amb a boles blanques i b negres. Les boles de l'urna s'escullen a l'atzar de manera que cadascuna d'elles té la mateixa probabilitat de ser escollida. Siguin

X_n="Nombre de boles blanques en les primeres n eleccions",
N_x="Nombre de boles que s'han d'escollir per obtenir-ne x de blanques".

Si hi ha reemplaçament cada nova elecció és un assaig Bernoulli amb $p = \frac{a}{a+b}$, i ja hem acabat:

$$P(X_n = x) = \binom{n}{x} \frac{a^x b^{n-x}}{(a+b)^n}, x = 0, 1, 2, \ldots, n.$$

Si no hi ha reemplaçament els assaigs no són independents. En aquest cas X_n té una distribució hipergeomètrica amb:

$$P(X_n = x) = \frac{\binom{a}{x}\binom{b}{n-x}}{\binom{a+b}{n}}.$$

En ambdos casos es verifica que

$$P(N_x = n) = \frac{x}{n} \cdot P(X_n = x)$$

perquè l'esdeveniment $\{N_x = n\} = \{X_n = x\} \cap A$, on A és l'esdeveniment *"l'ultima bola escollida és blanca"*, i $P(A|X_n = x) = x/n$.

$\square$

En la pràctica la distribució hipergeomètrica es pot aproximar per la binomial. Sigui $X \sim HG(A, B, n)$, $p = A/(A+B)$ i $Y \sim B(n, p)$. Si $n/(A+B)$ és petit (diguem-ne $n/(A+B) < 0.1$), aleshores

$$P(X = k) \approx P(Y = k), \text{ per tot } k \in \mathbb{N} \cup \{0\}.$$

2.5.2 Models de variables aleatòries contínues

Distribució uniforme (contínua)

Definició 2.5.15. *Siguin $a, b \in \mathbb{R}$ amb $a < b$. Llavors la variable aleatòria contínua X segueix una* DISTRIBUCIÓ UNIFORME *en $[a, b]$ si la probabilitat que $X \in S \subset [a, b]$ és proporcional a l'amplada de S. Ho denotarem per $X \sim U([a, b])$. En aquest cas:*

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{si } a \leq x \leq b \\ 0 & \text{altrament} \end{cases}$$

Proposició 2.5.16. *Si $X \sim U([a,b])$ aleshores la seva funció de distribució és*

$$F(x) = \begin{cases} 0 & si\ x < a, \\ \frac{x-a}{b-a} & si\ a \leq x < b, \\ 1 & si\ b \leq x. \end{cases}$$

A més a més,

$$E(X) = \frac{a+b}{2},\ Var(X) = \frac{(b-a)^2}{12}.$$

Per demostrar la proposició només cal escriure les definicions i fer les integrals involucrades.

Distribució exponencial. Relació amb la Poisson

Definició 2.5.17. *La variable aleatòria X té distribució* EXPONENCIAL *de paràmetre $\lambda > 0$, i escriurem $X \sim Exp(\lambda)$, si la seva funció de densitat és*

$$f(x) = \lambda e^{-\lambda x}, x > 0.$$

Proposició 2.5.18. *Si $X \sim Exp(\lambda)$ aleshores la seva funció de distribució és*

$$F(x) = \begin{cases} 0 & si\ x < 0, \\ 1 - e^{\lambda x} & si\ 0 \leq x. \end{cases}$$

A més a més,

$$E(X) = \frac{1}{\lambda},\ Var(X) = \frac{1}{\lambda^2}.$$

L'àmbit d'aplicació d'aquesta distribució és, entre d'altres, el mesurament del temps fins que succeeix un esdeveniment rar: l'espera d'un client, el temps fins que una màquina s'espatlla, el temps de vida d'un ésser viu, etc.

Una propietat important de la distribució exponencial és que és l'única variable aleatoria contínua que no té memòria, és a dir:

$$\Pr(X \geq t+h | X \geq t) = \frac{e^{-\lambda(t+h)}}{e^{-\lambda t}} = e^{-\lambda h} = \Pr(X \geq h).$$

Per tant, si modelitzem el temps de vida d'una màquina (o d'una peça) amb una distribució exponencial, estem suposant implícitament que aquesta màquina no envelleix: tant probable és que es trenqui en els 5 primers dies d'ús, com que es trenqui en els propers 5 dies, si ja porta funcionant 10 anys.

Relació entre la distribució exponencial i la distribució de Poisson

Considerem un fenomen que produeix esdeveniments al llarg del temps de forma aleatòria. Per exemple, podem pensar en l'arribada de clients a una botiga al llarg del dia o en l'emissió de partícules per una font radioactiva al llarg del seu període de desintegració.

Suposem que disposem d'un comptador $N(t)$ que per a cada instant $t \geq 0$ ens diu quants d'aquests esdeveniments impredictibles han ocorregut fins a l'instant t inclòs (els que han aparegut a instants de $[0, t]$).

Són també d'interès les variables $T_1, T_2, \ldots$ que indiquen els temps en que es produeixen esdeveniments i els temps esdeveniments consecutius $X_1, X_2, \ldots$, amb

$$X_n = T_n - T_{n-1},\ n \geq 1,$$

(hem definit $T_0 = 0$).

A l'exemple de les arribades de clients a la botiga, T_n és l'instant d'arribada del client n-èssim. X_n és el temps entre les arribades dels clients $(n-1)$ i n.

Hi ha situacions com les descrites que comparteixen 2 propietats importants:

1. Si $a < b \leq c < d$, el nombre d'esdeveniments apareguts a $(a, b]$ és independent del nombre d'esdeveniments apareguts a $(c, d]$.

2. Per a intervals *curts* $(a, a + \epsilon]$, la probabilitat que aparegui un esdeveniment entre a i $a + \epsilon$ és proporcional a ϵ.

En aquests casos direm que estem davant d'un Procés de Poisson de paràmetre λ. Es pot provar que aleshores:

- els temps entre arribades consecutives tenen distribucions exponencials de paràmetre λ i són variables aleatorias independents.

- el nombre d'esdeveniments que es produeixen en un interval de longitud t és una Poisson de paràmetre λt.

Les distribucions de Poisson (discreta) i exponencial (contínua) estan intimant lligades al procés de Poisson.

Distribució normal

Definició 2.5.19. *Siguin μ i σ nombres reals amb $\sigma > 0$. Llavors una variable aleatòria contínua X amb funció de densitat*

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2} \left(\frac{x - \mu}{\sigma} \right)^2 \right),\ -\infty < x < \infty,$$

es diu que té una distribució NORMAL *amb paràmetres* μ *i* σ^2. *Escriurem* $X \sim N(\mu, \sigma^2)$.

(**Nota:** La demostració de que aquesta funció $f(x)$ realment és una funció de densitat no és trivial i queda fora dels objectius del curs.)

(**Nota:** Més endavant veurem que si $X \sim N(\mu, \sigma^2)$ aleshores $E(X) = \mu$ i $\text{Var}(X) = \sigma^2$.)

Empíricament es demostra que molts mesuraments presenten distribucions contínues aproximadament normals. El TEOREMA DEL LÍMIT CENTRAL proporciona una explicació parcial d'aquest fenomen (ho veurem més endavant).

La funció de densitat té forma de campana i és simètrica respecte de μ: $f(x - \mu) = f(x + \mu)$. Com a conseqüència, $(X - \mu) \sim -(X - \mu)$.

La dispersió de la distribució depèn de σ, com es dedueix de la següent taula:

Interval	Probabilitat
$[\mu - 1\sigma, \mu + 1\sigma]$	0.68
$[\mu - 1.96\sigma, \mu + 1.96\sigma]$	0.95
$[\mu - 2\sigma, \mu + 2\sigma]$	0.955
$[\mu - 3\sigma, \mu + 3\sigma]$	0.997

Definició 2.5.20. *Una variable aleatòria normal amb paràmetres* $\mu = 0$ *i* $\sigma^2 = 1$ *es diu que té una distribució* NORMAL ESTÀNDARD.

La funció de densitat d'una variable aleatòria amb distribució normal estàndard és:

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-x^2}{2}\right), -\infty < x < \infty,$$

i la seva funció de distribució acumulada és:

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-u^2}{2}\right) du.$$

Com que aquesta integral no es pot avaluar algebraicament, és a dir, no té solució explícita, per tal de calcular els valors de la funció de distribució acumulada han d'utilitzar-se mètodes d'integració numèrica i/o taules (incorporades a tots els paquets estadístics).

Proposició 2.5.21. $\Phi(-x) = 1 - \Phi(x)$, $P(-a < X < a) = 2 \cdot \Phi(a) - 1$.

Remarquem que la primera propietat ens diu que per als càlculs només necessitarem taules per $x \geq 0$.

Proposició 2.5.22 (Estandardització)**.** *Sigui* $\mu \in \mathbb{R}$ *i* $\sigma > 0$.

1. *Si* $Z \sim N(0,1)$ *aleshores* $X = \mu + \sigma Z \sim N(\mu, \sigma^2)$.
2. *Si* $X \sim N(\mu, \sigma^2)$ *aleshores* $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$.

Demostració:

1. Sigui $Z \sim N(0,1)$ i sigui $X = \mu + \sigma Z$. Aleshores la funció de distribució de X és

$$F_X(x) = \Pr(X \leq x) = \Pr\left(\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right) = \Pr\left(Z \leq \frac{x-\mu}{\sigma}\right) = \Phi\left(\frac{x-\mu}{\sigma}\right).$$

Derivem ara $F_X(x)$ per a obtenir la funció de densitat de X:

$$f_X(x) = F'(x) = \phi\left(\frac{x-\mu}{\sigma}\right)\frac{1}{\sigma} = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2},$$

i per tant $X \sim N(\mu, \sigma^2)$.

2. Sigui $X \sim N(\mu, \sigma^2)$ i sigui $Z = \frac{X-\mu}{\sigma}$. Com que aquesta transformació és monòtona, tenim que

$$F_Z(z) = \Pr(Z \leq z) = \Pr\left(\frac{X-\mu}{\sigma} \leq z\right) = \Pr\left(X \leq \mu + \sigma z\right) = F_X(\mu+\sigma z).$$

Derivem ara $F_Z(z)$ per a obtenir la funció de densitat de Z:

$$f_Z(z) = F_Z'(z) = f_X(\mu + \sigma z)\sigma = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}(\frac{z\sigma+\mu-\mu}{\sigma})^2}\sigma$$

$$= \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2} = \phi(z),$$

i per tant $Z = \frac{X-\mu}{\sigma} \sim \mathrm{N}(0,1)$.

$\square$

Proposició 2.5.23. *Sigui* $X \sim N(\mu, \sigma^2)$. *Aleshores* $E(X) = \mu$ *i* $Var(X) = \sigma^2$.

Demostració: Sigui $Z \sim N(0,1)$. Com que $\mu + \sigma Z \sim N(0,1)$, aleshores

$$E(X) = \mu + \sigma E(Z),\ \mathrm{Var}(X) = \sigma^2\mathrm{Var}(Z).$$

Aleshores només cal provar que $E(Z) = 0$ i que $\mathrm{Var}(Z) = 1$.

$$E(Z) = \int_{-\infty}^{\infty} z \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz = \int_{-\infty}^{0} z \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz + \int_{0}^{\infty} z \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz$$

$$= -\int_{0}^{\infty} u \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du + \int_{0}^{\infty} z \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz = 0,$$

on hem fet el canvi de variable $u = -z$. Per simetria de $z^2\psi(z)$ al voltant del 0,

$$E(Z^2) = \int_{-\infty}^{\infty} z^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz = 2\int_{0}^{\infty} z^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz.$$

Fem integració per parts,

$$u = z,\ du = dz;\ dv = ze^{-\frac{1}{2}z^2} dz,\ v = -e^{-\frac{1}{2}z^2},$$

i temim que

$$\mathrm{Var}(Z) = E(Z^2) = \frac{\sqrt{2}}{\sqrt{\pi}} \left\{ \left(-ze^{-\frac{1}{2}z^2} \right)\Big|_0^{\infty} + \int_0^{\infty} e^{-\frac{1}{2}z^2} dz \right\}$$

$$= 2\int_0^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz = 1.$$

□

Proposició 2.5.24. *Si $X \sim N(\mu, \sigma^2)$, i $a \neq 0$, llavors $aX + b \sim N(a\mu + b, a^2\sigma^2)$.*

Demostració: Sigui $Z \sim N(0, 1)$. Aleshores $-Z$ també és $N(0, 1)$, per simetria. Com a conseqüència per a qualsevol real d, tenim que $(-d)Z \sim dZ \sim |d|Z$. Observeu que

$$Y = aX + b = (a\sigma)\left(\frac{X-\mu}{\sigma}\right) + (b + a\mu) \sim |a|\sigma Z + (b + a\mu) \sim N(b + a\mu, a^2\sigma^2),$$

on hem fet servir la Proposició 2.5.22 d'estandardització. □

Proposició 2.5.25 (Combinacions lineals)**.** *Si $X_1, \ldots, X_n$ són variables aleatòries independents amb $X_i \sim N(\mu_i, \sigma_i^2), i = 1, \ldots, n$, aleshores*

$$\sum_{i=1}^{n} a_i X_i \sim N\left(\sum_{i=1}^{n} a_i \mu_i, \sum_{i=1}^{n} a_i^2 \sigma_i^2\right)$$

per a constants reals $a_1, \ldots, a_n$.

No demostrarem aquest resultat. Fixem-nos que en particular el teorema ens diu que si $X_1, \ldots, X_n$ són independents i idènticament distribuïdes segons $\mathrm{N}(\mu, \sigma^2)$ llavors

$$\sum_{i=1}^{n} X_i \sim \mathrm{N}(n\mu, n\sigma^2),$$

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \sim \mathrm{N}\left(\mu, \frac{\sigma^2}{n}\right).$$

Càlcul de probabilitats en distribucions normals

la Proposició 2.5.22 d'estandardització justifica que per conèixer els valors de la funció de distribució de qualsevol distribució normal $X \sim N(\mu, \sigma^2)$ només calgui conèixer els valors pel cas de la normal estàndard Z:

$$P(X \leq x) = P\left(\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right) = P\left(Z \leq \frac{x-\mu}{\sigma}\right) = \Phi\left(\frac{x-\mu}{\sigma}\right).$$

Així doncs, només cal tenir la taula de la distribució de la normal estàndard per poder fer càlculs amb qualsevol distribució normal.

De manera similar podem transformar quantils de la normal estàndard en quantils d'una normal qualsevol. Sigui $z_\alpha = \Phi^{-1}(\alpha)$ el quantil d'ordre α de la normal estàndard, i sigui $X \sim N(\mu, \sigma^2)$. Aleshores

$$\alpha = \Phi(z_\alpha) = \Pr(Z \leq z_\alpha) = \Pr(\mu + \sigma Z \leq \mu + \sigma z_\alpha) = \Pr(X \leq \mu + \sigma z_\alpha)$$

i es dedueix que

$$x_\alpha = \mu + \sigma z_\alpha$$

és el quantil d'ordre α de $X \sim N(\mu, \sigma^2)$.

La distribució log-normal

Molt sovint ens trobarem amb dades que presenten una distribució molt asimètrica i que mitjançant la seva transformació logarítmica es fan més simètriques. De fet, donada una distribució asimètrica, si considerem la transformació logarítmica sovint ens trobem amb una distribució aproximadament normal.

Definició 2.5.26. *Sigui X una variable aleatòria tal que $\log X \sim N(\mu, \sigma^2)$, aleshores la llei de X és diu* LOG-NORMAL *de paràmetres μ i σ^2. Escriurem que $X \sim logN(\mu, \sigma^2)$.*

El Teorema del Límit Central (que estudiarem més endavant) justifica que sigui aproximadament log-normal tota variable aleatòria que reculli l'efecte multiplicatiu de diverses causes positives independents i totes elles d'una mida similar:

$$R_n = \prod_{i=1}^{n} D_i \Rightarrow Y_n = \log R_n = \sum_{i=1}^{n} \log D_i \approx N(\mu, \sigma^2) \Rightarrow R_n \approx \mathrm{logN}(\mu, \sigma^2).$$

Aquest és el cas de moltes variables econòmiques, com ara el rendiment relatiu d'una acció al cap d'un mes, que és el producte dels rendiments relatius diaris de l'acció al llarg de tots els dies del mes.

Estudiem la distribució d'una variable aleatòria X amb llei log-normal. Sigui $Y = \log X$, és a dir $X = e^Y$. Com que la transformació és bijectiva tenim que X pren valors a $\mathbb{R}^+$ i llavors per $x > 0$ tenim

$$F_X(x) = P(X \leq x) = P(e^Y \leq x) = P(Y \leq \log x) = F_Y(\log x).$$

Derivant obtenim la funció de densitat de X:

$$f_X(x) = f_Y(\log x)\frac{1}{x} = \frac{1}{\sqrt{2\pi}\sigma x} e^{-\frac{(\log x-\mu)^2}{2\sigma^2}}.$$

La Figura 2.2 (a dalt a l'esquerra) mostra la funció de densitat de la log-normal per diferents combinacions dels paràmetres μ i σ (denotats per `m` i `s`, respectivament, al gràfic).

Una variable aleatòria X log-normal té moments de tots els ordres. Calculem-ne alguns:

$$\mathbb{E}(X) = \mathbb{E}(e^Y) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(z - \frac{(z-\mu)^2}{2\sigma^2}\right) dz$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(z-(\sigma^2+\mu))^2}{2\sigma^2} + \frac{\sigma^2(\sigma^2+2\mu)}{2\sigma^2}\right) dz = e^{\frac{\sigma^2+2\mu}{2}} = e^{\mu+\frac{\sigma^2}{2}}.$$

Per altra banda es pot calcular

$$\mathbb{E}(X^2) = e^{2\mu+2\sigma^2}$$

i per tant

$$\mathrm{Var}(X) = e^{2\mu+2\sigma^2} - e^{\mu+\frac{\sigma^2}{2}}{}^{2} = e^{2\mu+\sigma^2}(e^{\sigma^2}-1).$$

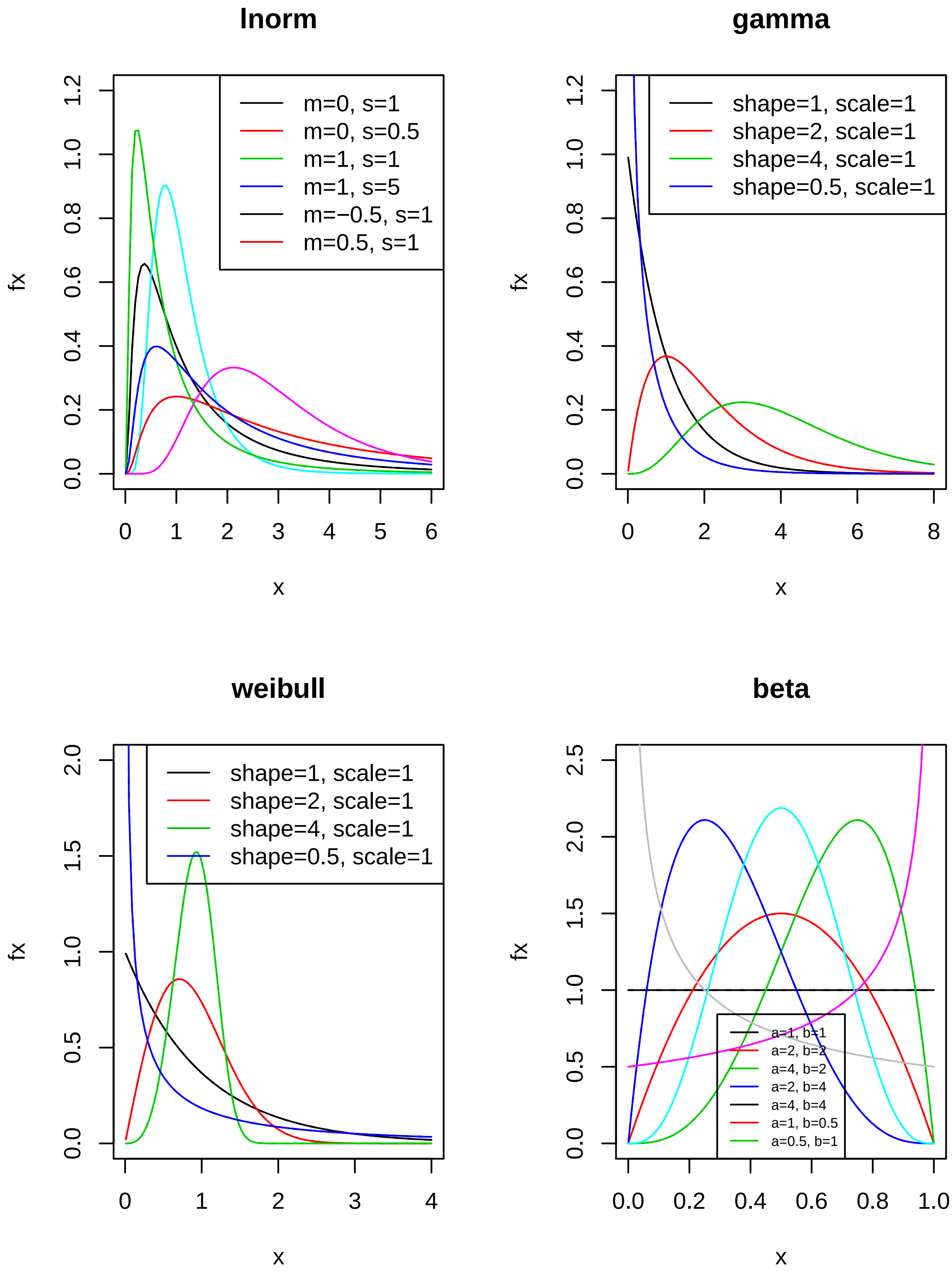


Figura 2.2: Funcions de densitat de diverses distribucions i combinacions de paràmetres.

La família de distribucions Gamma

Definició 2.5.27 (Funció Gamma)**.** *Si $\alpha > 0$ definim la* FUNCIÓ GAMMA *com*

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x}\, dx.$$

Propietats de la funció Gamma

- $\Gamma(1) = 1$.
- $\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1)$, es veu integrant per parts:

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x}\, dx = \left\{ \begin{array}{ll} u = x^{\alpha-1} & du = (\alpha-1)x^{\alpha-2} \\ dv = e^{-x}dx & v = -e^{-x} \end{array} \right\}$$

$$= \left(-e^{-x}x^{\alpha-1}\right)\Big|_0^\infty + \int_0^\infty (\alpha-1)x^{\alpha-2}e^{-x}\, dx = (0-0) + (\alpha-1)\Gamma(\alpha-1).$$

- $\Gamma(n) = (n-1)!$.
- $\Gamma(1/2) = \sqrt{\pi}$:

$$\Gamma(1/2) = \int_0^\infty x^{(1/2)-1} e^{-x}\, dx \overset{\left\{ \begin{array}{c} x = \frac{1}{2}y^2 \\ dx = y\, dy \end{array} \right\}}{=} \int_0^\infty \sqrt{2}\frac{1}{y} e^{-\frac{1}{2}y^2} y\, dy =$$

$$\sqrt{2}\sqrt{2\pi}\int_0^\infty \frac{1}{2\pi} e^{-\frac{1}{2}y^2}\, dy = \sqrt{2}\sqrt{2\pi}\frac{1}{2} = \sqrt{\pi}.$$

Hem fet servir que $\int_{-\infty}^\infty \frac{1}{2\pi} e^{-\frac{1}{2}y^2}\, dy = 1$.

Definició 2.5.28. *Una variable aleatòria X segueix una* LLEI GAMMA *de paràmetres $\alpha > 0$ i $\beta > 0$ si X té una funció de densitat definida per*

$$f(x|\alpha,\beta) = \left\{ \begin{array}{ll} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} & x > 0, \\ 0 & \textit{altrament}, \end{array} \right. ,$$

on α i β són els anomenats paràmetres de forma i d'escala respectivament. Escriurem que $X \sim \gamma(\alpha,\beta)$.

La llei Gamma és una distribució útil per a modelar durades de vida. Per exemple serveix per a mesurar el temps de supervivència de rates exposades a un alt nivell de radiació, les durades de transistors en una prova de vida accelerada, les durades de components electròniques o el temps entre microterratrèmols.

En determinats llibres podeu trobar una altra parametrització d'aquesta distribució, que s'obté definint $b = 1/\beta$:

$$f(x|\alpha, b) = \begin{cases} \frac{1}{\Gamma(\alpha)b^\alpha} x^{\alpha-1} e^{-\frac{x}{b}} & x > 0 \\ 0 & \text{altrament} \end{cases},$$

La Figura 2.2 (a dalt a la dreta) mostra la funció de densitat de la Gamma per diferents combinacions dels paràmetres α i $b = 1/\beta$ (denotats per `shape` i `scale`, respectivament, al gràfic).

Exemple 2.5.2.

Suposem $Z \sim \mathrm{N}(0, 1)$. Si definim $X = Z^2$ llavors

$$F_X(x) = \Pr(X \le x) = \Pr(-\sqrt{x} \le Z \le \sqrt{x}) = \Phi(\sqrt{x}) - \Phi(-\sqrt{x}).$$

$$f_X(x) = \varphi(\sqrt{x})\frac{1}{2\sqrt{x}} + \varphi(-\sqrt{x})\frac{1}{2\sqrt{x}} = \frac{1}{\sqrt{x}}\varphi(\sqrt{x}) = \frac{1}{\sqrt{2\pi}} x^{-\frac{1}{2}} e^{-\frac{x}{2}}.$$

Notem que prenent $\alpha = \beta = 1/2$ hem provat que la llei de X és $\gamma(1/2, 1/2)$ i que

$$\frac{\beta^\alpha}{\Gamma(\alpha)} = \frac{1}{\sqrt{2\pi}} \Rightarrow \Gamma(1/2) = \sqrt{\pi}.$$

La distribució $\gamma(1/2, 1/2)$ també s'anomena χ^2_1.

□

Un canvi en les unitats de mesura implica un canvi en el valor del paràmetre β però no afecta la forma de la distribució.

Exemple 2.5.3.

Sigui $X \sim \gamma(\alpha, \beta)$ i $Y = aX$, $a > 0$. Llavors

$$f_Y(y) = f_X\left(\frac{y}{a}\right)\frac{1}{a} = \frac{\beta^\alpha}{\Gamma(\alpha)}\left(\frac{y}{\alpha}\right)^{\alpha-1} e^{-\beta\frac{y}{a}}\frac{1}{a}$$

$$= \frac{\beta^\alpha}{\Gamma(\alpha)} y^{\alpha-1}\frac{1}{a^\alpha} e^{-\beta\frac{y}{a}} = \left(\frac{\beta}{a}\right)^\alpha \frac{1}{\Gamma(\alpha)} y^{\alpha-1} e^{-\frac{\beta}{a}y}.$$

Per tant $Y \sim \gamma(\alpha, \frac{\beta}{a})$.

□

Observació: La funció de distribució d'una Gamma no admet una expressió analítica:

$$F(x|\alpha,\beta) = \int_0^x \frac{\beta^\alpha}{\Gamma(\alpha)} u^{\alpha-1} e^{-\beta u}\, du.$$

Si definim la FUNCIÓ GAMMA INCOMPLETA (que està escrita en taules):

$$I(k,x) = \frac{1}{\Gamma(k)} \int_0^x u^{k-1} e^{-u}\, du,$$

es pot comprovar que $F(x|\alpha,\beta) = I(\alpha, \beta x)$.

Càlcul d'alguns moments

Calculem moments importants d'una variable aleatòria X que segueix una llei Gamma:

$$\mathbb{E}(X) = \int_0^\infty x\; f(x|\alpha,\beta)\, dx = \int_0^\infty x \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}\, dx$$

$$= \int_0^\infty \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha} e^{-\beta x}\, dx = \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha+1)}{\beta^{\alpha+1}} \int_0^\infty \frac{\beta^{\alpha+1}}{\Gamma(\alpha+1)} x^{\alpha} e^{-\beta x}\, dx$$

$$= \frac{\Gamma(\alpha+1)}{\Gamma(\alpha)} \frac{1}{\beta} \int_0^\infty f(x|\alpha+1,\beta)\, dx = \frac{\alpha\Gamma(\alpha)}{\Gamma(\alpha)} \frac{1}{\beta} = \frac{\alpha}{\beta}.$$

$$\mathbb{E}(X^2) = \frac{\alpha(\alpha+1)}{\beta^2}.$$

Amb els moments calculats podem trobar la variància fàcilment:

$$\mathrm{Var}(X) = \frac{\alpha(\alpha+1)}{\beta^2} - \frac{\alpha^2}{\beta^2} = \frac{\alpha}{\beta^2}.$$

Raonant com abans i aplicant el principi de inducció, és fàcil provar que

$$\mathbb{E}(X^k) = \frac{\alpha(\alpha+1)\cdots(\alpha+k-1)}{\beta^k}$$

per a tot k natural.

Casos particulars de les distribucions Gamma

- Si $\alpha = 1$ i $\beta > 0$ tenim la llei exponencial.

- Si $\alpha = \frac{n}{2}$, amb $n \in \mathbb{N}$ i $\beta = 1/2$ tenim una χ^2_n. En particular, recordeu que si $Z \sim N(0,1)$ aleshores $Z^2 \sim \chi^2_1$. Fent servir la propietat que diu que la suma de Gammas independents és també Gamma (ho provarem més endavant) es té que si $Z_1, \ldots, Z_n$ són $N(0,1)$ independents, aleshores
$$\sum_{i=1}^{n} Z_i^2 \sim \chi^2_n.$$
- Si $\beta = 1$ la funció de densitat és $f(x) = \frac{x^{\alpha-1}e^{-x}}{\Gamma(\alpha)}$ i llavors tenim una distribució Gamma amb un paràmetre: $\gamma(\alpha)$.
- Si $X \sim \gamma(\alpha, \beta)$ llavors $\beta X \sim \gamma(\alpha)$.
- Si $Y \sim \gamma(k)$ aleshores $2Y \sim \chi^2_{2k} \equiv \gamma(\frac{2k}{2}, 1/2) = \gamma(k, 1/2)$.
  Fixem-nos que si $2Y \sim \gamma(k, 1/2)$ llavors $\frac{1}{2}2Y = Y \sim \gamma(k)$.

Exemple 2.5.4.
Càlcul del coeficient d'apuntament de la Normal
Sigui $X \sim N(\mu, \sigma^2)$ i sigui $Z = (X - \mu)/\sigma$. El coeficient d'apuntament de X es defineix com $\mathbb{E}(Z^4)$. Sigui $Y = Z^2$. Hem vist que $Y \sim \chi^2_1 \equiv \gamma(1/2, 1/2)$. Per tant
$$\mathbb{E}(Y) = \frac{1/2}{1/2} = 1, \; V(Y) = \frac{1/2}{(1/2)^2} = 2.$$
Observeu que
$$\text{CAp}(X) = \mathbb{E}(Z^4) = \mathbb{E}(Y^2) = V(Y) + \mathbb{E}(Y)^2 = 2 + 1 = 3.$$
□

La distribució de Weibull

Aquest és un model bàsic per explicar el comportament de la vida d'un producte. És molt útil per estudiar duracions i fiabilitat de components.

Si una màquina està formada per diverses parts i la màquina falla, tant bon punt falla una de les seves parts, suposant que totes les avaries provenen d'una distribució similar, el model de Weibull és apropiat.

Aquesta distribució s'obté quan fem una potència d'una variable aleatoria exponencial. Sigui $X \sim \text{Exp}(\lambda)$. Aleshores la seva funció de densitat és
$$f_X(x) = \lambda e^{-\lambda x} I_{[0,\infty)}(x)$$

i la seva funció de distribució és

$$F_X(x) = \left(1 - e^{-\lambda x}\right) I_{[0,\infty)}(x).$$

Sigui $b > 0$. Definim $Y = X^b$. Calculem les seves funcions de distribució i de densitat. Si $y < 0$, $F_Y(y) = 0$ perquè $Y \geq 0$. Si $y \geq 0$,

$$F_Y(y) = \Pr(Y \leq y) = \Pr(X^b \leq y) = \Pr(X \leq y^{1/b}) = F_X(y^{1/b}) = 1 - e^{-\lambda y^{1/b}}.$$

Derivem per obtenir la funció de densitat de Y. Si $y < 0$ tenim $f_Y(y) = 0$. Si $y \geq 0$,

$$f_Y(y) = F'_Y(y) = \frac{\lambda}{b} y^{(1/b)-1} e^{-\lambda y^{1/b}}$$

Usualment, en comptes de fer servir els paràmetres λ i b, es fan servir els paràmetres

$$\beta = \frac{1}{b} \text{ i } \alpha = \frac{1}{\lambda^b}.$$

Amb aquesta parametrització es té la següent definició.

Definició 2.5.29. *Una variable aleatòria X segueix una* Llei Weibull *amb paràmetres $\alpha > 0$ i $\beta > 0$ si*

$$f(x|\alpha, \beta) = \begin{cases} \frac{\beta}{\alpha^\beta} x^{\beta-1} e^{-(\frac{x}{\alpha})^\beta} & \textit{si } x > 0 \\ 0 & \textit{altrament} \end{cases},$$

on α s'anomena paràmetre d'escala o vida característica i β s'anomena paràmetre de forma.

La Figura 2.2 (abaix a l'esquerra) mostra la funció de densitat de la Weibull per diferents combinacions dels paràmetres α i β (denotats per `scale` i `shape`, respectivament, al gràfic).

Si $X \sim \text{Weibull}(\alpha, \beta)$ la seva funció de distribució acumulada és

$$F(x|\alpha, \beta) = 1 - e^{-(\frac{x}{\alpha})^\beta}$$

si $x \geq 0$, i val 0 si $x < 0$.

Casos particulars

- Si $\beta = 1$ llavors $\text{W}(\alpha, 1)$ és una exponencial de mitjana α.
- Si $\beta = 2$ llavors $\text{W}(\alpha, 2) = \text{Rayleigh}(\alpha)$.
- Si $3 \leq \beta \leq 4$ llavors $\text{W}(\alpha, \beta) \cong$Normal.

Altres parametritzacions

- Si $\lambda = \frac{1}{\alpha^\beta}$ llavors
$$f(x|\lambda,\beta) = \lambda\beta x^{\beta-1}e^{-\lambda x^\beta},$$
$$F(x|\lambda,\beta) = 1 - e^{-\lambda x^\beta}.$$
- Si $\theta = \frac{1}{\lambda} = \alpha^\beta$ llavors
$$f(x|\theta,\beta) = \frac{\beta}{\theta}x^{\beta-1}e^{-\frac{x^\beta}{\theta}},$$
$$F(x|\theta,\beta) = 1 - e^{-\frac{x^\beta}{\theta}}.$$

Propietats

1. α és el percentil del 63.2% per a qualsevol valor de β.

   En efecte,
$$0.632 = F(t_{0.632}) = 1 - e^{-(\frac{t_{0.632}}{\alpha})^\beta}$$
$$0.368 = e^{-(\frac{t_{0.632}}{\alpha})^\beta}$$
$$\ln 0.368 = -\left(\frac{t_{0.632}}{\alpha}\right)^\beta,$$
   tenint en compte que $\ln 0.368 \approx -1$,
$$1 = \left(\frac{t_{0.632}}{\alpha}\right)^\beta$$
   Per tant $\alpha = t_{0.632}$.

2. Si $X \sim \mathrm{W}(\alpha,\beta)$ aleshores $Y = X^\beta \sim \mathrm{Exp}(\alpha^{-\beta})$, és a dir: una exponencial de mitjana α^β.

   En efecte,
$$F_Y(y) = \Pr(Y \leq y) = \Pr(X^\beta \leq y) = \Pr(X \leq y^{1/\beta}) = F_X(y^{1/\beta}) =$$
$$1 - e^{-\left(\frac{y^{1/\beta}}{\alpha}\right)^\beta} = 1 - e^{-\frac{y}{\alpha^\beta}},$$
   i per tant $Y \sim \mathrm{Exp}(\alpha^{-\beta})$.

Càlcul de moments importants

$$\mathbb{E}(X) = \int_0^\infty x\frac{\beta}{\alpha^\beta}x^{\beta-1}e^{-(\frac{x}{\alpha})^\beta}\,dx = \int_0^\infty \frac{\beta}{\alpha^\beta}x^\beta e^{-(\frac{x}{\alpha})^\beta}\,dx$$

$$= \left\{ \begin{array}{l} (\frac{x}{\alpha})^\beta = z \\ \beta \frac{x^{\beta-1}}{\alpha^\beta} dx = dz \\ x = \alpha z^{\frac{1}{\beta}} \end{array} \right\} = \int\limits_0^\infty x e^{-z}\, dz = \int\limits_0^\infty \alpha z^{\frac{1+\beta}{\beta}-1} e^{-z}\, dz$$

$$= \alpha\Gamma(\frac{1}{\beta}+1) \underbrace{\int\limits_0^\infty \frac{1}{\Gamma(\frac{1}{\beta}+1)} z^{\frac{1+\beta}{\beta}-1} e^{-z}\, dz}_{1} = \alpha\Gamma\left(\frac{1}{\beta}+1\right).$$

Anàlogament trobem $\mathbb{E}(X^2)$ i, després, $\mathrm{Var}(X)$:

$$\mathrm{Var}(X) = \alpha^2 \left(\Gamma\left(1+\frac{2}{\beta}\right) - \Gamma^2\left(1+\frac{1}{\beta}\right) \right).$$

La família de distribucions Beta

Siguin $\alpha > 0$ i $\beta > 0$. Calcularem el valor de $\Gamma(\alpha)\Gamma(\beta)$:

$$\Gamma(\alpha)\Gamma(\beta) = \int_0^\infty x^{\alpha-1} e^{-x} dx \int_0^\infty y^{\beta-1} e^{-y} dy = \int_0^\infty \int_0^\infty x^{\alpha-1} y^{\beta-1} e^{-(x+y)} dy dx$$

$$\left\{ \begin{array}{c} s = x+y \\ y = s-x \\ ds = dy \end{array} \right\} \overset{}{=} \int_0^\infty \int_x^\infty x^{\alpha-1}(s-x)^{\beta-1} e^{-s} ds dx = \int_0^\infty \int_0^s x^{\alpha-1}(s-x)^{\beta-1} e^{-s} dx ds$$

$$= \int_0^\infty s^{\alpha+\beta-2} e^{-s} \int_0^s \frac{x^{\alpha-1}}{s^{\alpha-1}} \frac{(s-x)^{\beta-1}}{s^{\beta-1}} dx ds$$

$$= \int_0^\infty s^{\alpha+\beta-1} \frac{1}{s} e^{-s} \int_0^s \left(\frac{x}{s}\right)^{\alpha-1} \left(1-\frac{x}{s}\right)^{\beta-1} dx ds$$

$$\left\{ \begin{array}{c} t = x/s \\ dt = (1/s)dx \end{array} \right\} = \int_0^\infty s^{\alpha+\beta-1} e^{-s} \int_0^1 t^{\alpha-1}(1-t)^{\beta-1} dt ds$$

$$= \int_0^\infty s^{\alpha+\beta-1} e^{-s} ds \int_0^1 t^{\alpha-1}(1-t)^{\beta-1} dt = \Gamma(\alpha+\beta) \int_0^1 t^{\alpha-1}(1-t)^{\beta-1} dt.$$

A partir d'aquesta igualtat, es defineix la funció *beta*:

$$B(\alpha,\beta) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1} dt = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}.$$

Observeu també que

$$\int_0^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} t^{\alpha-1}(1-t)^{\beta-1} dt = 1,$$

d'on es segueix que la funció

$$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} t^{\alpha-1}(1-t)^{\beta-1} I_{[0,1]}(t)$$

és una funció de densitat. Això permet fer la següent definició.

Definició 2.5.30. *Una variable aleatòria X direm que segueix una* LLEI BETA *amb paràmetres $\alpha > 0$ i $\beta > 0$ si*

$$f(x|\alpha,\beta) = \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} & \textit{si } 0 < x < 1 \\ 0 & \textit{si } x \leq 0 \end{cases}.$$

Un cas particular d'aquesta distribució és $\mathrm{U}[0,1]=\mathrm{B}(1,1)$.

La Figura 2.2 (abaix a la dreta) mostra la funció de densitat de la beta per diferents combinacions dels paràmetres a i b (a les funcions d'R corresponents a la distribució beta s'anomenen `shape1` i `shape1`).

Es pot provar que

$$\mathbb{E}(X) = \frac{\alpha}{\alpha+\beta}, \ \mathrm{Var}(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}.$$

Famílies exponencials

La variable aleatòria X PERTANY A LA FAMÍLIA EXPONENCIAL si la seva funció de densitat f (si és contínua) o la seva funció de probabilitat (si és discreta), depèn d'un paràmetre $\theta \in \Theta \subseteq \mathbb{R}^p$ i es pot escriure així

$$f(x|\theta) = h(x)c(\theta)\exp\left(\sum_{j=1}^{k} w_j(\theta)t_j(x)\right)$$

per a certes funcions h, c, w_j y t_j, $j = 1, \ldots, k$. Si $p = k$ i $w_j(\theta) = \theta_j$, $j = 1, \ldots, p$, aleshores direm que la família exponencial està PARAMETRITZADA DE FORMA NATURAL.

Exemple 2.5.5.

Exemples de famílies exponencials són aquests: binomial, geomètrica, Poisson, binomial negativa, exponencial, normal, gamma, beta.

□

2.6 Funcions d'una variable aleatòria univariant

Si X és una variable aleatòria amb funció de distribució F_X aleshores qualsevol funció *prou bona* (mesurable, tècnicament) de X, $g(X)$, és també una variable aleatòria (g ha de ser mesurable de $\mathbb{R}$ a $\mathbb{R}$).

Denotem per $Y = g(X)$ i intentem descriure el comportament probabilístic de Y en termes del de X i de la funció g.

Per qualsevol conjunt A tenim que $P(Y \in A) = P(g(X) \in A)$ i per tant la distribució de Y depèn de F_X i de g.

De fet es pot considerar que g és una funció del suport de X en el suport de Y (que anomenarem $\mathcal{X}$ i $\mathcal{Y}$, respectivament):

$$g : \mathcal{X} \longrightarrow \mathcal{Y}.$$

Associem ara a g una funció inversa g^{-1} definida per subconjunts borelians de $\mathcal{Y}$, tal que si $A \subseteq \mathcal{Y}$, aleshores

$$g^{-1}(A) = \{x \in \mathcal{X} : g(x) \in A\} \subset \mathcal{X}.$$

En particular $g^{-1}(\{y\}) = \{x \in \mathcal{X} : g(x) = y\}$. Si $g^{-1}(\{y\}) = \{x\}$ únic, escriurem $g^{-1}(y) = x$.

Per tant, per qualsevol conjunt borelià A:

$$P(Y \in A) = P(g(X) \in A) = P(X \in \{x \in \mathcal{X} : g(x) \in A\}) = P(X \in g^{-1}(A)),$$

i això ens defineix la llei de Y.

Teorema 2.6.1 (Cas discret)**.** *Si X és variable aleatòria discreta, $\mathcal{X}$ és numerable i $Y = g(X)$ també és discreta amb $\mathcal{Y} = \{y : y = g(x), x \in \mathcal{X}\}$, llavors*

$$P(Y = y) = \begin{cases} \sum_{x \in g^{-1}(y)} P(X = x) & \text{si } y \in \mathcal{Y} \\ 0 & \text{si } y \notin \mathcal{Y} \end{cases}$$

És a dir: es tracta d'identificar $g^{-1}(y)$ per cada $y \in \mathcal{Y}$ i sumar les probabilitats corresponents. Així,

$$F_Y(y) = P(Y \leq y) = P(g(x) \leq y) = P(x \in \mathcal{X} : g(x) \leq y) = \sum_{x : g(x) \leq y} P(X = x).$$

Exemple 2.6.1.
Sigui X definida com segueix:

$$X = \begin{cases} -2 & \text{amb probabilitat } p_1 \\ -1 & \text{amb probabilitat } p_2 \\ 0 & \text{amb probabilitat } p_3 \\ 1 & \text{amb probabilitat } p_4 \\ 2 & \text{amb probabilitat } p_5 \end{cases}$$

Llavors la funció de distribució de X ve donada per:

$$F_X(x) = \begin{cases} 0 & \text{si } x < -2 \\ p_1 & \text{si } -2 \leq x < -1 \\ p_1 + p_2 & \text{si } -1 \leq x < 0 \\ p_1 + p_2 + p_3 & \text{si } 0 \leq x < 1 \\ p_1 + p_2 + p_3 + p_4 = 1 - p_5 & \text{si } 1 \leq x < 2 \\ 1 & \text{si } 2 \leq x \end{cases}$$

Considerem ara $Y = g(X) = X^2$, llavors la següent taula resumeix la situació:

x	-2	-1	0	1	2
$P(X = x)$	p_1	p_2	p_3	p_4	p_5
$y = g(x) = x^2$	4	1	0	1	4

Tenim doncs que $g^{-1}(0) = \{0\}, g^{-1}(1) = \{-1, 1\}, g^{-1}(4) = \{-2, 2\}$. Llavors:
$P(Y = 0) = P(X = 0) = p_3$,
$P(Y = 1) = P(X = 1) + P(X = -1) = p_2 + p_4$,
$P(Y = 4) = P(X = -2) + P(X = 2) = p_1 + p_5$.
En forma de taula:

y	0	1	4
$P(Y = y)$	p_3	$p_2 + p_4$	$p_1 + p_5$

Finalment, la funció de distribució de Y és aquesta:

$$F_Y(y) = P(Y \leq y) = \begin{cases} 0 & \text{si } y < 0 \\ p_3 & \text{si } 0 \leq y < 1 \\ p_3 + p_2 + p_4 & \text{si } 1 \leq y < 4 \\ 1 & \text{si } 4 \leq y \end{cases}$$

□

Teorema 2.6.2 (Cas continu)**.** *Sigui X una variable aleatòria contínua amb funció de distribució F_X i densitat f_X i sigui $g : \mathbb{R} \longrightarrow \mathbb{R}$. Aleshores la funció de distribució de $Y = g(X)$ és*

$$F_Y(y) = P(Y \leq y) = P(g(X) \leq y) = \int_{\{x \in \mathcal{X} : g(x) \leq y\}} f_X(x)\, dx$$

Si F_Y és contínua i diferenciable, es defineix la densitat de Y com

$$f_Y(y) = \frac{d}{dy} F_Y(y).$$

Exemple 2.6.2.

Sigui $X \sim U[-1, 1]$ i considerem $Y = X^2$. Llavors:

$$f_X(x) = \begin{cases} \frac{1}{2} & \text{si } -1 < x < 1 \\ 0 & \text{altrament} \end{cases}$$

$$F_X(x) = \int_{-1}^{x} f(u)\, du = \frac{1}{2} u \Big|_{-1}^{x} = \begin{cases} 0 & \text{si } x < -1 \\ \frac{1}{2}(x+1) & \text{si } -1 \leq x \leq 1 \\ 1 & \text{si } x > 1 \end{cases}$$

Calculem ara la funció de distribució i la funció de densitat de Y, conegudes les de X. Obtenim:

$$F_Y(y) = P(Y \leq y) = P(X^2 \leq y) = P(-\sqrt{y} \leq X \leq \sqrt{y}) = F_X(\sqrt{y}) - F_X(-\sqrt{y})$$

$$= \frac{1}{2}(\sqrt{y}+1) - \frac{1}{2}(-\sqrt{y}+1) = \begin{cases} 0 & \text{si } y \leq 0 \\ \sqrt{y} & \text{si } 0 \leq y \leq 1 \\ 1 & \text{si } y \geq 1 \end{cases}$$

$$f_Y(y) = \begin{cases} \frac{1}{2\sqrt{y}} & \text{si } 0 \leq y \leq 1 \\ 0 & \text{altrament} \end{cases}$$

$\square$

Teorema 2.6.3. *Sigui X una variable aleatòria amb funció de distribució F i funció de densitat (o de probabilitat) f i suport $\mathcal{X}$.*

Sigui $g : \mathbb{R} \longrightarrow \mathbb{R}$ i $Y = g(X)$, amb suport $\mathcal{Y} = g(\mathcal{X}) = \{y : y = g(x), x \in \mathcal{X}\}$.

(a) Si g és creixent sobre $\mathcal{X}$, llavors

$$F_Y(y) = F_X(g^{-1}(y)), \text{ per tot } y \in \mathcal{Y}.$$

(b) Si g és decreixent sobre $\mathcal{X}$ i X és contínua, llavors

$$F_Y(y) = 1 - F_X(g^{-1}(y)), \text{ per tot } y \in \mathcal{Y}.$$

Demostració: En ambdós casos g és bijectiva i per tant per cada x hi ha un únic y i viceversa.

(a) Com que g és creixent:

$$\{x \in \mathcal{X} : g(x) \leq y\} = \{x \in \mathcal{X} : g^{-1}(g(x)) \leq g^{-1}(y)\} =$$
$$\{x \in \mathcal{X} : x \leq g^{-1}(y)\}.$$

Llavors

$$F_Y(y) = P(Y \leq y) = P(X \in \{x \in \mathcal{X} : g(x) \leq y\}) =$$
$$P(X \in \{x \in \mathcal{X} : x \leq g^{-1}(y)\}) = P(X \leq g^{-1}(y)) = F_X(g^{-1}(y)).$$

(b) Com que g és decreixent:

$\{x \in \mathcal{X} : g(x) \leq y\} = \{x \in \mathcal{X} : g^{-1}(g(x)) \geq g^{-1}(y)\} = \{x \in \mathcal{X} : x \geq g^{-1}(y)\}$.

Llavors $F_Y(y) = P(Y \leq y) = 1 - F_X(g^{-1}(y))$, on hem usat la continuïtat de X en l'última igualtat.

□

Exemple 2.6.3.

Sigui $X \sim U[0,1]$. Llavors:

$$f_X(x) = \begin{cases} 1 & \text{si } 0 < x < 1 \\ 0 & \text{altrament} \end{cases} \qquad F_X(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ x & \text{si } 0 < x < 1 \\ 1 & \text{si } x \geq 1 \end{cases}$$

Considerem $Y = -\log X = g(X)$. Si volem calcular la funció de distribució de Y necessitem càlculs previs:

Fixem-nos que $g'(x) = -\frac{1}{x} < 0$. Per tant g és decreixent.

A més $\mathcal{X} = [0,1]$ i $\mathcal{Y} = [0,\infty)$.

Si $y > 0, y = -\log x$. Llavors $x = e^{-y} = g^{-1}(y)$.
Finalment:

$$F_Y(y) = 1 - F_X(g^{-1}(y)) = 1 - F_X(e^{-y}) = \begin{cases} 0 & \text{si } y \leq 0 \\ 1 - e^{-y} & \text{si } y > 0 \end{cases}$$

Per tant veiem que $Y \sim \exp(1)$. $\square$

Teorema 2.6.4. *Sigui X una variable aleatòria amb funció de densitat $f_X(x)$ i sigui $Y = g(X)$ essent g una funció monòtona. Siguin $\mathcal{X}, \mathcal{Y}$ els suports de X i de Y respectivament.*

Suposem que f_X és contínua sobre $\mathcal{X}$ i que $g^{-1}(y)$ té derivada contínua sobre $\mathcal{Y}$. Aleshores la funció de densitat de Y ve donada per

$$f_Y(y) = \begin{cases} f_X(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right| = f_X(g^{-1}(y)) \frac{1}{|g'(g^{-1}(y))|} & \textit{si } y \in \mathcal{Y} \\ 0 & \textit{altrament} \end{cases}$$

Demostració: El resultat s'obté fàcilment usant el teorema anterior:

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \begin{cases} f_X(g^{-1}(y)) \frac{d}{dy} g^{-1}(y) & \text{si } g \text{ creixent} \\ -f_X(g^{-1}(y)) \frac{d}{dy} g^{-1}(y) & \text{si } g \text{ decreixent} \end{cases}$$

$\square$

Si definim $h(y) = g^{-1}(y)$, aleshores l'expressió anterior es pot escriure com

$$f_Y(y) = f_X(h(y))|h'(y)|.$$

Observació: Una transformació d'una variable aleatòria contínua no produeix necessàriament una variable aleatòria contínua.

Suposem que X és una variable aleatòria contínua que pren valors a $[a, b]$. Sigui c, d tals que $a < c < d < b$ i $P(c < X < d) > 0$.

Definim $r(X) = Y$ tal que $r(x) = k = \frac{c+d}{2}$ per a tot $x \in [c, d]$ i $r(x) = x$ altrament.

$$Y = r(X) = \begin{cases} k = \frac{c+d}{2} & \text{si } x \in [c, d] \\ x & \text{altrament} \end{cases}$$

Llavors $P(Y = \frac{c+d}{2}) = P(c \leq X \leq d) > 0$ i per tant Y no és contínua.

Exemple 2.6.4.

Sigui $X \sim U[0, 1]$. Definim:

$$Y_i = r_i(X) = \begin{cases} X^2 & \text{per } i = 1 \\ -X^3 & \text{per } i = 2 \\ \sqrt{X} & \text{per } i = 3 \end{cases}$$

Volem calcular la funció de distribució i la funció de densitat, si existeix, de Y_i:

(a) $Y_1 = r_1(X) = X^2$.

És clar que r_1 és estrictament creixent en [0,1]. Llavors:

$$G_1(y) = P(Y_1 \leq y) = P(X^2 \leq y) = P(X \leq \sqrt{y})$$

$$= \int_0^{\sqrt{y}} f(u)\, du = \begin{cases} 0 & \text{si } y \leq 0 \\ \sqrt{y} & \text{si } 0 < y < 1 \\ 1 & \text{si } y \geq 1 \end{cases}$$

Considerem $s_1(y) = r_1^{-1}(y) = \sqrt{y}$, la inversa de la transformació r_1. Ara podem escriure:

$$g_1(y) = \begin{cases} f_1(s_1(y)) \left| \frac{d}{dy} s_1(y) \right| = f_1(\sqrt{y}) \left| \frac{d}{dy} \sqrt{y} \right| = \frac{1}{2\sqrt{y}} & \text{si } 0 < y < 1 \\ 0 & \text{altrament} \end{cases}$$

(b) $Y_2 = r_2(X) = -X^3$.

Tenim que si $0 \leq X \leq 1$ llavors $-1 \leq Y \leq 0$. Considerem aquí $s_2(y) = (-y)^{\frac{1}{3}}$. Llavors:

$$g_2(y) = f_1(s_2(y)) \left| \frac{d}{dy} s_2(y) \right| = f_1(\sqrt[3]{-y}) \left| \frac{d}{dy} \sqrt[3]{-y} \right|$$

$$= 1 \cdot \left| \frac{1}{3} (-y)^{\frac{1}{3}-1} \right| = \begin{cases} \frac{1}{3}(-y)^{\frac{-2}{3}} & \text{si } -1 < y < 0 \\ 0 & \text{altrament} \end{cases}$$

(c) $Y_3 = r_3(X) = \sqrt{X}$.

Tenim ara $s_3(y) = y^2$. Anàlogament als casos anteriors obtenim:

$$g_3(y) = \begin{cases} 2y & \text{si } 0 < y < 1 \\ 0 & \text{altrament} \end{cases}$$

$\square$

2.6.1 Transformacions basades en la funció de distribució

Sigui F la funció de distribució acumulada d'una variable aleatòria contínua X. Suposarem que F és estrictament creixent per tot x tal que $0 < F(x) < 1$.

Donats uns valors $x_1, x_2, x_3, \ldots \in \mathcal{X}$, podem obtenir els valors $u_1, u_2, u_3, \ldots$ mitjançant la transformació de la integral de la probabilitat $u_i = F(x_i)$.

Recíprocament, donats $u_1, u_2, \ldots$ valors entre 0 i 1 podem obtenir els valors $x_1, x_2, \ldots$ mitjançant la transformació inversa $x_i = F^{-1}(u_i)$.

Teorema 2.6.5. *Si $X \sim F$ aleshores $Y = F(X) \sim U[0,1]$. Si $U \sim U[0,1]$, aleshores $X = F^{-1}(U) \sim F$.*

Demostració: Per veure la primera afirmació considerem la transformació $g \equiv F$, és a dir $Y = F(X)$.

Com que F és estrictament creixent, existeix un únic x tal que $F(x) = y$, i a més: $Y \leq y \leftrightarrow X \leq x$. Per tant, per $0 < y < 1$ tenim:

$$G(y) = P(Y \leq y) = P(F(X) \leq y) = P(X \leq F^{-1}(y)) = F(F^{-1}(y)) = y.$$

Veiem doncs que G és la funció de distribució d'una variable aleatòria $U[0,1]$.

Recíprocament, suposem que $U \sim U[0,1]$. Sigui $X = F^{-1}(U)$. Llavors

$$P(X \leq x) = P(F^{-1}(U) \leq x) = P(U \leq F(x)) = F(x),$$

d'on veiem que la funció de distribució de X és F. $\square$

Aplicació a una prova d'ajustament

Hem observat $x_1, x_2, \ldots, x_n$ i volem saber si és raonable pensar que les observacions x_i provenen d'una variable aleatòria amb funció de distribució F (suposant F coneguda).

Comencem calculant $u_i = F(x_i), i = 1, \ldots, n$ i comprovem si els valors $u_1, \ldots, u_n$ semblen una mostra aleatòria de nombres aleatoris entre 0 i 1.

Per exemple, suposem que $F(x) = 1 - e^{\frac{-x}{\theta}}$.

Calculem $u_i = 1 - e^{\frac{-x_i}{\theta}}, i = 1, 2, \ldots, n$ i si els valors $u_1, \ldots, u_n$ estan aleatòriament i uniformement repartits entre 0 i 1 podrem concloure que provenen d'una $U[0,1]$ i per tant les x_i provindran d'una exponencial de paràmetre θ.

Si el valor de θ fos desconegut llavors el calcularíem com $\overline{\theta} = \overline{X}_n$ *(més endavant veurem perquè).*

Simulació d'una distribució contínua

Mitjançant un generador de nombres aleatoris, per exemple amb la comanda RANDOM d'alguns paquets informàtics, obtenim n observacions independents $u_1, u_2, \ldots, u_n$ d'una $U[0,1]$.

Llavors els valors $x_i = F^{-1}(u_i)$, $i = 1, 2, \ldots, n$, són observacions independents d'una variable aleatòria amb funció de distribució F.

Exemple 2.6.5.
Generem 10 valors d'una exponencial de mitjana $\theta = \frac{1}{2}$. Fem servir

$$F(x) = 1 - e^{\frac{-x}{\theta}} = 1 - e^{-2x}, \text{ si } x > 0.$$

Fixem-nos que

$$Y = F(X) = 1 - e^{\frac{-X}{\theta}} \Rightarrow 1 - Y = e^{-\frac{X}{\theta}} \Rightarrow X = -\theta \cdot \log(1 - Y).$$

Per tant, si $Y \sim U[0, 1]$ llavors $-\theta \cdot \log(1 - Y) \sim \exp(\theta)$.

Seleccionem 10 valors u_i, per $i = 1, \ldots, 10$, a l'atzar:

$$0.2214, 0.8259, 0.3403, 0.2439, 0.1343, 0.9385, 0.2584, 0.767, 0.0007, 0.6333.$$

Llavors els x_i corresponents per $\theta = \frac{1}{2}$ són:

$$0.1251, 0.8741, 0.208, 0.1398, 0.0721, 1.3944, 0.1495, 0.7284, 0.0004, 0.5016.$$

□

Exemple 2.6.6.
Generem 3 valors independents d'una variable aleatòria X amb funció de densitat $f(x) = \frac{1}{2}(2 - x)$, si $0 < x < 2$.

La funció de distribució de X és

$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ \frac{1}{2}(2x - \frac{x^2}{2}) = x - \frac{x^2}{4} & \text{si } 0 \leq x < 2 \\ 1 & \text{si } x \geq 2 \end{cases}$$

Considerem $Y = F(X) = X - \frac{X^2}{4}$.

Tenim: $y = x - \frac{x^2}{4} \Leftrightarrow x^2 - 4x + 4y = 0 \Leftrightarrow x = 2 \pm 2\sqrt{1 - y}$. Com que $0 < x < 2$, només és possible la solució $x = 2 - 2\sqrt{1 - y}$.

Per tant $X = 2(1 - \sqrt{1 - Y})$. Ara si escollim a l'atzar els valors 0.6808, 0.0423, 0.0155, els valors generats són:

$$x_1 = 0.87, x_2 = 0.04, x_3 = 0.016.$$

□