Probabilitat

J. M. Aroca
A. Miralles

PID_00160375

Universitat Oberta de Catalunya
www.uoc.edu

Índex

1. Introducció: Tècniques de comptar

En moltes experiències, el càlcul d'una probabilitat va lligat a la quantitat de possibilitats diferents que té un cert aspecte de l'experiència. Per exemple, sabem que en llançar un dau perfecte la probabilitat que surti un 2 és $\frac{1}{6}$ ja que hi ha 6 possibles resultats. Però aquest és un cas molt senzill i de vegades aquest recompte de resultats no és tan simple. Per exemple, si ens preguntem de quantes maneres podem anar caminant a un cinema que es troba a 10 cruïlles de casa nostra, haurem d'especificar si volem anar pel camí més curt o no, i si volem passar per un cert punt del recorregut, etc. Aquest primer capítol el dediquem a explicar les tècniques de comptar més bàsiques. A partir d'exemples, anirem introduint els conceptes.

Exemple 1.1 *Comencem pensant en un conjunt de 10 elements:* $A = \{a_1, \cdots a_{10}\}$*. Considerem* 4 *elements d'aquest conjunt:* ***mostra de mida*** 4*. Escrivim* 5 *mostres d'exemple que anomenem:* $m1$*,* $m2$*,* $m3$*,* $m4$ *i* $m5$*.*

$$m1 = a_1a_1a_2a_3 \quad m2 = a_1a_1a_3a_2 \quad m3 = a_1a_{10}a_3a_2 \quad m4 = a_1a_{10}a_2a_3 \quad m5 = a_1a_9a_6a_5$$

Ens fixem en alguns aspectes. Hi ha mostres que tenen elements repetits com $m1$ i $m2$. Hi ha mostres en què l'única diferència que tenen entre elles és l'ordre dels elements, com $m3$ i $m4$. A l'hora de comptar el nombre de mostres que podem fer haurem de tenir en compte aquests aspectes.

A continuació veurem els tipus de mostres de m elements que es poden formar en un conjunt de n elements $A = \{a_1, a_2, \cdots a_n\}$.

- **Definició 1.1** *Mostra de mida m, ordenada i sense repetició (o reemplaçament).*

  Per formar la mostra no podem repetir els elements del conjunt A i si tenim dues mostres que tenen els mateixos elements però ordenats de forma diferent, les considerem diferents. A l'exemple 1.1, $m3$, $m4$ i $m5$ són mostres d'aquest tipus.

- **Definició 1.2** *Mostra de mida m, ordenada i amb repetició (o reemplaçament).*

  Per formar la mostra podem repetir els elements del conjunt A, i si tenim dues mostres que tenen els mateixos elements però ordenats de forma diferent, les considerem diferents. A l'exemple 1.1 totes les mostres són d'aquest tipus.

Repetició i reemplaçament

Els mots **repetició** i **reemplaçament** s'utilitzen indistintament. El mot repetició ens diu que hi pot haver elements repetits dins d'una mateixa mostra. També s'utilitza el mot reemplaçament perquè aquest fet de vegades va lligat a la manera que s'ha realitzat l'experiència. Per exemple, si en una experiència hem de treure dues cartes d'una baralla i després de treure la primera carta anotem el resultat i la tornem a deixar a la baralla (remplaçament), a la segona extracció podem obtenir la mateixa carta que abans.

- **Definició 1.3** *Mostra de mida m, no ordenada i sense repetició (o reemplaçament).*

  Per formar la mostra no podem repetir els elements del conjunt A, i si tenim dues mostres que tenen els mateixos elements però ordenats de forma diferent, les considerem la mateixa. A l'exemple 1.1, $m3$ i $m4$ són la mateixa mostra.

- **Definició 1.4** *Mostra de mida m, no ordenada i amb repetició (o reemplaçament).*

  Per formar la mostra podem repetir els elements del conjunt A, i si tenim dues mostres que tenen els mateixos elements però ordenats de forma diferent, les considerem la mateixa. A l'exemple 1.1, $m1$ i $m2$ representen la mateixa mostra i $m3$ és la mateixa mostra que $m4$.

Veiem quantes mostres podem formar de cada un dels tipus anteriors.

1.1. Mostres ordenades amb repetició.Variacions amb repetició

Si ens fixem en l'exemple 1.1, podem pensar que per formar una mostra d'aquest tipus hem d'omplir $m = 4$ posicions. A la primera posició podem posar qualsevol dels 10 elements del conjunt A, tenim 10 possibilitats. Un cop hem omplert la primera posició, a la segona posició també podem posar qualsevol dels 10 elements del conjunt A i per cada una d'aquestes possibilitats en tenim 10 de diferents de la primera posició. Seguint aquest raonament, veiem que podem formar $10 \cdot 10 \cdot 10 \cdot 10$ mostres. Anomenem variacions amb repetició de 10 elements agafats de 4 en 4, $VR_{10,4} = 10^4$.

En general, si partim d'un conjunt $A = \{a_1, a_2, \cdots a_n\}$ amb n elements, el nombre de mostres de mida m ordenades i amb repetició que es poden formar és

$$VR_{n,m} = n^m.$$

Exemple 1.2 *Quantes paraules de mida* 3 *es poden formar amb els elements del conjunt* $\{0,1\}$*?*

En un conjunt de 2 *elements, hem de trobar les mostres de mida* 3 *ordenades i amb repetició,* $VR_{2,3} = 2^3 = 8$.

000 001 010 100 011 101 110 111

1.2. Mostres ordenades sense repetició. Variacions. Permutacions

Tornem a l'exemple 1.1. A partir del conjunt $A = \{a_1,a_2,\cdots a_{10}\}$ volem formar mostres de mida 4 que no tinguin elements repetits. Hem d'omplir $m = 4$ posicions. A la primera posició podem posar qualsevol dels 10 elements del conjunt A, tenim 10 possibilitats. Un cop hem omplert la primera posició, a la segona posició només podem posar 9 elements del conjunt A, ja que no podem repetir l'element que hem posat a la primera posició. Seguint aquets raonament veiem que podem formar $10 \cdot 9 \cdot 8 \cdot 7$ mostres. Anomenem aquesta quantitat variacions de 10 elements agafats de 4 en 4, $V_{10,4} = 10 \cdot 9 \cdot 8 \cdot 7 = 5040$.

En general, si partim del conjunt $A = \{a_1,a_2,\cdots a_n\}$, el nombre de mostres de mida m ($m \leq n$), ordenades i sense repetició que es poden formar és

$$V_{n,m} = n \cdot (n - 1) \cdots (n - m + 1).$$

En el cas particular que $m = n$, $V_{n,n} = n \cdot (n - 1) \cdots 1 = n!$, **factorial** de n. Aquest nombre ens dóna les maneres d'ordenar n elements. Per al cas $n = 0$ s'adopta el conveni $0! = 1$.

Exemple 1.3 *Hem de connectar* 4 *cables diferents a* 3 *endolls diferents. Quantes possibilitats tenim?*

Sigui el conjunt dels 4 *cables,* $A = \{a,b,c,d\}$*. Una mostra la podem pensar com a acb, on la posició de la lletra indica un endoll determinat. Per exemple, si considerem la mostra acb volem indicar que el cable a és a l'endoll* 1*, el cable c a l'endoll* 2 *i el cable b a l'endoll* 3*. Si pensem en cab, és una mostra diferent de l'anterior, ja que ara és el cable c el que és a l'endoll* 1*. Hem de comptar el nombre de mostres de mida* 3*, ordenades i sense repetició que es poden formar en un conjunt de* 4 *elements. Així,* $V_{4,3} = 4 \cdot 3 \cdot 2 = 24$.

Aquestes són totes les mostres:

abc	*acb*	*bac*	*bca*	*cab*	*cba*
abd	*adb*	*bad*	*bda*	*dab*	*dba*
acd	*adc*	*cad*	*cda*	*dac*	*dca*
bcd	*bdc*	*cbd*	*cdb*	*dbc*	*dcb*

1.3. Mostres no ordenades sense repetició. Combinacions

Ens fixem en l'exemple 1.3 i el modifiquem lleugerament.

Exemple 1.4 *Hem de connectar* 4 *cables diferents a* 3 *endolls* ***iguals*** *(indistingibles). Quantes possibilitats tenim?*

Si ens fixem en les mostres que hem escrit a l'exemple 1.3, observem que en aquest nou exemple, totes les mostres que hi ha a la mateixa fila són la mateixa ja que l'unic que importa és el ***conjunt de tres*** *cables que hem triat per connectar. Per tant, hem de dividir el nombre de mostres que tenim en una fila per* 3!. *En tenim, doncs,* $\frac{V_{4,3}}{3!} = \frac{24}{6} = 4$.

$$abc \quad abd \quad acd \quad bcd$$

Anomenem combinacions de 4 *elements agafats de* 3 *en* 3*:* $C_{4,3} = \binom{4}{3} = \frac{V_{4,3}}{3!} = 4$.

En general, en un conjunt de n elements, el nombre de mostres de mida m ($m \leq n$), no ordenades i sense repetició és:

$$C_{n,m} = \binom{n}{m} = \frac{n \cdot (n-1) \cdots (n-m+1)}{m!} = \frac{n!}{m!(n-m)!}$$

Tal com hem comentat, el nombre combinatori $\binom{n}{m}$ ens dóna el nombre de subconjunts de m elements que podem formar d'un conjunt que en té n.

Proposició 1.1 *Propietats*

1) $\binom{n}{0} = 1$.

Per provar-ho pensem que el nombre de subconjunts de 0 *elements, que té un conjunt de n elements és 1, el conjunt buit.*

2) $\binom{n}{1} = n$.

*Aquesta igualtat és evident ja que el nombre de subconjunts d'*1 *element, que té un conjunt de n elements és n.*

3) $\binom{n}{m} = \binom{n}{n-m}$.

Per fer la prova fem el següent raonament: Podem formar el mateix nombre de subconjunts de m elements que de n - m elements, ja que cada cop que comptem un subconjunt de m elements, també estem comptant un subconjunt de n - m elements, $n = m + (n - m)$.

4) $\binom{n}{m} = \binom{n-1}{m} + \binom{n-1}{m-1}$.

Per fer la prova pensem en un conjunt A que té n elements. Si ens fixem en un element en concret, x, podem escriure el conjunt A com una unió $A = (A - \{x\}) \cup \{x\}$. *El nombre de subconjunts de m elements que podem formar en A serà la suma dels* *subconjunts que no tenen x més la dels que sí que tenen x. El nombre de subconjunts amb m elements on no hi és x és* $\binom{n-1}{m}$*, agafem els m elements del conjunt* $A - \{x\}$ *que té* $n - 1$ *elements. Els subconjunts de m elements que tenen x els formem afegint a l'element x,* $m - 1$ *elements del conjunt* $A - \{x\}$ *que té* $n - 1$ *elements, és a dir,* $\binom{n-1}{m-1}$.

1.4. Mostres no ordenades amb repetició

Exemple 1.5 *Tenim* 4 *boles iguals i les volem posar en* 3 *caixes diferents. Quantes possibilitats tenim?*

Si anomenem les caixes A, B i C, pensem la mostra AAAA com el cas en què les 4 boles es troben dins la caixa A, la mostra AABB com el cas en què hi ha dues boles a la caixa A i les altres dues a la caixa B. La mostra AABB és la mateixa que BABA, ja que les boles són iguals (indistingibles), per tant només l'hem de comptar un cop. Veiem que des d'aquest punt de vista (primer model), tenim mostres de mida 4 *(boles indistingibles) amb repetició i no ordenades. Ara bé, per calcular la quantitat de mostres d'aquest tipus, és millor pensar cada una d'aquestes mostres des d'un altre punt de vista. Pensem que hem d'omplir* 6 *espais amb* 4 *símbols del tipus • i* 2 *símbols del tipus |. La raó de que sigui així la veiem a continuació: Ens imaginem les tres caixes seguint aquest ordre, A|B|C, i ara, per simplificar, només cal que ens imaginem les separacions entre les caixes. Cada símbol | representa una separació entre dues caixes consecutives i per tant, només necessitem* 2 *separacions. De les* 6 *posicions que tenim, en triem dues per posar les separacions i a les altres posicions posem els símbols •. El que acabem d'explicar ho podem veure en algunes mostres*

primer model	*segon model* *omplim 6 espais* <u>1</u> <u>2</u> <u>3</u> <u>4</u> <u>5</u> <u>6</u>	*posicions de les separacions*
AAAA	• • • • | |	*{5,6}*
AAAB	• • • | • |	*{4,6}*
AABC	• • | • | •	*{3,5}*
CCCC	| | • • • •	*{1,2}*

*Cada mostra queda caracteritzada per la posició de les dues separacions entre les 6 que podem triar. Observem que el nombre de posicions per triar és la suma (*boles + separacions*)* = 4 + (3 - 1) = 6. *Donar dues posicions és el mateix que donar un subconjunt de* 2 *elements dins d'un conjunt de* 6 *elements. Per tant, el que estem comptant és el nombre de subconjunts de* 2 *elements que podem formar en un conjunt de* 6 *elements,* $\binom{3-1+4}{4} = \binom{6}{2} = \binom{6}{4} = 15$. *El fet que* $\binom{6}{2} = \binom{6}{4}$*, reflecteix el fet que és el mateix començar triant la posició de les separacions que la posició de les boles.*

AAAA	*BBBB*	*CCCC*	*AAAB*	*AAAC*	*BBBA*
BBBC	*CCCA*	*CCCB*	*AABB*	*AACC*	*BBCC*
AABC	*BBAC*	*CCAB*			

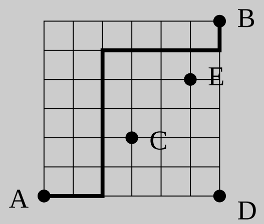
En general, en un conjunt de n elements, el nombre de mostres de mida m, no ordenades i amb repetició és:

$$CR_{n,m} = C_{n-1+m,m} = \binom{n-1+m}{m} = \binom{n-1+m}{n-1}. \qquad (1)$$

En diem, combinacions amb repetició de n elements agafats de m en m.

1.5. Altres exemples

1) Es vol connectar (cablejar) els punts A i B, de manera que el camí segueixi la quadrícula que marca el dibuix. Només és permès anar a la dreta (1) i a dalt (0). Al gràfic teniu representat un dels camins possibles, que vindria descrit per la seqüència 110000011110.



a) Calculeu el nombre de camins possibles entre A i B.

Volem conèixer el nombre de mostres del tipus 110000011110, on hem de mantenir el nombre de zeros. Cada 0 ocupa una posició que és determinada per un nombre del conjunt $A = \{1,2,\cdots,12\}$; així doncs, a la mostra 110000011110 li fem correspondre el subconjunt de 6 elements $\{3,4,5,6,7,12\}$. El nombre de subconjunts de 6 elements que podem formar amb els elements de A és $\binom{12}{6} = 924$.

b) Calculeu el nombre de camins possibles entre A i B que passin per C.

De A a C hi ha $\binom{5}{2}$ possibilitats i de C a B $\binom{7}{4}$ possibilitats. En total hi haurà $\binom{5}{2}\binom{7}{4} = 350$.

c) Calculeu el nombre de camins possibles entre A i B que passin per C i per E.

$$\binom{5}{2}\binom{4}{2}\binom{3}{2} = 180.$$

2) Considereu totes les solucions de l'equació $x_1+x_2+x_3+x_4 = 50$, on x_1,x_2,x_3,x_4 prenen valors naturals.

a) Quantes n'hi ha?

$$\binom{50+4-1}{4-1} = 23426.$$

b) Quantes en què, una i només una de les incògnites sigui 0?

$$4\binom{(50-3)+3-1}{3-1} = 4704.$$

c) Quantes hi ha, de manera que x_1,x_2,x_3,x_4 prenguin valors parells?

$$\binom{\frac{50}{2}+4-1}{4-1} = 3276.$$

d) Quantes hi ha, de manera que x_1,x_2,x_3,x_4 prenguin valors senars?

$$\binom{\frac{50-4}{2}+4-1}{4-1} = 2600.$$

3) Volem omplir la quadrícula següent amb 20 fitxes diferents.

	1	2	3	4	5
a					
b					
c					
d					
e					

a) De quantes maneres ho podem fer si podem posar totes les fitxes que vulguem dins d'un mateix quadre? $VR_{25,20} = 25^{20}$.

b) De quantes maneres, si cada quadre només pot tenir com a màxim una fitxa? $V_{25,20} = 25!/5!$.

c) De quantes maneres, si cada quadre només pot tenir com a màxim una fitxa i volem deixar una única fila buida? $5 \cdot V_{20,20} = 5 \cdot 20!$.

2. Espai de probabilitat

Molt sovint ens interessem per fenòmens on intervé l'atzar. Aquests fenòmens es caracteritzen perquè el resultat de les observacions varien d'una experiència a una altra.

La *probabilitat* que es realitzi un cert resultat en una experiència donada està relacionada amb la freqüència d'aquest resultat, si repetíssim *molts cops* l'experiència. Al llarg de la història s'han proposat diverses definicions matemàtiques de probabilitat (motivades principalment pels jocs d'atzar). Però no és fins a principi del segle XX que s'introdueix el model probabilístic de forma axiomàtica i així es formalitzen totes les anteriors idees.

Comencem el capítol donant les eines bàsiques necessàries per poder formalitzar el concepte de probabilitat.

2.1. Experiència aleatòria i successos. Operacions bàsiques i propietats

Definició 2.1 *Suposem que en repetir una determinada experiència en les mateixes condicions podem obtenir un conjunt de resultats diferents. Diem que l'experiència és aleatòria si és impossible de predir-ne el resultat. Per exemple:*

- *Observació del temps que triga un aparell nou abans d'espatllar-se.*
- *Observació del temps de durada de vida d'una persona anònima.*
- *Observació del nombre de peticions que arriben a un servidor no sobrecarregat.*

Exemple 2.1 *En llançar un dau podem obtenir un resultat qualsevol d'entre els següents* $\{1,2,3,4,5,6\}$*, però no podem predir-ne quin. Es tracta d'una experiència aleatòria. El conjunt format per tots els possibles resultats,* $\Omega = \{1,2,3,4,5,6\}$*, s'anomena espai mostral.*

Definició 2.2 *Anomenem espai mostral,* Ω*, el conjunt de resultats possibles d'una experiència aleatòria.*

Diem que "passa A" quan el resultat de l'experiment és un element de A.

Definició 2.3 *Donat un espai mostral,* Ω*, anomenem succés (esdeveniment),* A*, qualsevol subconjunt de l'espai mostral,* $A \subset \Omega$*. Un succés es diu elemental quan té un únic element.*

Exemple 2.2 *Continuant amb l'exemple 2.1, veiem alguns successos i algunes maneres de descriure'ls:*

$A = \{$ *nombre parell* $\} = \{2,4,6\}$, $B = \{$*nombre més gran que* $3\} = \{4,5,6\}$

En aquest exemple tenim 6 *successos elementals o successos que tenen un sol element:* $\{1\}, \{2\}, \{3\}, \{4\}, \{5\}$ *i* $\{6\}$.

Exemple 2.3 *Rebem un missatge binari (format amb elements de* $\{0,1\}$*), de llargada* 3 *(o de mida* 3*).*

- *L'espai mostral és* $\Omega = \{000,001,010,100,011,101,110,111\}$. *Com que té* 8 *elements, diem que el cardinal d'omega és* 8 *i escrivim* $|\ \Omega\ |= 8$.
- *Alguns successos:* $A = \{000,001,010\}$, $B = \{$*missatges amb un sol* $0\}$, $C = \{011,101\}$, $D = \{010,100,011,111\}$.

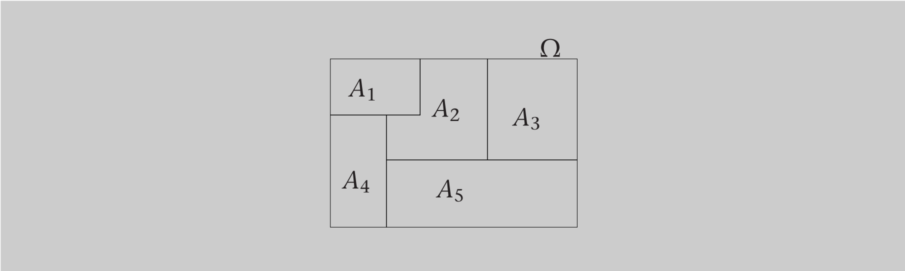
{000,001,010} = {010,000,001}. No importa l'ordre en què escrivim els elements d'un conjunt.

Definició 2.4 *Donats dos conjunts A i B,* $A,B \subset \Omega$:

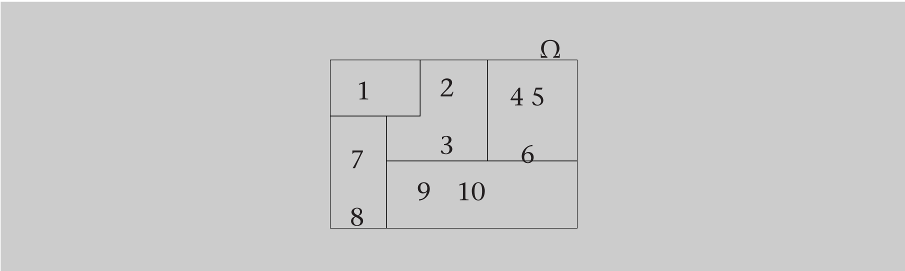
- *Complementari de A,* A^c, *és el conjunt que té per elements tots els de* Ω *que no són de A.*
- *A unió B,* $A \cup B$, *és el conjunt que té tots els elements de A i també els de B.*
- *A intersecció B,* $A \cap B$, *és el conjunt que té tots els elements de A que alhora també són de B.*
- $\varnothing$ *és el conjunt buit i* Ω *és el conjunt total.*
- *Diem que dos conjunts A i B són disjunts, quan* $A \cap B = \varnothing$, *és a dir, no tenen cap element en comú.*
- *Diem que els conjunts* $A_1, A_2, \cdots A_n$ *formen una partició de* Ω *quan els conjunts són disjunts dos a dos, i la unió de tots ells és el conjunt total. És a dir,*

$$A_i \bigcap A_j = \varnothing \quad per \quad i \neq j \quad i \quad \bigcup_{i=1}^{n} A_i = \Omega.$$

Exemple 2.4 *En aquest exemple representem una partició,* A_1, A_2, A_3, A_4, A_5, *d'un conjunt* $\Omega = \{1,2,3,4,5,6,7,8,9,10\}$, *amb* $A_1 = \{1\}$, $A_2 = \{2,3\}$, $A_3 = \{4,5,6\}$, $A_4 = \{7,8\}$ *i* $A_5 = \{9,10\}$. *La unió de tots ells és el total,* $A_1 \cup A_2 \cup A_3 \cup A_4 \cup A_5 = \Omega$, *i la intersecció entre dos qualssevol és buida. Vegem dues representacions,*



Els conjunts A_1, A_2, A_3, A_4 i A_5 formen una partició del conjunt total Ω



Els conjunts $\{1\}$, $\{2,3\}$, $\{4,5,6\}$, $\{7,8\}$ i $\{9,10\}$ formen una partició del conjunt total Ω

Vegem les anteriors definicions en termes probabilístics:

- El succés contrari de A, A^c, es realitza quan no ho fa A.
- El succés $A \cup B$ es realitza, si passa A, passa B o passa A i B alhora.
- El succés $A \cap B$ es realitza, si passa A i B alhora.
- $\varnothing$ és el succés impossible i Ω és el succés segur.
- Diem que A i B són dos successos incompatibles quan $A \cap B = \varnothing$.
- Diem que $A_1, A_2, \cdots A_n$, formen un sistema complet de successos, si formen una partició.

Els anteriors conceptes els hem resumit a la taula següent:

En termes de probabilitat	En termes de conjunts	Notació
Succés segur	Conjunt total	Ω
Succés impossible	Conjunt buit	$\varnothing$
Succés contrari	Conjunt complementari	A^c, també $\overline{A}$
A i B	Intersecció	$A \cap B$
A o B	Unió	$A \cup B$
Successos incompatibles	Conjunts disjunts	$A \cap B = \varnothing$
Sistema complet de successos	Partició de Ω	$A_i \cap A_j = \varnothing$ $\cup A_i = \Omega$

Considerant els conjunts de l'exemple 2.3, podem escriure:

$$
\begin{aligned}
A \text{ complementari, o complementari de } A, A^c &= \{000,001,010,011,101,110\} \\
A \text{ unió } B, A \cup B &= \{000,001,010,100\} \\
A \text{ unió } C, A \cup C &= \{000,001,010,011,101\} \\
A \text{ intersecció } C, A \cap C &= \varnothing
\end{aligned}
$$

2.2. Definició axiomàtica de probabilitat. Espai finit equiprobable

El resultat d'una experiència aleatòria no es pot preveure amb certitud. La teoria de la probabilitat dóna un *pes* a cada un dels possibles resultats, és a dir, un nombre que avalua la certesa que tenim que un resultat es doni.

Definició 2.5 *A partir d'una experiència aleatòria amb l'espai mostral Ω, considerem el conjunt format per tots els subconjunts de Ω, $\mathcal{P}(\Omega)$. Una probabilitat sobre Ω és una aplicació que a cada subconjunt $A \subset \Omega$ ($A \in \mathcal{P}(\Omega)$), li assigna un nombre real, $P(A)$, que verifica:*

1) $0 \leq P(A) \leq 1$.

2) $P(\Omega) = 1$.

3) *Si $A \cap B = \varnothing$ aleshores $P(A \cup B) = P(A) + P(B)$.*

Diem que tenim un *espai de probabilitat* quan tenim un conjunt Ω on hem definit una probabilitat.

Dels axiomes anteriors es dedueixen les següents propietats:

Proposició 2.1 *Propietats:*

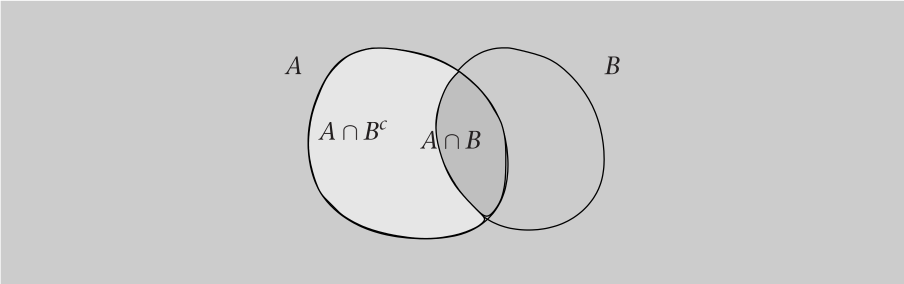
1) *La probabilitat del succés impossible és* 0.

2) *Donat un succés qualsevol A, es verifica $P(A^c) = 1 - P(A)$. Aquesta propietat és clara si pensem que A i A^c formen una partició del total, Ω.*

3) *Donats dos successos A i B, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.*

La prova de les tres primeres propietats és immediata. Fem la prova d'aquesta última propietat. Posem el conjunt A com a unió de dos conjunts disjunts,

$$A = (A \cap B) \cup (A \cap B^c)$$



Representació de $A \cap B^c$ i $A \cap B$

Com que tenim una unió de dos conjunts disjunts (ho podem veure a la figura), la probabilitat és suma de probabilitats,

$$P(A) = P(A \cap B) + P(A \cap B^c) \quad (2)$$

De manera semblant, escrivim $A \cup B$ *com a unió de dos conjunts disjunts,*

$$A \cup B = (A \cap B^c) \cup B$$

llavors,

$$P(A \cup B) = P(A \cap B^c) + P(B) \quad (3)$$

de les equacions 2 i 3 obtenim $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Definició 2.6 *Espai finit equiprobable*

En un espai finit equiprobable que té per espai mostral $\Omega = \{a_1, a_2, \cdots a_n\}$*, cada un dels successos elementals té la mateixa probabilitat. Així,*

$$P(\{a_1\}) = P(\{a_2\}) = \cdots = P(\{a_n\}) = p$$

i com que s'ha de verificar que

$$P(\{a_1\}) + P(\{a_2\}) + \cdots + P(\{a_n\}) = np = 1$$

es té que la probabilitat de cada succés elemental és $P(\{a_i\}) = \dfrac{1}{n}$.

Proposició 2.2 *Llei de Laplace*

En un espai equiprobable, la probabilitat d'un succés A és el quocient entre el nombre d'elements de A i el nombre d'elements de l'espai mostral. S'acostuma a dir

$$P(A) = \frac{\textit{nombre de casos favorables}}{\textit{nombre de casos possibles}}.$$

Exemple 2.5 *Considerem l'experiència de llençar una moneda tres cops seguits. L'espai mostral és* $\Omega = \{\bigcirc\bigcirc\bigcirc, \bigcirc\bigcirc+, +\bigcirc\bigcirc, \bigcirc+\bigcirc, ++\bigcirc, +\bigcirc+, \bigcirc++, +++\}$. *Si la moneda és perfecta es tracta d'un espai* ***equiprobable***. *Siguin els següents successos:* $A = \{$*han sortit dues cares*$\}$, $B = \{$*no ha sortit cap cara*$\}$, $C = \{$*ha sortit una creu*$\}$, $D = \{$*almenys ha sortit una creu*$\}$. *Calculem algunes probabilitats. El fet que l'espai sigui equiprobable ens permet aplicar la llei de Laplace. En cada cas cal comptar el nombre d'elements que té el conjunt i dividir per* 8.

Probabilitat que surtin dues cares, $P(A) = \frac{3}{8}$.
Probabilitat que no surti cap cara, $P(B) = \frac{1}{8}$.
Probabilitat que surti una creu, $P(C) = \frac{3}{8}$.
Probabilitat que almenys surti una creu, $P(A \cup D) = P(D) = \frac{7}{8}$.
Probabilitat que surtin dues cares i alhora cap cara (és impossible), $P(A \cap B) = 0$.
Probabilitat que surtin dues cares o bé almenys una creu, $P(A \cup D) = P(A) = \frac{3}{8}$.
Probabilitat que surtin dues cares o cap cara, $P(A \cup B) = \frac{4}{8}$.
Probabilitat que no surti cap cara i almenys una creu, $P(B \cap D) = P(B) = \frac{1}{8}$.

2.3. Probabilitat condicionada. Successos independents

Parlem de probabilitat condicionada quan ja s'ha realitzat l'experiència i ens donen una pista sobre el resultat obtingut. Vegem un exemple.

Exemple 2.6 *Considerem el mateix espai de probabilitat que a l'exemple 2.5. Es realitza l'experiència i ens donen la pista que almenys ha sortit una creu. Quina és la probabilitat que hagin sortit dues cares?*

És clar que ara l'espai total ha quedat reduït al conjunt

$$\{\bigcirc\bigcirc+, +\bigcirc\bigcirc, \bigcirc+\bigcirc, ++\bigcirc, +\bigcirc+, \bigcirc++, +++\}$$

Per tant, ara, la probabilitat que hagin sortit dues cares és $\frac{3}{7}$.

Ho escrivim com a $P(A/D) = \frac{3}{7}$ *i diem probabilitat de A condicionada a D.*

A continuació en donem la definició.

Definició 2.7 *Donats dos conjunts $A,B \subset \Omega$, amb $P(B) \neq 0$, definim la probabilitat del conjunt A condicionada a B com a:*

$$P(A/B) = \frac{P(A \cap B)}{P(B)} \tag{4}$$

Hem reduït l'espai mostral.

Es tracta de trobar una probabilitat, ***sabent*** *que s'ha realitzat B.*

Exemple 2.7 *Considerem el mateix espai de probabilitat que a l'exemple 2.5 i calculem algunes probabilitats condicionades.*

- *Sabent que ha sortit almenys una creu, probabilitat que hagin sortit dues cares. És a dir,* $P(A/D) = \frac{P(A \cap D)}{P(D)} = \frac{\frac{3}{8}}{\frac{7}{8}} = \frac{3}{7}$*. Ja ho havíem trobat a l'exemple anterior.*
- *Sabent que ha sortit almenys una creu, probabilitat que hagi sortit exactament una creu. És a dir,* $P(C/D) = \frac{P(C \cap D)}{P(D)} = \frac{3/8}{7/8} = \frac{3}{7}$
- *Sabent que ha sortit almenys una creu, probabilitat que no hagi sortit cap cara. És a dir,* $P(B/D) = \frac{P(B \cap D)}{P(D)} = \frac{1/8}{7/8} = \frac{1}{7}$
- *Sabent que ha sortit una creu, probabilitat que hagin sortit dues cares. De fet, els successos A i C són el mateix. Per tant,* $P(A/C) = 1$.
- *Sabent que no ha sortit cap cara, probabilitat que hagin sortit dues cares. És a dir,* $P(A/B) = \frac{P(A \cap B)}{P(B)} = 0$

Definició 2.8 *Sigui $A,B \subset \Omega$. El succés A és independent del succés B, quan la probabilitat de A no es modifica en conèixer alguna informació de la realització de B. És a dir,*

$$P(A/B) = P(A). \tag{5}$$

Volem veure que si A és independent del succés B llavors el succés B també és independent de A. De l'equació (4) tenim

$$P(A)P(B/A) = P(B \cap A) = P(A \cap B) = P(B)P(A/B)$$

i si tenim en compte (5) i subsituïm l'últim terme de la igualtat,

$$P(A)P(B/A) = P(B)P(A/B) = P(B)P(A) \tag{6}$$

ara simpliquem $P(A) \neq 0$ i ens queda

$$P(B/A) = P(B) \tag{7}$$

Per tant, si A és independent de B, B també és independent de A i es verifica

$$P(A \cap B) = P(A)P(B/A) = P(A)P(B) \tag{8}$$

i, per tant, si A i B són independents es verifica

$$P(A \cap B) = P(A)P(B). \tag{9}$$

Vegem un exemple numèric.

Exemple 2.8 *Siguin A i B dos successos de Ω i sabem que $P(A \cup B) = 0.52$, $P(A \cap B) = 0.08$ i $P(A) = 0.4$. Veurem que A i B són independents.*

Per a això veiem si es verifica la igualtat $P(A \cap B) = P(A)P(B)$. Aplicant la propietat $P(A \cup B) = P(A) + P(B) - P(A \cap B)$, tenim

$$0.52 = 0.4 + P(B) - 0.08 \implies P(B) = 0.2 \implies P(A)P(B) = 0.4\ 0.2 = 0.08 = P(A \cap B).$$

2.4. Teorema de la probabilitat total. Teorema de Bayes

Exemple 2.9 *Un aparell electrònic ha de treballar dins del rang de temperatures $[10^oC,40^oC]$. S'ha observat que quan la temperatura es troba a l'interval $T_1 = [10^oC,20^oC]$ té un comportament òptim el 75% de les vegades, quan treballa a temperatures de l'interval $T_2 = (20^oC,30^oC]$ un 55% de les vegades, i quan treballa a temperatures dins el rang $T_3 = (30^oC,40^oC]$ un 45% de les vegades. També coneixem la freqüència de cada un d'aquests rangs de temperatura. El 25% de les vegades la temperatura és dins T_1, el 60% dins T_2 i el 15% dins T_3. Ens preguntem quina és la probabilitat que, en un moment donat a una temperatura qualsevol dins el rang $[10^oC,40^oC]$, l'aparell tingui un comportament òptim.*

T_1, T_2 i T_3 formen una partició del conjunt de temperatures possible $[10^oC,40^oC]$ perquè $T_1 \cup T_2 \cup T_3 = [10^oC,40^oC]$ i $T_1 \cap T_2 = \emptyset$, $T_1 \cap T_3 = \emptyset$ i $T_2 \cap T_3 = \emptyset$. Si anomenem el succés O={Funcionament òptim }, podem escriure

$$O = (O \cap T_1) \cup (O \cap T_2) \cup (O \cap T_3)$$

i com que aquests conjunts són disjunts,

$$P(O) = P(O \cap T_1) + P(O \cap T_2) + P(O \cap T_3).$$

Però no coneixem el valor numèric de les probabilitats d'aquestes interseccions. De l'enunciat sabem que $P(O/T_1) = 0.75$, $P(O/T_2) = 0.55$, *i* $P(O/T_3) = 0.45$. *També coneixem* $P(T_1) = 0.25$, $P(T_2) = 0.60$ *i* $P(T_3) = 0.15$. *Podem deduir el valor d'aquestes probabilitats*

$$\begin{aligned} P(O \cap T_1) &= P(O/T_1)P(T_1) = 0.1875 \\ P(O \cap T_2) &= P(O/T_2)P(T_2) = 0.33 \\ P(O \cap T_3) &= P(O/T_3)P(T_3) = 0.0675 \end{aligned}$$

Així $P(O) = 0.1875 + 0.33 + 0.0675 = 0.585$.

En aquest exemple hem aplicat el teorema de la probabilitat total que enunciem a continuació:

Teorema 2.1 *Si* $A_1, A_2, \cdots A_n$ *és un sistema complet de successos de* Ω *i* $B \subset \Omega$, *podem escriure el succés B com a unió de parts disjuntes dos a dos*

$$B = (A_1 \cap B) \cup (A_2 \cap B) \cup \cdots (A_n \cap B).$$

Com que les parts són disjuntes, la probabilitat la trobem sumant les probabilitats

$$P(B) = P(A_1 \cap B) + P(A_2 \cap B) + \cdots + P(A_n \cap B)$$

i tenint en compte 4, obtenim el teorema de la probabilitat total

$$P(B) = P(B/A_1)P(A_1) + P(B/A_2)P(A_2) + \cdots + P(B/A_n)P(A_n).$$

Exemple 2.10 *Considerem el mateix enunciat que a l'exemple 2.9 i ens fem la següent pregunta: Sabent que el funcionament de l'aparell ha estat òptim, quina és la probabilitat que ens trobem en un rang de temperatures corresponent a* T_1*?*

És clar que el que ens demanen és $P(T_1/O)$*, que és precisament el contrari de les dades que ens donen, ja que el que coneixem és* $P(O/T_1)$, $P(O/T_2)$ *i* $P(O/T_3)$.

Com coneixem la relació, $P(T_1/O) = \dfrac{P(T_1 \cap O)}{P(O)}$*, i els valors numèrics ja els hem obtingut a l'exemple anterior, tenim que* $P(T_1/O) = \dfrac{0.1875}{0.585} = 0.3205$.

Acabem d'aplicar el que s'anomena teorema de Bayes. Vegem-ho de forma general.

Teorema 2.2 Teorema de Bayes

Si $A_1, A_2, \cdots A_n$ *és un sistema complet de successos de* Ω *i* $B \subset \Omega$,

$$P(B \cap A_i) = P(B/A_i)P(A_i) = P(A_i/B)P(B)$$

i obtenim la fórmula de Bayes:

$$P(A_i/B) = \frac{P(B/A_i)P(A_i)}{P(B/A_1)P(A_1) + P(B/A_2)P(A_2) + \cdots + P(B/A_n)P(A_n)}.$$

Exemple 2.11 *Hi ha tres empreses, A, B i C, que fabriquen la mateixa peça d'avió en les proporcions següents, respecte del total de peces fabricades:* 40%, 25% *i* 35%. *El* 10% *de peces que fabrica l'empresa A són defectuoses, mentre que aquest percentatge és de* 5% *per a l'empresa B ,i d'*1% *per a A. Dins de la producció total de les tres empreses, es tria una peça a l'atzar i s'observa que és defectuosa. Calculem la probabilitat que hagi estat fabricada per l'empresa A.*

Definim els següents successos:

$$D = \{\textit{la peça és defectuosa}\}$$
$$A = \{\textit{la peça ha estat fabricada per A}\}$$
$$B = \{\textit{la peça ha estat fabricada per B}\}$$
$$C = \{\textit{la peça ha estat fabricada per C}\}$$

A, B i C formen una partició, i coneixem $P(A) = 0.4$, $P(B) = 0.25$ *i* $P(C) = 0.35$. *L'enunciat també ens dóna les dades sobre la probabilitat que la peça sigui defectuosa segons on ha estat fabricada:* $P(D/A) = 0.1$, $P(D/B) = 0.05$ *i* $P(D/C) = 0.01$.

Segons el teorema de la probabilitat total,

$$P(D) = P(D/A)P(A) + P(D/B)P(B) + P(D/C)P(C) = 0.056$$

Amb el teorema de Bayes obtenim el que ens demanen,

$$P(A/D) = \frac{P(D/A)P(A)}{P(D)} = \frac{0.4\ 0.1}{0.056} = 0.714$$

2.5. Diagrames d'arbre

A l'hora d'aplicar els teoremes de la probabilitat total i Bayes, ens podem ajudar amb el que anomenem diagrames d'arbre. Vegem a la següent figura la manera de representar l'experiència de l'exemple 2.11.

Exemple 2.12 *Diagrama d'arbre.*

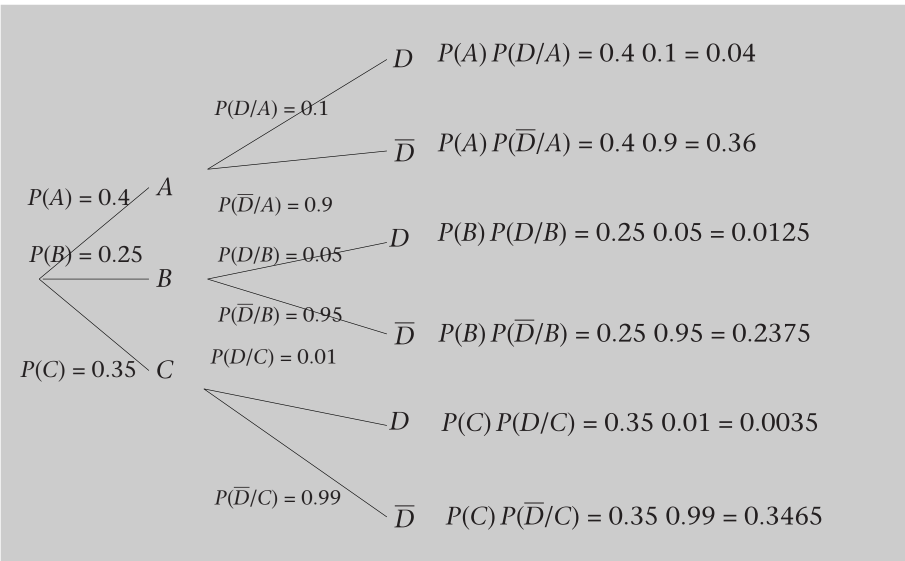


Diagrama d'arbre

Algunes consideracions sobre aquests diagrames:

- *Ens imaginem, temporalment, l'experiència, d'esquerra a dreta.*
- *Cada un dels camins, des de l'inici fins al final, representa una possibilitat de l'experiència.*
- *A la dreta del diagrama queden representades totes les possibilitats i per tant, la suma és 1.*
- *Cada un dels segments representa un "pas" de l'experiència.*
- *La probabilitat que indiquem en cada un d'aquests segments està condicionada a la part del camí ja realitzada.*
- *La suma de les probabilitats de tots els segments que parteixen d'un mateix punt és 1.*

Exemple 2.13 *S'envia una paraula de mida* 12 *formada amb elements del conjunt* {0,1}*, i cada un d'aquests elements és el que s'anomena* bit. *Una possible paraula podria ser* 111000111000. *Sabem que la probabilitat que un bit, independent dels altres, arribi erroni al receptor és de* 0.1. *Enviem una paraula i ens plantegem les següents questions:*

1) *Quina és la probabilitat que no arribi cap bit erroni?*

2) *Quina és la probabilitat que arribi un bit erroni?*

3) *Quina és la probabilitat que arribin dos bits erronis?*

4) *Quina és la probabilitat que arribin tres bits erronis?*

5) *Quina és la probabilitat que arribi, almenys, un bit erroni?*

6) *Quina és la probabilitat que arribi, com a mínim, un bit erroni?*

7) *Quina és la probabilitat que arribi, com a màxim, un bit erroni?*

Ens podem representar cada possibilitat com una seqüència de dotze lletres del conjunt $\{e,n\}$*, segons el bit arribi erroni o no arribi erroni. Per exemple:*

eeeeeeeeeeee	*tots els bits arriben erronis*
nneeeeeeeeee	*tots els bits arriben erronis menys els dos primers*
nennnnnnnnnn	*el segon bit arriba erroni i els altres no*
enennnnnnnnn	*el primer i tercer bit arriben erronis i els altres no*

Com que la probabilitat que un bit qualsevol arribi erroni és 0.1*, i aquesta probabilitat és independent del que els passa als altres bits, la probabilitat de cada una de les seqüències només depèn de la quantitat de* ***e*** *o* ***n***. *Ara podem respondre les anteriors preguntes.*

1) $P(nnnnnnnnnnnn) = 0.9^{12} = 0.28$.
2) *Hem de tenir en compte que el bit erroni pot ser a la primera posició, o a la segona,..., o a la dotzena posició, és a dir,*

$$ennnnnnnnnnn, nennnnnnnnnn, \ldots, nnnnnnnnnnne.$$

Com que la probabiltat de cada un d'aquests casos és $0.1\ 0.9^{11}$*, llavors*

$$P(\text{arriba un bit erroni}) = 12\ 0.1\ 0.9^{11} = 0.375$$

3) *De la mateixa manera que hem fet abans, hem de calcular quantes paraules es poden formar amb dos bits erronis, per exemple eennnnnnnnnn, enennnnnnnnn,... El nombre de paraules d'aquest tipus és* $\binom{12}{2}$. *Així,*

$$P(\text{arriben dos bits erronis}) = \binom{12}{2} 0.1^2 0.9^{10} = 0.23$$

4) *Amb un raonament semblant al cas anterior, obtenim ara*

$$P(\textit{arriben tres bits erronis}) = \binom{12}{3} 0.1^3 0.9^9 = 0.085$$

5) *Els casos que tenen* ***almenys*** *un bit erroni són els que tenen un bit erroni més els casos que tenen dos bits erronis, etc. És a dir, són tots els casos menys el cas on no hi ha cap bit erroni. Com que la probabilitat de tots els casos és 1, tenim* $P(\textit{arriba almenys un bit erroni}) = 1 - 0.9^{12} = 0.717$.

6) *Ens demanen el mateix que en el cas anterior, així,*

$$P(\textit{arriba com a mínim un bit erroni}) = 1 - 0.9^{12}$$

7) *En aquest cas només hem de comptar els casos que no tenen cap bit erroni més els casos on hi ha un bit erroni,*

$$P(\textit{arriba com a màxim un bit erroni}) = 0.9^{12} + 12 0.1 0.9^{11} = 0.65$$

3. Variables aleatòries

Molt sovint és necessari relacionar el resultat d'una experiència amb un nombre. En aquest capítol introduïm el concepte de variable aleatòria: variable aleatòria discreta i variable aleatòria contínua. Veurem com estudiar-les i en mostrarem els casos que apareixen més habitualment. Treballarem, en particular, amb les distribucions més importants que estan relacionades amb les telecomunicacions. Definirem els conceptes de valor mitjà, $E(X)$ i variància, $Var(X)$, d'una variable aleatòria X.

A partir d'una experiència aleatòria vam definir l'espai mostral, Ω, com el conjunt de tots els possibles resultats associats a aquesta experiència. Una variable aleatòria, X, assigna un nombre a cada un d'aquests resultats. Vegem alguns exemples:

Exemple 3.1 *Considerem l'experiència de llençar una moneda. L'espai mostral és* $\Omega = \{\text{cara,creu}\}$. *Podem assignar a cada un d'aquests resultats els valors* 0 *o* 1, *segons el resultat de l'experiència sigui* cara *o* creu. *Escrivim, doncs, que* $X(\text{cara}) = 0$ *i* $X(\text{creu}) = 1$. *La variable* X *pot prendre els valors* $\{0,1\}$.

X pren els valors $\{0,1\}$. És una variable aleatòria discreta.

Exemple 3.2 *Suposem que un aparell elèctric emet un senyal aleatori en mV cada segon i dins l'interval* $[0,2]$. *En aquest cas l'espai mostral està format per valors numèrics. Podem definir la variable aleatòria com l'aplicació identitat. A cada resultat de l'experiència li assigna el mateix valor. La variable* X *pot prendre un valor qualsevol de l'interval* $[0,2]$.

X pren els valors a $[0,2]$. És una variable aleatòria contínua.

Definició 3.1 *Una variable aleatòria, X, és una funció que assigna un nombre a cada element de l'espai mostral.*

Als exemples 3.1 i 3.2 veiem la diferència entre una variable aleatòria discreta i una contínua. Una **variable aleatòria discreta** (v. a. d.) pren valors d'un conjunt finit $\{a_1,a_1,\cdots,a_n\}$ o bé numerable infinit $\{a_1,a_2,\cdots\}$. Una **variable aleatòria contínua** (v. a. c.) pot prendre valors a tot $\mathbb{R}$. Això fa que el tractament matemàtic sigui molt diferent en els dos casos.

Observació

Escrivim $X = 0$ quan volem indicar que X pren el valor 0. A l'exemple 3.1

$$\begin{aligned} P(X=0) &= \frac{1}{2} \\ P(X=1) &= \frac{1}{2} \\ P(X\leq 1) &= 1 \end{aligned}$$

3.1. Variable aleatòria discreta

De forma natural la probabilitat que tenim definida a l'espai Ω es trasllada als valors que pren X.

A l'exemple 3.1 escrivim $P(X = 0) = P(\text{cara}) = \frac{1}{2}$ i $P(X = 1) = P(\text{creu}) = \frac{1}{2}$.

Definició 3.2 *S'anomena funció de probabilitat el conjunt de valors* $P(X = a_i)$.

Definició 3.3 *Diem que una v. a. d., que pren els valors* $\{a_1,a_2,\cdots,a_n\}$*, és* ***uniforme****, quan* $P(X = a_i) = \frac{1}{n} \quad \forall i \in \{1,2,\cdots n\}$.

A continuació veiem les distribucions discretes més importants.

3.1.1. Distribucions més importants

Distribució de Bernoulli: $B(p)$

Partim d'una experiència aleatòria i distingim entre els resultats $\Omega = \{A,\overline{A}\}$. El resultat A s'anomena èxit i definim $X(A) = 1$ i $X(\overline{A}) = 0$. La variable aleatòria pren només els dos valors $\{0,1\}$. Només cal donar la probabilitat assignada a un d'aquests valors i la distribució queda definida completament. Si $P(A) = p$ llavors $P(X = 1) = p$ i $P(X = 0) = 1 - p$.

Diem que X és una variable aleatòria de Bernoulli amb probabilitat d'èxit p i escrivim $X \sim B(p)$.

La variable de l'exemple 3.1 segueix una distribució $B(\frac{1}{2})$.

Exemple 3.3 *En comunicacions binàries X pot indicar l'error en la transmissió d'un bit. L'espai mostral de l'experiència és determinat per* $\Omega = \{error, no\ error\}$ *i la variable aleatòria pren els valors* $X = 1$ *si hi ha error i* $X = 0$ *si no n'hi ha.* $P(X = 1) = P(error)$ *i* $P(X = 0) = P(no\ error)$.

Exemple 3.4 *En llençar un dau ens fixem en la màxima puntuació possible i definim* $A = \{surt\ un\ 6\}$*, llavors* $\overline{A} = \{no\ surt\ un\ 6\}$*. Definim* $X(A) = 1$ *i* $X(\overline{A}) = 0$*. La variable X segueix una distribució* $B(\frac{1}{6})$.

Exemple 3.4

A l'exemple 3.4:

$$\begin{aligned} P(X = 0) &= \frac{5}{6} \\ P(X = 1) &= \frac{1}{6} \\ P(X \leq 0) &= \frac{5}{6} \\ P(X \leq 1) &= 1 \end{aligned}$$

Distribució Binomial: $Bin(n,p)$

Exemple 3.5 *Una persona, emissor, ha d'enviar un missatge de 10 elements, triats del conjunt* $\{0,1\}$ *i ordenats. Un missatge d'aquest tipus podria ser la paraula* 0011111101 *(formada amb 10 bits). Suposem que cada cop que la persona tria un bit per formar la paraula, la probabilitat que sigui un* 0 *és* 0.1 *i per tant, la que sigui un* 1 *és* 0.9*. Amb aquesta idea ens vénen al cap tota una sèrie de preguntes, com per exemple:*

1) *Quina és la probabilitat que l'emissor envïi exactament la paraula* 0011111101*?*

2) *Quina és la probabilitat que l'emissor envïi una paraula que tingui exactament tres zeros?*

3) *Quina és la probabilitat que l'emissor enviï una paraula que tingui exactament n zeros? (Busquem una expressió simbòlica de (3.5)).*

4) *Quina és la probabilitat que l'emissor enviï una paraula que tingui com a màxim tres zeros?*

Les anteriors qüestions les podem resoldre aplicant el que heu après al tema anterior. S'obtenen els resultats següents:

1) $(0.1)^3(0.9)^7$

2) *L'anterior resultat ens dóna la probabilitat per a una paraula que té* 3 *zeros i* 7 *uns. Vam veure al capítol anterior que el nombre de paraules que es poden formar amb* 3 *zeros i* 7 *uns és* $\binom{10}{3}$. *Llavors la resposta és* $\binom{10}{3}(0.1)^3(0.9)^7$.

3) *És clar que una paraula de mida* 10 *pot tenir entre* 0 *i* 10 *zeros. Així,* $0 \leq n \leq 10$. *Fent el mateix raonament que a l'apartat anterior tenim* $\binom{10}{n}(0.1)^n(0.9)^{10-n}$.

4) *Ara hem de tenir en compte aquells casos on el nombre de zeros sigui menor o igual a* 3. *De l'expressió anterior sumem els casos on n pren els valors* 0, 1, 2 *i* 3. *Obtenim:*

$$\begin{aligned}\sum_{k=0}^{3}\binom{10}{k}(0.1)^k.(0.9)^{10-k} &= \binom{10}{0}(0.1)^0(0.9)^{10-0}+\binom{10}{1}(0.1)^1(0.9)^{10-1}\\ &+ \binom{10}{2}(0.1)^2(0.9)^{10-2}+\binom{10}{3}(0.1)^3(0.9)^{10-3}\end{aligned}$$

A l'anterior exemple, si a la variable X li assignem el nombre de zeros que té cada paraula, els anteriors apartats els podem escriure utilitzant la X:

2) $P(X = 3) = \binom{10}{3}(0.1)^3(0.9)^7$

3) $P(X = n) = \binom{10}{n}(0.1)^n(0.9)^{10-n}$

4) $P(X \leq n) = \sum_{k=0}^{n}\binom{10}{k}(0.1)^k(0.9)^{10-k}$.

La **distribució binomial**, $Bin(n,p)$, es dóna quan repetim n cops i de forma independent una experiència $B(p)$. A cada resultat, la v. a. X li assigna el nombre d'èxits que han sortit. Així X pren els valors $\{0,1,2,\cdots,n\}$. La distribució de probabilitats (probabilitat que té cada un dels valors que pren la variable X) és determinada per:

$$P(X = k) = \binom{n}{k}p^k(1-p)^{n-k} \quad \text{amb} \quad k \in \{0,1,2,\cdots,n\}$$

Diem que X és una variable aleatòria binomial i escrivim $X \sim Bin(n,p)$, on n és el nombre de cops que repetim l'experiència de Bernoulli $B(p)$, de probabilitat d'èxit p.

A l'exemple 3.5, X segueix una distribució $Bin(10,0.1)$.

Exemple 3.6 *Enviem una paraula de n bits on cada bit pot portar error o no, independentment dels altres. La variable* $X \sim Bin(n,p)$ *pren el valor del nombre de bits erronis que hi ha a la paraula i per tant, els valors possibles són* $\{0,1,2,\cdots,n\}$. *p és la probabilitat que un bit sigui erroni.*

Distribució geomètrica: *Geom*(*p*)

Exemple 3.7 *Per enviar missatges per Internet, els missatges es divideixen en paquets i després s'envien per la xarxa. Si la xarxa està congestionada els paquets es poden perdre. Suposem que en una xarxa molt congestionada la probabilitat de perdre un paquet és* 0.8. *Això significa que el paquet no es perd en una transmissió, amb una probabilitat de* 0.2. *El paquet es transmet repetidament fins que el receptor el rep. Ens fem les següents preguntes:*

1) *Quina és la probabilitat que el paquet hagi de ser enviat almenys tres cops?*

2) *Quina és la probabilitat que haguem d'enviar el paquet com a màxim* 5 *cops perquè el receptor el rebi?*

Si la v. a. X compta el nombre de vegades que cal enviar un paquet, pren valors del conjunt $\{1,2,3,\cdots\}$ *(X és una variable discreta infinita). Podem escriure:*

1) $P(X \geq 3) = 1 - P(X \leq 2) = 1 - (P(X = 1) + P(X = 2)) = 1 - (0.2 + 0.2\ 0.8) = 0.64$

2) $P(X \leq 5) = \sum_{k=1}^{5} 0.2(0.8)^{k-1} = 0.2\left(1 + 0.8 + (0.8)^2 + (0.8)^3 + (0.8)^4\right) = 0.67232$

La **distribució geomètrica**, *Geom*(p), es dóna quan repetim, de forma independent, una experiència $B(p)$, fins a obtenir el primer èxit. X compta el nombre de cops que cal fer l'experiència per obtenir el primer èxit. Per tant, X pren els valors $\{1,2,3,\cdots\}$ i la seva distribució de probabilitats és determinada per:

$$P(X = k) = (1 - p)^{k-1}p,$$

Diem que X és una variable aleatòria geomètrica amb probabilitat d'èxit p i escrivim $X \sim Geom(p)$.

A l'exemple 3.7, X segueix una distribució $Geom(0.2)$.

Distribució de Poisson: $Poiss(\alpha)$

La variable aleatòria de Poisson s'utilitza per modelar alguns fenòmens com els següents:

- El nombre d'accidents en un encreuament donat i per un interval de temps fixat.
- El nombre de trucades que arriben a una centraleta en un cert interval de temps.
- El nombre de peticions que arriben a un servidor en un cert interval de temps.
- El nombre d'electrons o forats que travessen una barrera de potencial.
- El nombre de defectes de fabricació d'un producte d'unes dimensions determinades.

El nombre mitjà d'èxits dins de l'interval de mesura considerat (temps, àrea, etc.) ens dóna el paràmetre de la distribució, α.

X pren els valors $\{0,1,2,3,\cdots\}$ i la distribució de probabilitats és determinada per:

$$P(X = k) = \frac{\alpha^k}{k!}e^{-\alpha}$$

Diem que X és una variable aleatòria de Poisson de paràmetre α, i escrivim $X \sim Poiss(\alpha)$.

Exemple 3.8 *Sabem que a un servidor hi arriben de mitjana 5 peticions per segon. Quina és la probabilitat que en un segon no hi arribi cap petició? Quina és la probabilitat que en un segon hi arribin una o més peticions?*

La dada que ens dóna l'enunciat és precisament el paràmetre de la distribució de Poisson, així $\alpha = 5$. *Ara ja podem donar les respostes:*

$$P(X = 0) = \frac{5^0}{0!}e^{-5} = e^{-5} = 0.0067$$

i

$$P(X \geq 1) = 1 - P(X = 0) = 1 - 0.0067 = 0.9933$$

3.1.2. Paràmetres: Valor mitjà i variància

Hi ha certs paràmetres que ens proporcionen una informació global del comportament de les variables.

Definició 3.4 *Sigui X una variable aleatòria discreta que pren els valors següents: $\{a_1,a_2,\cdots,a_n\}$ (per al cas que X prengui infinits valors es fa de manera semblant, però en lloc d'una suma finita tenim una sèrie numèrica).*

El valor mitjà o esperança de X:

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

El moment d'ordre 2:

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

El moment d'ordre n:

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

La variància:

$$Var(X) = E((X - E(X))^2) = \sum_{i=1}^{n} (a_i - E(X))^2 P(X = a_i) = E(X^2) - E(X)^2.$$

Desviació típica:

$$\sigma = \sqrt{Var(X)}$$

L'esperança $E(X)$ ens dóna el valor mitjà dels valors que pren X tenint en compte la probabilitat de cada un d'aquests valors. La variança $Var(X)$ ens dóna el valor mitjà de les diferències de cada valor respecte a $E(X)$ i al quadrat. Per obtenir una mesura de la dispersió amb unitats equivalents a les dels valors de la variable, fem $\sqrt{Var(X)}$.

Vegem-ho al següent exemple:

Exemple 3.9 *Siguin X_1 i X_2 dues variables aleatòries equiprobables que prenen els valors $\{4,5,6\}$ i $\{0,5,10\}$, respectivament. Aquests valors podrien ser les tres notes obtingudes en una determinada assignatura per dos alumnes diferents. Suposant que les tres notes tenen el mateix pes, calculem l'esperança i variància de cada una de les variables:*

$$
\begin{aligned}
E(X_1) &= 4\frac{1}{3} + 5\frac{1}{3} + 6\frac{1}{3} = 5 \\
E(X_2) &= 0\frac{1}{3} + 5\frac{1}{3} + 10\frac{1}{3} = 5 \\
Var(X_1) &= (4-5)^2\frac{1}{3} + (5-5)^2\frac{1}{3} + (6-5)^2\frac{1}{3} = \frac{2}{3} \quad i \quad \sigma_{X_1} = 0.82 \\
Var(X_2) &= (0-5)^2\frac{1}{3} + (5-5)^2\frac{1}{3} + (10-5)^2\frac{1}{3} = 16.67 \quad i \quad \sigma_{X_2} = 4.08
\end{aligned}
$$

El que podem dir és que els dos alumnes tenen la mateixa nota de mitjana $E(X_1) = E(X_2)$, però el segon alumne presenta més dispersió en les seves notes, $\sigma_{X_2} > \sigma_{X_1}$, és menys regular.

Vegem la propietat de linealitat per l'esperança, ja que ens pot facilitar molts càlculs. (La variança no presenta aquesta propietat.)

Proposició 3.1 *Si X és una v. a. d. llavors $E(aX + b) = aE(X) + b$, $\forall\ a,b \in \mathbb{R}$.*

Proposició 3.2 *Per cada una de les distribucions vistes, s'obtenen els següents valors de l'esperança i la variància:*

Distribucions de variables aleatòries discretes

$X \sim$	k	$P(X = k)$	$E(X)$	$Var(X)$
$B(p)$	$\{0,1\}$	$P(X=1) = p$ $P(X=0) = 1 - p$	p	$p(1 - p)$
$Bin(n,p)$	$\{0,1,2,\ldots,n\}$	$P(X=k) = \binom{n}{k}p^k(1-p)^{n-k}$	np	$np(1 - p)$
$Geom(p)$	$\{1,2,3,\ldots\}$	$P(X=k) = (1-p)^{k-1}p$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
$Poiss(\lambda)$	$\{0,1,2,\ldots\}$	$P(X=k) = \frac{\lambda^k}{k!}e^{-\lambda}$	λ	λ

Comprovem els resultats de la taula anterior que corresponen a $X \sim B(p)$.

El valor mitjà, $E(X) = 1p + 0(1 - p) = p$.

El moment d'ordre 2, $E(X^2) = 1^2p + 0^2(1 - p) = p$.

La variància, $Var(X) = p - p^2 = p(1 - p)$.

3.1.3. Funció de distribució

Una manera de donar el valor de les probabilitats acumulades és a partir de la funció de distribució.

Definició 3.5 *La funció de distribució d'una variable aleatòria X es defineix com a:*

$$F_X(x) = P(X \leq x\} \quad \forall x \in \mathbb{R} \tag{10}$$

i verifica les següents propietats.

1) $0 \leq F_X(x) \leq 1$.

2) $\lim_{x \to \infty} F_X(x) = 1$.

3) $\lim_{x \to -\infty} F_X(x) = 0$.

4) $F_X(x)$ és creixent, és a dir, si $a < b$ llavors $F_X(a) \leq F_X(b)$.

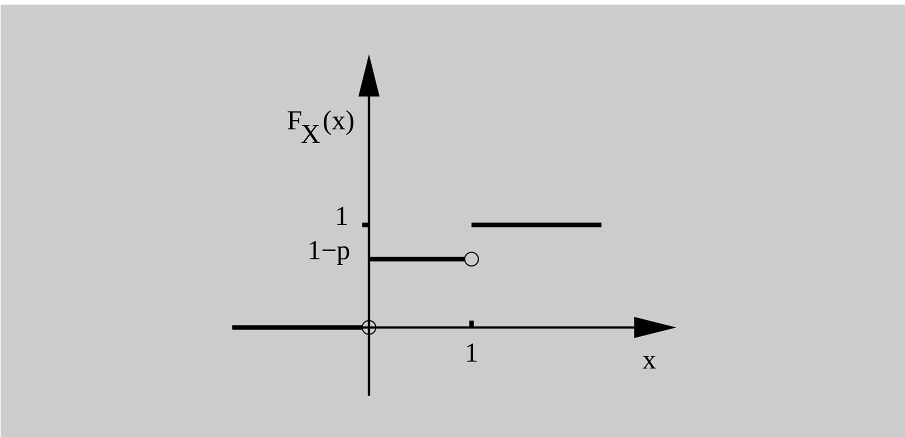
5) $P(a < X \leq b) = F_X(b) - F_X(a)$.

En els cas de variables aleatòries discretes, la funció de distribució és esglaonada. La funció experimenta un salt en cada nombre real que correspongui a un valor que pren X.

Vegem les anteriors propietats en alguns exemples:

Exemple 3.10 *Per al cas de $X \sim B(p)$. La funció de distribució $F_X(x)$ presenta una discontinuïtat de salt a* 0 *i a* 1.

$$F_X(x) = \begin{cases} 0 \quad si \quad x < 0 \\ 1 - p \quad si \quad 0 \leq x < 1 \\ 1 \quad si \quad 1 \leq x \end{cases}$$



Funció de distribució de $X \sim B(p)$

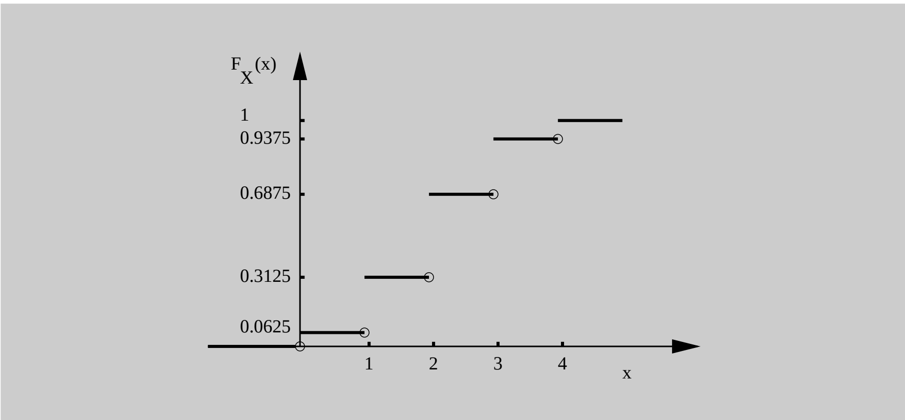
Exemple 3.11 *En el cas $X \sim Bin(4,1/2)$, $F_X(x)$ presenta una discontinuïtat de salt en els valors del conjunt $\{0,1,2,3,4\}$. Tenint en compte que $P(X = k) = \binom{4}{k}(\frac{1}{2})^k(\frac{1}{2})^{4\text{-}k}$ i*

$$F_X(x) = P(X \leq k) = \sum_{i=0}^{k} \binom{4}{i} (\frac{1}{2})^i (\frac{1}{2})^{4\text{-}i} = \sum_{i=0}^{k} \binom{4}{i} (\frac{1}{2})^4,$$

obtenim els valors més significatius de la funció de distribució:

$$\begin{aligned}
F_X(0) &= 0.0625 \\
F_X(1) &= 0.0625 + 0.250 = 0.3125 \\
F_X(2) &= 0.0625 + 0.250 + 0.375 = 0.6875 \\
F_X(3) &= 0.0625 + 0.2500 + 0.3750 + 0.2500 = 0.9375 \\
F_X(4) &= 0.0625 + 0.2500 + 0.3750 + 0.2500 + 0.0625 = 1
\end{aligned}$$

$$F_X(x) = \begin{cases} 0 \quad si \quad x < 0 \\ 0.0625 \quad si \quad 0 \leq x < 1 \\ 0.3125 \quad si \quad 1 \leq x < 2 \\ 0.6875 \quad si \quad 2 \leq x < 3 \\ 0.9375 \quad si \quad 3 \leq x < 4 \\ 1 \quad si \quad 4 \leq x \end{cases}$$



Funció de distribució de $X \sim Bin(4, \frac{1}{2})$

Les 4 primeres propietats de la funció de distribució, $F(X)$, s'observen fàcilment a la figura.

Els valors de les probabilitats acumulades els obtenim directament de la funció de distribució. Donem alguns valors com a exemple:

$$\begin{aligned} P(X \leq 0.5) &= F_X(0.5) = 0.0625 \\ P(X \leq 1.7) &= F_X(1.7) = 0.3125 \\ P(X \leq 2.4) &= F_X(2.4) = 0.6875 \\ P(1.7 < X \leq 2.4) &= F_X(2.4) - F_X(1.7) = 0.375 \end{aligned}$$

3.2. Variable aleatòria contínua

A l'exemple 3.2 hem vist que la variable aleatòria contínua X pot prendre un valor qualsevol de l'interval [0,2]. En aquest cas, si suposem que cap dels valors dins [0,2] té preferència, podríem trobar els següents resultats de forma intuïtiva:

1) Quina és la probabilitat que el senyal emès es trobi entre 0 i 1 mV? És a dir, $P(0 \leq X \leq 1)$? Tot ens fa pensar que és $\frac{1}{2}$ ja que estem jugant amb la meitat de possibilitats.

2) Quina és la probabilitat que el senyal emès es trobi entre 3 i 4 mV? És a dir, $P(3 \leq X \leq 4)$? Com que sabem que això no succeirà mai diem que és 0.

3) Quina és la probabilitat que el senyal emès sigui exactament d'1 mV? Aquest cas ens caracteritza les distribucions de v. a. c. Diem que $P(X = 1) = 0$. En una distribució de variable aleatòria contínua, la probabilitat en qualsevol punt x, és zero.

Si $X \sim$ **v. a. contínua,** $\forall x \in R$
$P(X = x) = 0$.

3.2.1. Funció de distribució i funció de densitat

La funció de distribució es defineix de la mateixa manera que per a una v. a. d.

Definició 3.6 *La funció de distribució d'una variable aleatòria X es defineix com a:*

$$F_X(x) = P(X \leq x\} \quad \forall x \in \mathbb{R} \tag{11}$$

i verifica les següents propietats:

1) $0 \leq F_X(x) \leq 1$.

2) $\lim_{x \to \infty} F_X(x) = 1$.

3) $\lim_{x \to -\infty} F_X(x) = 0$.

4) *$F_X(x)$ és creixent, és a dir, si $a < b$ llavors $F_X(a) \leq F_X(b)$.*

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

En aquest cas de v. a. c., $F_X(x)$ és contínua i derivable (llevat potser d'una quantitat finita de punts). Això ens permet definir la funció de densitat:

Definició 3.7 *Si X és una v. a. c. amb funció de distribució $F_X(x)$, la funció de densitat es defineix com a:*

$$f_X(x) = \frac{dF_X(x)}{dx} \quad \forall x \in \mathbb{R} \tag{12}$$

S'obtenen les següents relacions:

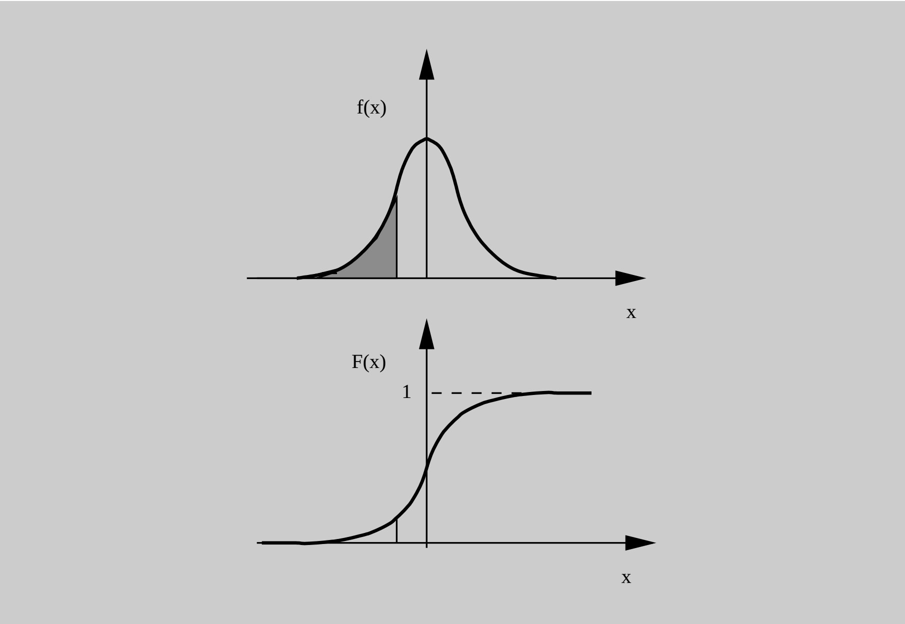
Proposició 3.3

1) $f_X(x) \geq 0$. *Això és clar si observem l'equació (12) i pensem que $F_X(x)$ és creixent.*

2) $P(a < X \leq b) = F_X(b) - F_X(a) = \int_a^b f_X(x)\,dx$. *És a dir, la probabilitat entre dos punts a i b l'obtenim integrant la funció de densitat entre aquests dos punts. (Àrea per sota de la corba de la funció de densitat.)*

Si $X \sim$ v. a. c. llavors es verifica $P(a \leq X \leq b) = P(a < X < b)$.

3) $F_X(x) = \int_{-\infty}^x f_X(t)\,dt$. *La probabilitat acumulada també la podem pensar com una àrea per sota de la funció de densitat $f(x)$.*



Funció de densitat $f(x)$ i funció de distribució $F(x)$

L'àrea per sota de la corba de $f(x)$ correspon a un punt de $F(x)$.

3.2.2. Distribucions més importants

Quan treballem amb v. a. c. les distribucions queden caracteritzades per la seva funció de densitat. Vegem a continuació les més importants.

Distribució uniforme: $X \sim U(a,b)$

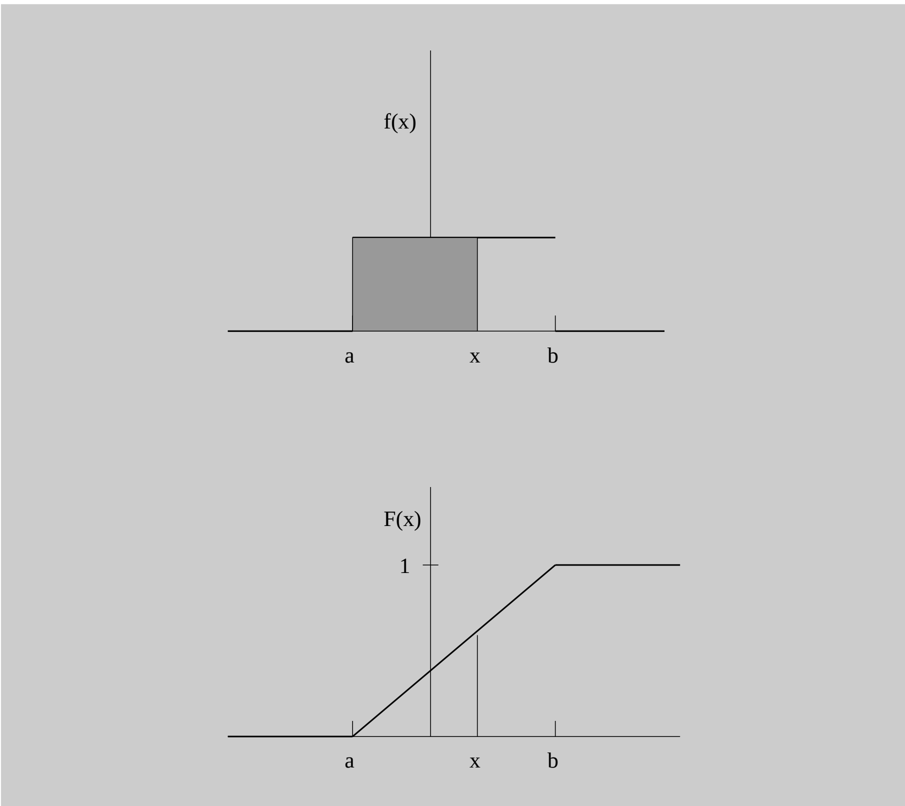
La variable X pot prendre un valor qualsevol de l'interval $[a,b]$ i de manera uniforme. Això ho indiquem amb la funció de densitat:

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

La funció de distribució serà, doncs:

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt = \begin{cases} 0 & \text{si} \quad x < a \\ \frac{1}{b\text{-}a}(x - a) & \text{si} \quad a \le x < b \\ 1 & \text{si} \quad b \le x \end{cases}$$

A la figura següent s'ha representat la funció de densitat i la funció de distribució. L'àrea indicada per sota de $f(x)$ correspon a un punt de $F(x)$.



Funció de densitat $f(x)$ i funció de distribució $F(x)$ de la variable aleatòria $X \sim U(a,b)$

L'exemple 3.2 segueix una distribució $X \sim U(0{,}2)$. Així, la probabilitat que el senyal emès es trobi entre 0 i 1 mV ens la dóna l'àrea per sota de la corba de la funció de densitat, que en aquest cas correspon a l'àrea d'un rectangle de base 1 i altura $\frac{1}{2}$. És a dir, $P(0 \leq X \leq 1) = \frac{1}{2}$.

Exemple 3.12 *Triem a l'atzar un nombre, X, a l'interval* (0,5). *La funció de densitat és:*

$$f(x) = \begin{cases} \frac{1}{5} & \text{si} \quad x \in (0{,}5) \\ 0 & \text{altrament} \end{cases}$$

Calculem algunes probabilitats.

1) *Probabilitat que el nombre sigui menor que 3,* $P(X < 3) = \int_0^3 \frac{1}{5}\,dx = \frac{3}{5}$.

2) *Sabent que el nombre és superior a 2, probabilitat que sigui menor que 3,*

$$P(X < 3/X > 2) = \frac{P(2 < X < 3)}{P(X > 2)} = \frac{1/5}{3/5} = \frac{1}{3}$$

Distribució exponencial: *Exp*(λ)

La distribució exponencial se sol utilitzar per modelar experiències on intervé un temps d'espera, com:

- Temps d'espera en una consulta sense cita prèvia.
- Temps d'espera en un servidor.
- La vida d'un component electrònic.

La distribució de Poisson està molt relacionada amb la distribució exponencial. Si un procés és de Poisson (succés aleatori en el temps), la variable temps que passa fins que es realitza el primer succés, és exponencial. Cal destacar que el paràmetre de la variable de Poisson val $\alpha = \lambda t$.

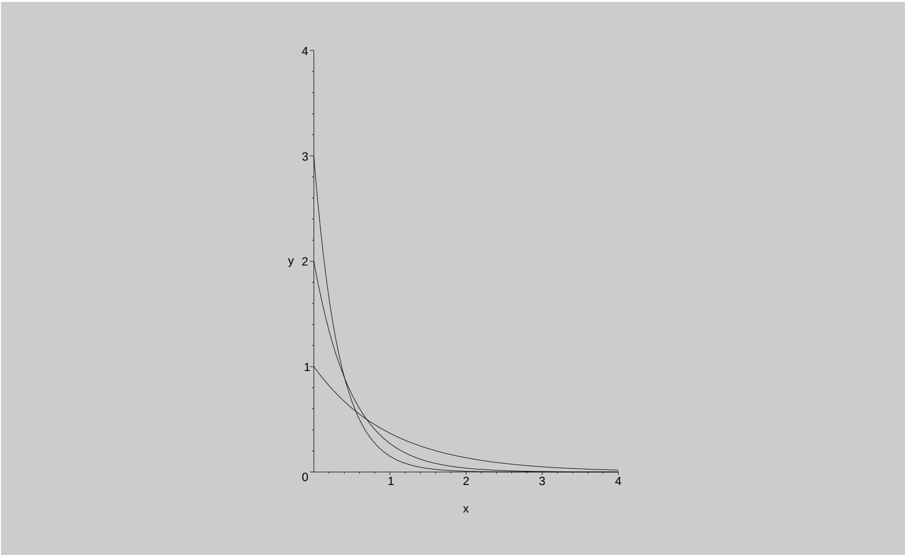
La distribució exponencial té per funció de densitat:

$$f(x) = \begin{cases} e^{-\lambda x} & \text{si} \quad x \geq 0 \\ 0 & \text{altrament} \end{cases}$$

Obtenim la funció de distribució integrant. Així,

$$F(x) = \begin{cases} 0 & \text{si} \quad x < 0 \\ \lambda e^{-\lambda x} & \text{si} \quad x \geq 0 \end{cases}$$

A la figura veiem la representació de la funció de densitat, per a tres valors diferents de λ. (No s'ha representat l'eix negatiu d'abscisses on la funció és 0.)



Funcions densitat de $X \sim Exp(1)$, $X \sim Exp(2)$ i $X \sim Exp(3)$

Vegem un exemple d'aplicació.

Exemple 3.13 *Suposem que el temps, en hores, que es necessita per arreglar un cert tipus d'avaria telefònica és una variable aleatòria, T, que segueix una llei exponencial de paràmetre $\lambda = 0.5$. En aquest cas tenim $f(t) = 0.5e^{-0.5t}$ i $F(t) = 1 - e^{-0.5t}$. Calculem algunes probabilitats.*

1) *Probabilitat que el temps de reparació passi de les 2 hores. És a dir,* $P(T > 2) = 1 - P(T < 2) = 1 - F(2) = e^{-1} = 0.368$.

2) *Sabent que el temps de reparació ja ha sobrepassat les 9 hores, quina és la probabilitat que la reparació trigui almenys 10 hores?*

En aquest cas es tracta de trobar una probabilitat condicionada, escrivim

$$P(T > 10/T > 9) = \frac{P((T > 10) \cap (T > 9))}{P(T > 9)} = \frac{P((T > 10)}{P(T > 9)} = \frac{1 - F(10)}{1 - F(9)} = e^{-0.5 \cdot 1} = 0.606.$$

Distribució normal o de Gauss: $N(m,\sigma)$

És una de les lleis de probabilitat més utilitzada. Depèn de dos paràmetres, el valor mitjà m (paràmetre de posició), i la desviació típica σ, paràmetre que ens mesura la dispersió de la variable aleatòria respecte de m (vegem la taula de la proposició 3.4).

Aquesta llei s'aplica a una gran quantitat de fenòmens de física i economia. A més, és la forma límit d'algunes distribucions discretes quan s'augmenta indefinidament el nombre de repeticions d'un experiment. Moltes variables aleatòries com pesos, alçades, talles, consums de gas, etc. segueixen una distribució normal perquè cadascuna d'elles és la suma d'un gran nombre de variables aleatòries independents. Així, l'alçada d'una persona és la suma de molts factors, hereditaris, alimentació, tipus de vida, etc.

Els errors, anomenats aleatoris, que es presenten en observacions atronòmiques, pesades d'una balança, etc. i, en general, en la majoria de mesures amb algun aparell, són la suma d'un gran nombre d'errors elementals independents com corrents d'aire, vibracions, error d'apreciació, etc. Per això, els errors aleatoris segueixen una distribució normal.

La distribució normal té per funció de densitat:

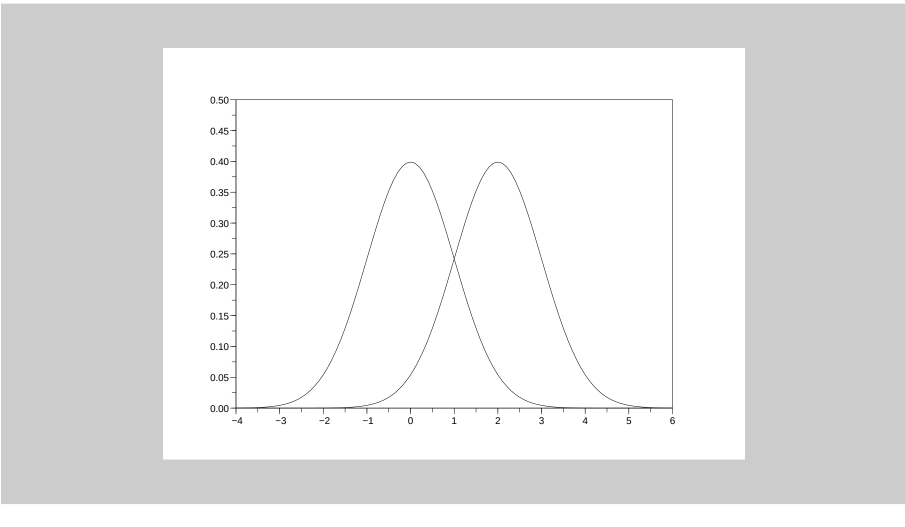
$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}} \quad \forall x \in \mathbb{R}.$$

En aquest cas la funció de distribució no es pot trobar integrant analíticament de forma senzilla com ho hem fet abans.

Vegem alguns gràfics de la funció de densitat en variar els paràmetres m i σ.

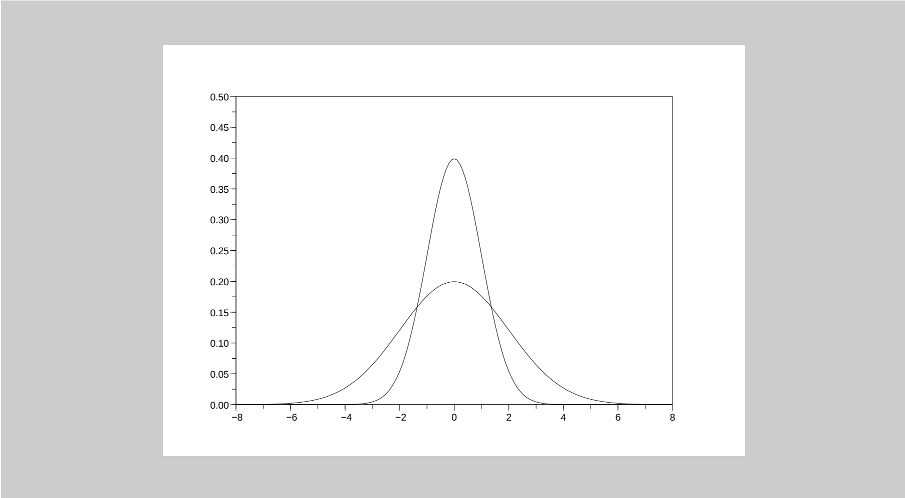
Per calcular probabilitats utilitzarem la calculadora scilab. També es solen usar taules.

A la següent figura hem representat $N(0,1)$ i $N(2,1)$. Observem que la forma de la funció no ha variat. $N(2,1)$ està traslladada dues unitats a la dreta respecte de $N(0,1)$.



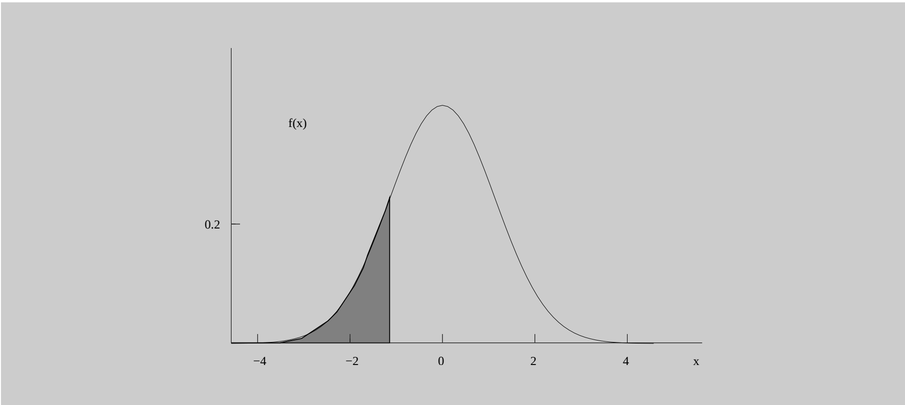
Funcions de densitat de $X \sim N(0,1)$ i $X \sim N(2,1)$

A la següent figura fixem el valor a $m = 0$ i modifiquem σ. Observem que per a un valor menor de σ, hi ha menys dispersió. Per a $\sigma = 2$ la funció de densitat és més punxeguda. (Observem que encara que canviï la forma de la funció, l'àrea total per sota de la corba és 1, la probabilitat total.)



Funcions de densitat de $X \sim N(0,1)$ i $X \sim N(0,2)$

Ja que no podem integrar analíticament $f(x)$, per trobar probabilitats cal utilitzar taules o bé algun programa de càlcul. En el nostre cas utilitzarem la calculadora scilab. Vegem a continuació algun exemple.



Funció de distribució de $X \sim N(0,1)$

Si el que volem calcular és l'àrea indicada a la figura, $P(X < n)$, per a un cert valor de n, i $N(0,1)$, escrivim a la calculadora:

```
y = integrate('1/sqrt(2*%pi)*exp(-t^2/2)','t',-10,n)
```

En el cas particular $n = 0$, obtenim una àrea de 0.5. Proveu en aquest cas la calculadora. Us sortirà:

```
y = integrate('1/sqrt(2*%pi)*exp(-t^2/2)','t',-10,0)

y = 0.5
```

Fixeu-vos que s'ha posat com a límit inferior -10. Proveu altres valors, i veureu que no cal considerar valors molt negatius ja que la funció és pràcticament zero.

3.2.3. Paràmetres: Valor mitjà (esperança) i variància

De forma semblant a com ho hem fet en v. a. d. definim els paràmetres més importants. En treballar amb variable contínua, cal transformar els sumatoris amb integrals.

Definició 3.8 *Sigui X una variable aleatòria contínua,*

El valor mitjà o esperança de X:

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

El moment d'ordre 2:

$$E(X^2) = \int_{-\infty}^{+\infty} x^2 f(x)\, dx.$$

El moment d'ordre n:

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

La variància:

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

Desviació típica:

$$\sigma = \sqrt{Var(X)}$$

Proposició 3.4 *Per cada una de les distribucións vistes, s'obtenen els valors de l'esperança i la variància:*

Distribucions de variables aleatòries contínues

$X \sim$	*Funció de densitat*	$E(X)$	$Var(X)$
$U(a,b)$	$f(x) = \begin{cases} \frac{1}{b-a} & \text{si } x \in (a,b) \\ 0 & \text{altrament} \end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
$N(m,\sigma)$	$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}} \ \forall x \in \mathbb{R}$	m	σ^2
$Exp(\lambda)$	$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x \geq 0 \\ 0 & \text{altrament} \end{cases}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$

Comprovem algun resultat per al cas de $X \sim Exp(\lambda)$

$$\begin{aligned}
E(X) &= \int_{-\infty}^{+\infty} xf(x)\,dx = \int_{0}^{+\infty} x\lambda e^{-\lambda x}\,dx = \left\{ \begin{array}{c} u = x \quad \text{i} \quad du = dx \\ dv = \lambda e^{-\lambda x} \quad \text{i} \quad v = -e^{-\lambda x} \end{array} \right\} = \\
&= \left\{-xe^{-\lambda x}\right\}_{0}^{+\infty} + \int_{0}^{+\infty} e^{-\lambda x}\,dx = \lim_{x \to +\infty}(-xe^{-\lambda x}) - \lim_{x \to 0}(-xe^{-\lambda x}) + \frac{1}{\lambda} = \\
&= 0 + 0 + \frac{1}{\lambda}.
\end{aligned}$$

3.3. Teorema central del límit. Aplicació

Tal com hem dit abans, la distribució normal és la forma límit d'algunes distribucions discretes, quan s'augmenta indefinidament el nombre de repeticions d'un experiment. Enunciem el teorema central del límit que reflecteix aquest fet:

Sigui (X_n) amb $n \geq 1$ una successió de variables aleatòries independents, que segueixen la mateixa llei de probabilitat, amb una esperança m i variància σ^2. Considerem la nova variable aleatòria definida per:

$$Y_n = \frac{\frac{1}{n}\sum_{i=1}^{n} X_i - m}{\frac{\sigma}{\sqrt{n}}} \tag{13}$$

Es té que la variable Y_n convergeix cap a la distribució $N(0,1)$ quan n tendeix a infinit.

Aquest resultat teòric ens permet aproximar una llei binomial per una llei normal.

3.3.1. Aproximació de llei binomial per la normal

Si estem treballant amb una variable $X \sim Bin(n,p)$, sota certes condicions, la podrem aproximar per una variable $Y \sim N(np, \sqrt{np(1-p)})$. On la variable Y té la mateixa esperança i variància que X.

Encara que a l'anterior teorema es parla d'aproximació quan n tendeix a infinit, a la pràctica aquesta aproximació és vàlida quan es compleix $np > 5$ i $n(1-p) > 5$.

Cal tenir en compte que passem d'una distribució discreta que pren valors a $\{0,1,2,\cdots,n\}$, a una variable contínua que pren valors a tot $\mathbb{R}$. A més, en el cas de la llei binomial, la probabilitat en un punt és diferent de zero, mentre que no és així en el cas de la llei normal.

Per aquestes raons quan aproximem una distribució binomial per una normal, cal fer una correció de continuïtat de la següent manera:

Si $X \sim Bin(n,p)$ amb $np > 5$ i $n(1 - p) > 5$ i volem calcular $P(a \leq X \leq b)$, considerem la variable $Y \sim N(np, \sqrt{np(1 - p)})$) i calculem $P(a - 0.5 < Y < b + 0.5)$.

Exemple 3.14 *En un magatzem s'ha analitzat durant un any el percentatge de peces defectuoses i se n'ha detectat un* 8% *(podem considerar que la probabilitat que una peça sigui defectuosa és de* 0.08*). S'agafa una mostra de* 100 *peces i es defineix la variable aleatòria X com el nombre de peces defectuoses dins de la mostra de* 100*. La variable aleatòria X segueix una llei binomial* $Bin(100,0.08)$*. Calculem la probabilitat que en les* 100 *peces n'hi hagi entre* 10 *i* 20 *de defectuoses:*

Primer ho calculem sense fer l'aproximació. Com que $X \sim Bin(100,0.08)$*,*

$$P(10 \leq X \leq 20) = \sum_{k=10}^{20} \binom{100}{k} 0.08^k 0.92^{100-k} = 0.28$$

Calculem fent l'aproximació. Com que $np = 8 > 5$ *i* $n(1\text{-}p) = 92 > 5$*, podem utilitzar l'aproximacio normal amb paràmetres* $m = np = 8$ *i* $\sigma = \sqrt{np(1 - p))} = 2.713$*. Si* $Y \sim N(8,2.713)$*, podem escriure:*

$$P(10 \leq X \leq 20) = P(9.5 < Y < 20.5) = 0.29$$

L'últim valor numèric l'hem obtingut amb la calculadora scilab, introduint les ordres indicades a continuació:

```
-->y = integrate('1/(2.713 * sqrt(2*%pi))*exp(-(t-8)^2/(2*2.713^2))','t',9.5,20.5)
 y = 0.2901661
```

4. Funcions de variables aleatòries

De vegades ens interessa estudiar el resultat que s'obté en modificar una variable aleatòria X. Per exemple, si fem passar un senyal, X, a través d'un rectificador, podem obtenir a la sortida un senyal, Y, amb una distribució diferent de X. Parlem d'una variable aleatòria Y que és funció d'una altra X i escrivim $Y = g(X)$. Aquest aspecte ja l'hem tractat de pasada quan hem considerat $E(X - E(X))$, ja que consideràvem la nova variable aleatòria $g(X) = X - E(X)$.

Tractarem alguns casos senzills i enunciarem el teorema de l'esperança.

4.1. Funció d'una variable aleatòria discreta

Suposem que X és una v. a. d. amb valors dins el conjunt $\{a_1,a_2,a_3,\cdots,a_n\}$ (el conjunt també pot ser infinit). Sigui Y una nova v. a. d. definida per una funció $Y = g(X)$. Ens interessa trobar la distribució de probabilitats de Y. Suposem que Y pren valors dins el conjunt $\{b_1,b_2,b_3,\cdots,b_m\}$. Per trobar la probabilitat de cada un d'aquests valors, b_j, hem de trobar la probabilitat del subconjunt de valors de $\{a_1,a_2,a_3,\cdots,a_n\}$ que tenen per imatge b_j. Ho escrivim com a

$$P(Y = b_j) = \sum_{a_i} P(X = a_i) \quad \text{on} \quad g(a_i) = b_j$$

Vegem un exemple.

Exemple 4.1 *Sigui X la v. a. d. que compta el nombre de zeros en un missatge de mida 3 format amb elements del conjunt $\{0,1\}$. És clar que X pren els valors $\{0,1,2,3\}$. Si la probabilitat que hi hagi un zero en una determinada posició és $\frac{1}{2}$, $X \sim Bin(3,\frac{1}{2})$. Definim $Y = g(X)$ com a*

$$g(x) = \begin{cases} 2 & si \quad x \leq 0 \\ 3 & altrament \end{cases}$$

Segons hem definit $g(x)$, Y pot prendre els valors $\{2,3\}$. Calcular la distribució de probabilitats de la variable Y és donar-ne totes les probabilitats. Així,

$$P(Y = 2) = P(X = 0) = \binom{3}{0}\frac{1}{2^3} = \frac{1}{8},$$

$P(Y = 3) = P(X = 1) + P(X = 2) + P(X = 3) = \binom{3}{1}\frac{1}{2^3} + \binom{3}{2}\frac{1}{2^3} + \binom{3}{3}\frac{1}{2^3} =$
$\frac{3+3+1}{16} = \frac{7}{8}.$

És clar que la suma és 1.

4.2. Funció d'una variable aleatòria contínua

Suposem que X és una v. a. amb funció de densitat coneguda $f_X(x)$. Definim una nova v. a. $Y = g(X)$ i el que voldríem és trobar la distribució de Y. Segons la definició de funció de distribució de Y,

$$F_Y(y) = P(Y \leq y) = P(g(X) \leq y)$$

L'equació anterior ens diu que per cada valor de y, hem de trobar la probabilitat de tots els valors de X que satisfan $g(X) \leq y$, per tant, prèviament cal determinar quins són els valors de X que satisfan $g(X) \leq y$. Vegem un exemple.

Exemple 4.2 *Si X segueix una distribució uniforme a l'interval* (8,10)*, vam veure que les seves funcions de densitat i de distribucio són:*

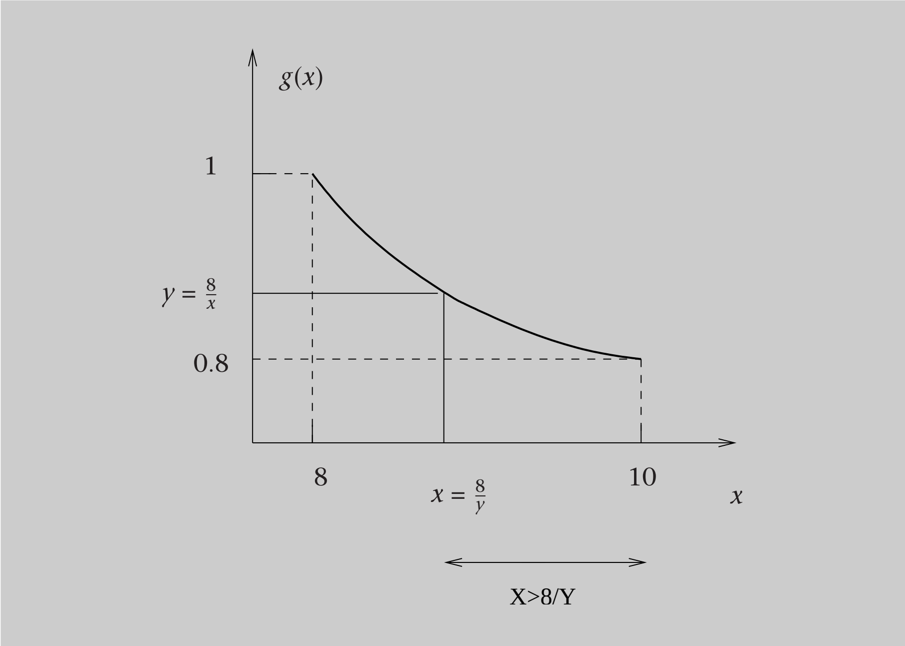
$$f(x) = \begin{cases} \frac{1}{10\text{-}8} & \text{si} \quad x \in (8,10) \\ 0 & \text{altrament} \end{cases}$$

$$F_X(x) = \begin{cases} 0 & \text{si} \quad x < 8 \\ \frac{1}{2}(x - 8) & \text{si} \quad 8 \leq x < 10 \\ 1 & \text{si} \quad 10 \leq x \end{cases}$$

Definim la nova variable $Y = 8/X$ i volem trobar quina distribució segueix.

$$F_Y(y) = P(Y \leq y) = P(\frac{8}{X} \leq y) = P(X \geq \frac{8}{y}) = \frac{10 - \frac{8}{y}}{2} = 5 - \frac{4}{y}$$

Les anteriors igualtats són vàlides quan $8 < \frac{8}{y} < 10$ *o bé* $0.8 < y < 1$.



Funció $g(x) = \frac{8}{x}$

La funció de distribució és:

$$F_Y(y) = \begin{cases} 0 & si \quad y \leq 0.8 \\ 5 - \frac{4}{y} & si \quad 0.8 < y \leq 1 \\ 1 & si \quad y \geq 1 \end{cases}$$

La funció de densitat de Y la trobem derivant la funció de distribució. Tenim

$$f_Y(y) = \begin{cases} \frac{4}{y^2} & si \quad 0.8 < y < 1 \\ 0 & altrament \end{cases}$$

Observem que Y no segueix una distribució uniforme.

Quan la funció $g(x)$ és derivable i esctrictament creixent o estrictament decreixent, podem trobar la funció de densitat de Y d'una altra manera. Vegem-ho cas per cas.

4.2.1. La funció $g(x)$ és estrictament creixent

El fet que $g(x)$ sigui estrictament creixent i derivable ens permet escriure

$$(g^{-1}(y))' = \frac{1}{g'(g^{-1}(y))} = \frac{1}{g'(x)} \quad \text{amb} \quad x = g^{-1}(y)$$

així podem escriure

$$F_Y(y) = P(Y \leq y) = P(g(X) \leq y) = P(X \leq g^{-1}(y)) = F_X(g^{-1}(y))$$

i si derivem aquesta expressió,

$$f_Y(y) = \frac{f_X(x)}{g'(x)} \quad \text{amb} \quad x = g^{-1}(y)$$

Exemple 4.3 *Considerem les v. a. X i* $Y = X^3$. *En aquest cas tenim* $g(x) = x^3$, *que és estrictament creixent i derivable. La relació entre les funcions de densitat és*

$$f_Y(y) = \frac{f_X(x)}{3x^2} = \frac{f_X(x)}{3x^2} \quad amb \quad x = y^{\frac{1}{3}}$$

4.2.2. La funció $g(x)$ és estrictament decreixent

Si la funció $g(x)$ és estrictament decreixent i derivable, obtenim una expressió semblant

$$F_Y(y) = P(Y \leq y) = P(g(X) \leq y) = P(X \geq g^{-1}(y)) = 1 - F_X(g^{-1}(y))$$

i si derivem aquesta expressió,

$$f_Y(y) = -\frac{f_X(x)}{g'(x)} \quad \text{amb} \quad x = g^{-1}(y)$$

Exemple 4.4 *Considerem les v. a. X i* $Y = -X^3$. *En aquest cas tenim* $g(x) = -x^3$, *que és estrictament decreixent i derivable. La relació entre les funcions de densitat és*

$$f_Y(y) = -\frac{f_X(x)}{-3x^2} = \frac{f_X(x)}{3x^2} \quad amb \quad x = (-y)^{\frac{1}{3}}$$

4.2.3. La funció $g(x)$ té extrems locals

En el cas més general on la funció $y = g(x)$ és derivable i tingui un nombre finit o d'extrems locals, cal separar el domini de $g(x)$ en intervals on la funció és monòtona. Això ho aconseguim resolent l'equació $g'(x) = 0$. Si obtenim n intervals, dins de cada un d'aquests intervals anomenem $x_1 = g_1^{-1}(y)$. $x_2 = g_1^{-1}(y), \ldots, x_n = g_n^{-1}(y)$. Per cada valor de $y = g(x)$, la funció de densitat de Y s'obté a partir de la següent expressió

$$f_Y(y) = \frac{f_X(x_1)}{|g'(x_1)|} + \cdots + \frac{f_X(x_n)}{|g'(x_n)|}$$

4.3. Teorema de l'esperança

Moltes vegades només ens interessa trobar el valor esperat o esperança de la variable transformada $Y = g(X)$ i no ens cal trobar prèviament la funció de densitat $f_Y(y)$. El teorema de l'esperança ens permet trobar la esperança, $E(Y)$, de la v. a. Y definida per $Y = g(X)$, encara que no coneixem $f_Y(y)$. Es té

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

Encara que hem enunciat aquest teorema per v. a. c., també és vàlid per v. a. d. i només cal tenir en compte que en lloc d'integrals tindrem sumatoris. Així, si tenim una v. a. d. X que pren valors al conjunt $\{a_1, a_2, \cdots, a_n\}$ i definim una nova v. a. d. $Y = g(X)$, el teorema de l'esperança és:

$$E(Y) = \sum_{i=1}^{n} g(a_i) P(X = a_i)$$

Exemple 4.5 *Amb el mateix enunciat que a l'exemple 4.2, trobarem l'esperança de la variable Y de dues maneres diferents.*

1) *A partir de la definició de $E(Y)$ i a partir de $f_Y(y)$. Havíem trobat*

$$f_Y(y) = \begin{cases} \frac{4}{y^2} & si \quad 0.8 < y < 1 \\ 0 & altrament \end{cases}$$

tenim

$$E(Y) = \int_{-\infty}^{+\infty} y\, f_Y(y)\, dy = \int_{0.8}^{1} y\, \frac{4}{y^2}\, dy = [4\, \ln(y)]_{0.8}^{1} = -4 \ln(0.8).$$

2) *Utilitzant el teorema de l'esperança, on només cal recordar la funció de densitat de la variable original, X:*

$$f_X(x) = \begin{cases} \frac{1}{2} & si \quad 8 < x < 10 \\ 0 & altrament \end{cases}$$

$$E(Y) = \int_{-\infty}^{+\infty} g(x)\, f_X(x)\, dx = \int_{8}^{10} \frac{8}{x} \frac{1}{2} = 4(\ln 10 - \ln 8) = -4\, \ln(0.8)$$

Proposició 4.1 *Propietat de l'esperança*

Aquesta propietat és vàlida tant si X és una v. a. c. o v. a. d.

- $E(a\,g(X) + b\,h(X)) = a\,E(g(X)) + b\,E(h(X))$, *amb* $a, b \in \mathbb{R}$ *i* h, g *funcions de* X.

  Amb aquesta propietat ara podem demostrar una propietat que ja havíem fet servir:

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

En el cas d'una constant c, **tenim que** $E(c) = c$.

Exemple 4.6 *El temps en anys, T, que triga a espatllar-se un component electrònic, segueix una distribució exponencial,* $Exp(1)$. *El cost, Y, de reparació del component durant el primer any és funció de* $2X$, *mentre que després, és de* $3X + 2$. *Trobem el valor mitjà del cost.*

Podem expressar el cost Y com a

$$Y = g(X) = \begin{cases} 2X & si \quad 0 < x < 1 \\ 3X + 2 & si \quad x \geq 1 \end{cases}$$

Designem per E_1, *la mitjana de X, a l'interval* $(0,1)$ *i per* E_2, *la mitjana de X, a l'interval* $(1, +\infty)$. *És a dir,*

$$E_1 = \int_0^1 e^{-x}\,dx = 1 - e^{-1} \quad E_2 = \int_1^{+\infty} e^{-x}\,dx = e^{-1}.$$

A causa de les propietats de la integral, podem dir que $E(X) = E_1(X) + E_2(X)$

$$E(Y) = E_1(2X) + E_2(3X + 2) = 2E_1(X) + 3E_2(X) + 2 = 2(1 - e^{-1}) + 3(e^{-1}) + 2 = 4 + e^{-1}.$$

5. Vectors aleatoris

De vegades ens trobem amb fenòmens que estan relacionats amb més d'una variable aleatòria alhora. Per exemple, en un circuit on la resistència, inductància i capacitat estiguin modelades com a variables aleatòries, haurem de treballar amb tres v.a. alhora. També, si en un circuit prenem 5 mesures d'un valor desconegut, com podria ser la intensitat del corrent, l'error en cada una d'aquestes mesures podria ser modelat per una v.a. i llavors hauríem de treballar amb cinc v.a. alhora. Introduirem el concepte de vector aleatori i estudiarem en particular el cas bidimensional.

5.1. Vector aleatori (X,Y), amb X i Y variables aleatòries discretes

Definició 5.1 *Si X i Y són dues v.a. discretes, s'anomena vector aleatori discret bidimensional al vector (X,Y).*

En general, donades n variables aleatòries discretes, $X_1,X_2,\cdots X_n$, cal treballar amb el vector aleatori discret n-dimensional $(X_1,X_2,\cdots X_n)$.

De vegades (X,Y) s'anomena variable aleatòria bidimensional, i $(X_1,X_2,\cdots X_n)$, variable aleatoria n-dimensional.

5.1.1. Probabilitat conjunta. Probabilitat marginal

Siguin X, Y dues v. a. d. on X pren els valors $\{a_1,a_2,\cdots,a_n\}$ i Y, $\{b_1,b_2,\cdots,b_m\}$.

Definició 5.2 *Per cada parella de valors a_i, b_j, tenim una probabilitat conjunta $P(\{X = a_i\} \cap \{Y = b_j\}) = P(X = a_i,Y = b_j)$. (Per comoditat, adoptem l'expressió de la segona part de la igualtat en lloc de la notació conjuntista.)*

Per comoditat, adoptem la notació $P(X = a_i,Y = b_j)$.

Per al cas particular on X pren els valors $\{a_1,a_2\}$ i Y, $\{b_1,b_2\}$, obtenim la següent taula de probabilitats conjuntes i marginals:

$Y \setminus X$	a_1	a_2	$P(Y = b_j)$
b_1	$P(X = a_1,Y = b_1)$	$P(X = a_2,Y = b_1)$	$P(Y = b_1)$
b_2	$P(X = a_1,Y = b_2)$	$P(X = a_2,Y = b_2)$	$P(Y = b_2)$
$P(X = a_i)$	$P(X = a_1)$	$P(X = a_2)$	1

Definició 5.3 *Per cada valor a_i, obtenim una probabilitat marginal de la variable X, $P(X = a_i) = \sum_{j=1}^{m} P(X = a_i, Y = b_j)$, és a dir, la suma de les probabilitat conjuntes fixat un valor de X i per a tots els valors de Y. A l'última fila de la taula anterior, obtenim les probabilitats marginals de X. Per exemple, $P(X = a_1) = P(X = a_1, Y = b_1) + P(X = a_1, Y = b_2)$. De manera semblant, per cada b_j, obtenim una probabilitat marginal de Y, $P(Y = b_j) = \sum_{i=1}^{n} P(X = a_i, Y = b_j)$. A l'última columna de la taula anterior, obtenim les probabilitats marginals de Y.*

Amb les probabilitats marginals treballem de la mateixa manera que amb una única variable i considerem els mateixos paràmetres que havíem definit al tema de variables aleatòries. En particular, σ_X i σ_Y són les desviacions típiques de X i de Y respectivament.

Exemple 5.1 *Un emissor envia un missatge binari (format amb elements de* $\{0,1\}$*), de mida* 2 *i a l'atzar. Pel canal de transmissió es poden produir errors. Sabem que la probabilitat que un bit arribi al receptor amb error és* $P(e) = 0.02$*. Definim les v. a. de la següent manera: la v.a.* X *compta el nombre de* 0 *que envia l'emissor i la v. a.* Y *el nombre de* 0 *que arriben al receptor. Calculem les probabilitats conjuntes i marginals, així com els paràmetres, esperança i desviació típica.*

Tant la variable X *com la* Y *poden prendre els valors* $\{0,1,2\}$.

Tenim les següents probabilitats conjuntes,

S'emet 11 *i arriba* 11. $P(X = 0, Y = 0) = \frac{1}{4}0.98^2 = 0.2401$.

S'emet 11 *i arriba* 01 *o* 10. $P(X = 0, Y = 1) = \frac{1}{4}2\ 0.98\ 0.02 = 0.0098$.

S'emet 11 *i arriba* 00. $P(X = 0, Y = 2) = \frac{1}{4}0.02^2 = 0.0001$.

S'emet 10 *i arriba* 11 *o s'emet* 01 *i arriba* 11. $P(X = 1, Y = 0) = \frac{1}{4}0.98\ 0.02 + \frac{1}{4}0.02\ 0.98 = 0.0098$.

S'emet 01 *i arriba* 01 *o s'emet* 01 *i arriba* 10 *o s'emet* 10 *i arriba* 10 *o s'emet* 10 *i arriba* 01. $P(X = 1, Y = 1) = \frac{1}{4}0.98^2\ 2 + \frac{1}{4}0.02^2\ 2 = 0.4804$.

S'emet 01 *o* 10 *i arriba* 00. $P(X = 1, Y = 2) = \frac{1}{4}0.02\ 0.98\ 2 = 0.0098$.

S'emet 00 *i arriba* 11. $P(X = 2, Y = 0) = \frac{1}{4}0.02^2 = 0.0001$.

S'emet 00 *i arriba* 01 *o* 10. $P(X = 2, Y = 1) = \frac{1}{4}0.98\ 0.02\ 2 = 0.0098$.

S'emet 00 *i arriba* 00. $P(X = 2, Y = 2) = \frac{1}{4}0.98^2 = 0.2401$.

Calculem ara les probabilitats marginals. És a dir, fixem el valor d'una variable i sumem per tots els valors de l'altra variable.

$P(X = 0) = 0.25$, $P(X = 1) = 0.5$, $P(X = 2) = 0.25$,

$P(Y = 0) = 0.25$, $P(Y = 1) = 0.5$, $P(Y = 2) = 0.25$.

Taula de probabilitats conjuntes:

$Y \setminus X$	0	1	2	$P(Y = b_j)$
0	0.2401	0.0098	0.0001	0.25
1	0.0098	0.4804	0.0098	0.5
2	0.0001	0.0098	0.2401	0.25
$P(X = a_i)$	0.25	0.5	0.25	1

Ara calculem alguns paràmetres que ens donen informació de cada una de les variables.

L'esperança o valor mitjà de la variable X:

$E(X) = 0\ P(X = 0) + 1\ P(X = 1) + 2\ P(X = 2) = 1$.

$E(X^2) = 0^2\ P(X = 0) + 1^2\ P(X = 1) + 2^2\ P(X = 2) = 1.5$.

La variància de la variable X:

$Var(X) = E(X^2) - E(X)^2 = 1.5 - 1^2 = 0.5$.

La desviació típica de la variable X que ens dóna informació de la dispersió dels valors que pren X, en les mateixes unitats:

$\sigma_X = \sqrt{0.5} = 0.707$.

Per la variable Y obtenim: $E(Y) = 1$, $Var(Y) = 0.5$ *i* $\sigma_Y = 0.707$.

5.1.2. Probabilitat condicionada. Independència

Definició 5.4 *La probabilitat que la variable X prengui el valor a_i sabent que Y pren el valor b_j és*

$$P(X = a_i / Y = b_j) = \frac{P(X = a_i,\ Y = b_j)}{P(Y = b_j)}.$$

Definició 5.5 *La probabilitat que la variable Y prengui el valor b_j sabent que X pren el valor a_i és*

$$P(Y = b_j/X = a_i) = \frac{P(X = a_i,\ Y = b_j)}{P(X = a_i)}.$$

Definició 5.6 *Les variables X i Y són independents si i només si,*

$$P(X = a_i,\ Y = b_j) = P(X = a_i)\ P(Y = b_j) \quad \forall\, a_i\ ,b_j$$

Exemple 5.2 *Amb el mateix enunciat que a l'exemple 5.1, ens fem les següents preguntes:*

1) *Sabent que el receptor ha rebut una paraula amb un zero ($Y = 1$), quina és la probabilitat que l'emissor hagi enviat la paraula 00?*

$P(X = 2/Y = 1) = \frac{P(X=2,\ Y=1)}{P(Y=1)} = \frac{0.0098}{0.5} = 0.0196$

2) *Les variables X i Y són independents?*

Com que $P(X = 2, Y = 1) = 0.0098 \neq P(X = 2)\,P(Y = 1) = 0.125$, no són independents.

Definim alguns paràmetres que ens relacionen les dues variables.

5.1.3. Paràmetres: Covariància i coeficient de correlació

Definició 5.7

Esperança del producte:

$$E(XY) = \sum_{a_i\ ,b_j} a_i b_j\, P(X = a_i,\ Y = b_j).$$

Covariància:

$$\begin{aligned} Cov(X,Y) &= E((X - E(X))(Y - E(Y))) = \\ &= \sum_{a_i\ b_j} (a_i - E(X))(b_j - E(Y))P(X = a_i,\ Y = b_j) = \\ &= E(XY) - E(X)E(Y). \end{aligned}$$

Aquest paràmetre ens dóna informació sobre la relació lineal entre X i Y, si les variables creixen conjuntament o no. Per donar una informació normalitzada, cal dividir el valor de la covariància pel producte de les desviacions típiques de les dues variables. Per a això definim el coeficient de correlació.

Coeficient de correlació:

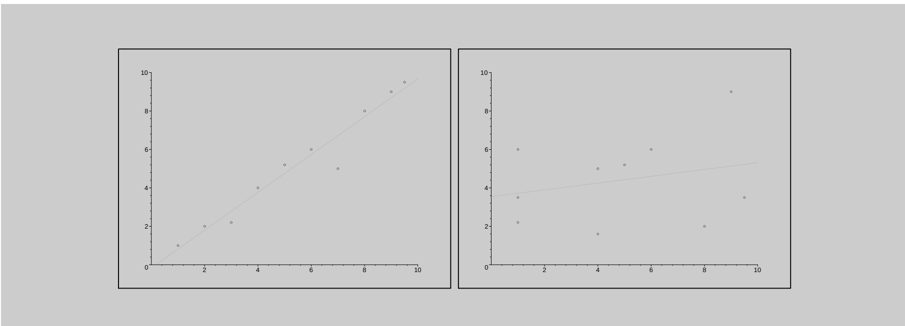
$$\rho = \frac{Cov(X,Y)}{\sigma_X \, \sigma_Y}.$$

Proposició 5.1 *Propietats.*

- $-1 \leq \rho \leq 1$
- *Si ρ es troba a prop d'1 o -1 diem que hi ha una forta correlació entre X i Y.*
- *Si X i Y augmenten o disminueixen conjuntament, $\rho > 0$.*
- *Si una de les variables augmenta en disminuir l'altra (o a l'inrevés), $\rho < 0$.*
- *Si ρ es troba a prop de 0, les variables presenten una correlació dèbil o no hi ha correlació. En el cas particular de $\rho = 0$, diem que les variables són incorrelades.*
- *Si X i Y són independents, llavors $Cov(X,Y) = 0$ ($\rho = 0$). La implicació contrària, en general, no és certa.*

A les següents figures mostrem dos exemples.

Exemple 5.3 *A la figura de l'esquerra hem representat les 10 parelles de valors (x_i,y_i) que pren un vector (X,Y). En aquest cas s'obté un coeficient de correlació de 0.9. Com que és un valor proper a 1, les variables X i Y presenten una correlació forta: a mesura que augmenta el valor d'una, també augmenta el valor de l'altra. El fet que tinguin una correlació lineal forta ens diu que els valors (x_i,y_i) es troben a prop d'una recta. La recta que hem representat l'hem obtinguda pel mètode dels mínims quadrats. Tot i que no és el nostre propòsit obtenir-la, l'hem representada per observar millor la correlació lineal. A la figura de la dreta hem donat un altre exemple on el comportament és completament diferent. En aquest altre cas s'obté un coeficient de corelació de 0.25, pràcticament no hi ha correlació.*



El coeficient de correlació del gràfic esquerre és 0.9 i el del gràfic de la dreta 0.25.

Exemple 5.4 *Continuem amb l'exemple 5.1 i calculem els paràmetres definits anteriorment.*

$$\begin{aligned} E(XY) &= 0\,0\,P(X=0,\,Y=0)+0\,1\,P(X=0,\,Y=1)+0\,2\,P(X=0,\,Y=2)+ \\ &+ 1\,0\,P(X=1,\,Y=0)+1\,1\,P(X=1,\,Y=1)+1\,2\,P(X=1,\,Y=2)+ \\ &+ 2\,0\,P(X=2,\,Y=0)+2\,1\,P(X=2,\,Y=1)+2\,2\,P(X=2,\,Y=2)= \\ &= 0.4804+2\,0.0098+2\,0.0098+4\,0.2401=1.4 \end{aligned}$$

Així $Cov(X,Y) = 1.4 - 1\,1 = 0.4$ *i el coeficient de correlació* $\rho = \frac{0.4}{\sqrt{0.5}\,\sqrt{0.5}} = 0.8$

5.2. Vector aleatori (X,Y), amb X i Y variables aleatòries contínues

Definició 5.8 *Si X i Y són dues v.a. contínues, s'anomena vector aleatori bidimensional continu al vector* (X,Y).

En general, donades n variables aleatòries contínues, $X_1,X_2,\cdots X_n$, cal treballar amb el vector aleatori continu n-dimensional $(X_1,X_2,\cdots X_n)$.

Ja vam veure en el tema de variables aleatòries que el tractament amb variables contínues és molt diferent que amb variables discretes. Comencem definint la funció de distribució conjunta.

5.2.1. Funció de distribució conjunta. Funció de densitat conjunta

Definició 5.9 *La funció de distribució conjunta,F, de les variables contínues X i Y és una aplicació de* $\mathbb{R}^2$ *a* $[0,1]$ *definida per:*

$$F_{XY}(x,y) = P(X \leq x,\, Y \leq y) \quad \forall x,\, y \in \mathbb{R}^2.$$

Si F_{XY} *és dues vegades derivable, la funció de densitat conjunta,* f_{XY}*, és*

$$f_{XY}(x,y) = \frac{\partial^2 F_{XY}}{\partial x \partial y}(x,y).$$

També podem escriure

$$F_{XY}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{XY}(u,v)\,du\,dv.$$

Generalitzant el que vam veure en una dimensió, la probabilitat $P(D)$, on D és un domini a $\mathbb{R}^2$, és determinada pel volum per sota de la funció de densitat conjunta $f_{XY}(x,y)$ i que determina D.

Definició 5.10 *Distribució uniforme*

Diem que el vector aleatori (X,Y) es distribueix uniformement a la regió D si la funció de densitat conjunta és

$$f_{XY}(x,y) = \begin{cases} \frac{1}{\text{àrea}(D)} & si \quad (x,y) \in D \\ 0 & altrament \end{cases}$$

En aquest cas, el càlcul de volums per sota de la funció de densitat pot no requerir la utilització de les integrals dobles (vegem l'exemple 5.6).

5.2.2. Funcions de densitat marginals. Esperances

Definició 5.11 *Densitat marginal de X: Fixat un valor de x integrem per tots els possibles valors de y. Densitat marginal de Y: Fixat un valor de y integrem per tots els possibles valors de x.*

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)\,dy \qquad f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)\,dx.$$

A partir de les anteriors expressions podem trobar els paràmetres que caracteritzen cada una de les variables:

Definició 5.12

Esperança de X i esperança de Y:

$$E(X) = \int_{-\infty}^{+\infty} x\,f_X(x)\,dx, \qquad E(Y) = \int_{-\infty}^{+\infty} y\,f_Y(y)\,dy.$$

Moments d'ordre 2:

$$E(X^2) = \int_{-\infty}^{+\infty} x^2\,f_X(x)\,dx, \qquad E(Y^2) = \int_{-\infty}^{+\infty} y^2\,f_Y(y)\,dy.$$

Variància de X i variància de Y:

$$\sigma_X^2 = E(X^2) - E(X)^2, \qquad \sigma_Y^2 = E(Y^2) - E(Y)^2.$$

5.2.3. Probabilitat condicionada. Variables independents

Definició 5.13 *Les variables contínues X i Y són independents si i només si,*

$$f_{XY}(x,y) = f_X(x)f_Y(y) \qquad \forall\, x, y.$$

Definició 5.14 *Es defineix la densitat de X condicionada a $Y = y$ com a $f(x/y) = \frac{f_{XY}(x,y)}{f_Y(y)}$.*

Igualment, la densitat de Y condicionada a $X = x$ com a $f(y/x) = \frac{f_{XY}(x,y)}{f_X(x)}$.

Exemple 5.5 *El senyal d'entrada, X (volts), en un canal de comunicacions, es troba distribuït uniformement a l'interval* [-2,2]. *El senyal de sortida, Y (volts), és la suma del senyal d'entrada més un soroll que es troba uniformement distribuït a l'interval* [-3,3]. *Calculem les probabilitats condicionades $P(Y \leq 0/X = 1)$, $P(Y \leq y/X = 1)$ i $P(Y \leq y/X = x)$.*

Si el senyal d'entrada és x, llavors Y es distribueix uniformement a $[x - 3, x + 3]$. *En particular, si el senyal d'entrada és $X = 1$, Y es distribueix uniformement a* [-2,4].

Llavors $P(Y \leq 0/X = 1) = \frac{2}{6}$. També $P(Y \leq y/X = 1) = \frac{y+2}{6}$, per a $-2 < y < 4$.

El cas més general, $P(Y \leq y/X = x) = \frac{y-x+3}{6}$, per a $x - 3 < y < x + 3$.

5.2.4. Paràmetres: Covariància i coeficient de correlació

Definim els mateixos paràmetres que hem vist per a vectors discrets. Ara els sumatoris esdevenen integrals.

Definició 5.15

Esperança del producte:

$E(XY) = \int_{-\infty}^{+\infty} x\, y\, f_{XY}(x, y)\, dx\, dy.$

Covariància:

$$\begin{aligned} Cov(X,Y) &= E((X - E(X))(Y - E(Y))) = \\ &= \int_{-\infty}^{+\infty} (x - E(X))\,(y - E(Y))\, f_{XY}(x, y)\, dx\, dy = \\ &= E(XY) - E(X)E(Y). \end{aligned}$$

Aquest paràmetre ens dóna informació sobre la relació lineal entre X i Y (si les variables creixen conjuntament o no). Per donar una informació quantificada, cal normalitzar la covariància i per això definim el coeficient de correlació.

Coeficient de correlació:

$$\rho = \frac{Cov(X,Y)}{\sigma_X\ \sigma_Y}.$$

Proposició 5.2 *Propietats.*

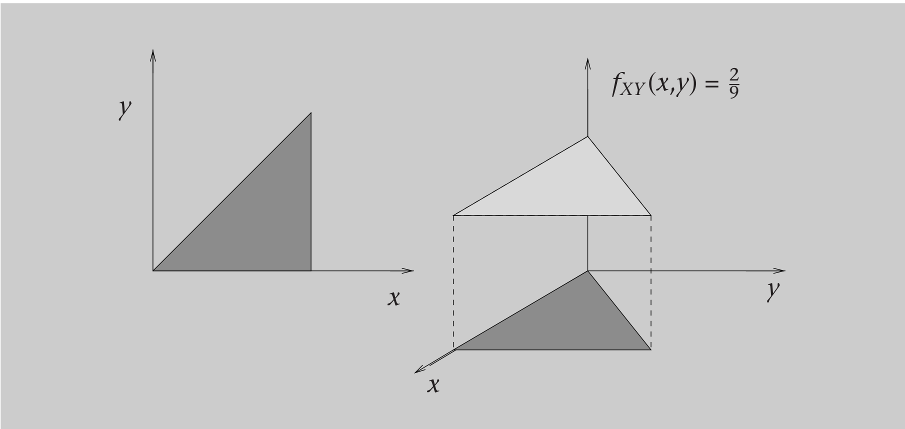
- $-1 \leq \rho \leq 1$.
- *Si ρ es troba a prop d'1 o -1 diem que hi ha una forta correlació entre X i Y.*
- *Si X i Y augmenten o disminueixen conjuntament, $\rho > 0$.*
- *Si una de les variables augmenta en disminuir l'altra (o a l'inrevés), $\rho < 0$.*
- *Si ρ es troba a prop de 0, les variables presenten una correlació dèbil o no hi ha correlació. En el cas particular de $\rho = 0$, diem que les variables són incorrelades.*
- *Si X i Y són independents, llavors $Cov(X,Y) = 0$ ($\rho = 0$). La implicació contrària, en general, no és certa.*

Exemple 5.6 *(X,Y) és un vector aleatori bidimensional uniforme a la regió limitada pel triangle, T, de costats sobre les rectes $y = 0$, $x = 3$ i $y = x$.*

1) *Trobem la funció de densitat conjunta, funcions de densitat marginals i el valor esperat de cada una de les variables. Les variables X i Y són independents?*

Com que la distribució és uniforme,

$$f_{XY}(x,y) = \begin{cases} \frac{1}{\text{àrea triangle}} = \frac{2}{9} & si \quad (x,y) \in T \\ 0 \quad altrament & \end{cases}$$



Densitat uniforme al domini T

Densitat marginal de X:

$$f_X(x) = \begin{cases} \int_{-\infty}^{+\infty} f(x,y)\, dy = \int_0^x \frac{2}{9}\, dy = \frac{2}{9}x \quad si \quad x \in [0,3]. \\ 0 \quad altrament \end{cases}$$

Densitat marginal de Y:

$$f_Y(y) = \begin{cases} \int_{-\infty}^{+\infty} f(x,y)\, dx = \int_y^3 \frac{2}{9}\, dx = \frac{2}{9}(3-y) \quad si \quad y \in [0,3]. \\ 0 \quad altrament \end{cases}$$

Valor esperat de X:

$$E(X) = \int_{-\infty}^{+\infty} x f_X(x)\, dx = \frac{2}{9}\int_0^3 x^2\ dx = \frac{2}{9}\frac{3^3}{3} = 2.$$

Valor esperat de Y:

$$E(Y) = \int_{-\infty}^{+\infty} y f_Y(y)\, dy = \frac{2}{9}\int_0^3 (3-y)y\ dy = \frac{2}{9}(\frac{3^3}{2} - \frac{3^3}{3}) = 1.$$

Les variables no són independents perquè $f_{XY}(x,y) \neq f_X(x) f_Y(y)$.

2) *Trobem el coeficient de correlació.*

Moment d'ordre 2 de X:

$$E(X^2) = \int_{-\infty}^{+\infty} x^2 f_X(x)\, dx = \frac{2}{9}\int_0^3 x^3\ dx = \frac{2}{9}\frac{3^4}{4} = \frac{9}{2}.$$

Moment d'ordre 2 de Y:

$$E(Y^2) = \int_{-\infty}^{+\infty} y^2 f_Y(y)\, dy = \frac{2}{9}\int_0^3 (3-y)y^2\ dy = \frac{2}{9}(3^3 - \frac{3^4}{4}) = \frac{3}{2}.$$

Variància de X i σ_X:

$$Var(X) = E(X^2) - E(X)^2 = \frac{9}{2} - 4 = \frac{1}{2}. \quad \sigma_X = \sqrt{\frac{1}{2}}.$$

Variància de Y i σ_Y:

$$Var(Y) = E(Y^2) - E(Y)^2 = \frac{3}{2} - 1 = \frac{1}{2}. \quad \sigma_Y = \sqrt{\frac{1}{2}}.$$

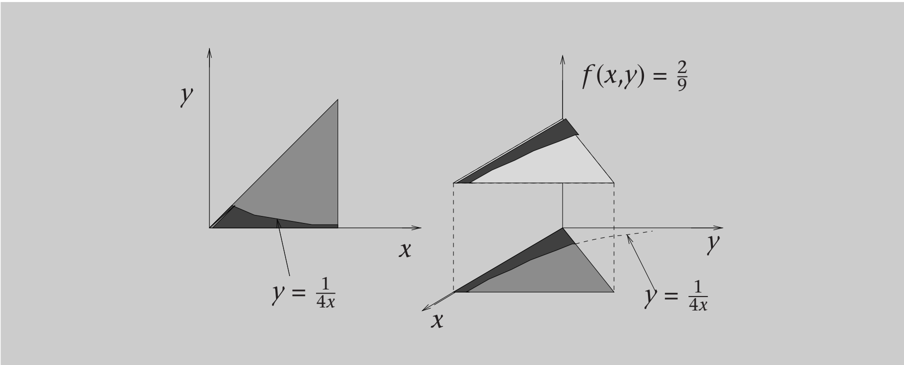
Esperança del producte:

$$E(XY) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x\,y\,f(x,y)\,dx\,dy =$$

$$= \frac{2}{9}\int_{x=0}^{3}\int_{y=0}^{x} x\,y\,dy\,dx = \frac{2}{9}\int_{x=0}^{3}\frac{x^3}{3}\,dx = \frac{2}{9}\frac{3^4}{12} = \frac{9}{4}.$$

Covariància i coeficient de correlació:

$$Cov(X,Y) = E(XY) - E(X)E(Y) = \frac{9}{4} - 2\;1 = \frac{1}{4} \qquad \rho = \frac{\frac{1}{4}}{\sqrt{\frac{1}{2}}\sqrt{\frac{1}{2}}} = \frac{1}{2}.$$



Funció de densitat uniforme al triangle T

3) *Trobem les probabilitats:* $P(X < \frac{1}{2})$, $P(Y < \frac{1}{2})$ *i* $P(X \cdot Y < \frac{1}{4})$.

$$P(X < \frac{1}{2}) = \int_{x=0}^{\frac{1}{2}} \frac{2}{9}\,x\,dx = \frac{1}{36}.$$

Representem el volum que determina la zona del triangle on $x < \frac{1}{2}$.

$$P(Y < \frac{1}{2}) = \int_{y=0}^{\frac{1}{2}} \frac{2}{9}\,(3 - y)\,dy = \frac{11}{36}.$$

Representem el volum que determina la zona del triangle on $y < \frac{1}{2}$.

$$P(X \cdot Y < \frac{1}{4}) = \int_{x=0}^{\frac{1}{2}}\int_{y=0}^{x} \frac{2}{9}\,dy\,dx + \int_{\frac{1}{2}}^{3}\int_{y=0}^{\frac{1}{4x}} \frac{2}{9}\,dy\,dx = \frac{1}{36} + \frac{\ln 3 + \ln 2}{18}.$$

A la figura mostrem la zona del triangle on $x \cdot y < \frac{1}{4}$ *i el volum que genera.*

6. Resum

Tècniques de comptar

Variacions amb repetició: $VR_{n,k} = n^k$.

Variacions: $V_{n,m} = n \cdot (n-1) \cdots (n-m+1)$.

Factorial de n: $V_{n,n} = n \cdot (n-1) \cdots 1 = n!$.

Combinacions: $C_{n,m} = \binom{n}{m} = \frac{n \cdot (n-1) \cdots (n-m+1)}{m!} = \frac{m!}{m!(n-m)!}$.

Probabilitat

Relació entre conjunts i successos:

En termes de probabilitat	En termes de conjunts	Notació
Succés segur	Conjunt total	Ω
Succés impossible	Conjunt buit	$\varnothing$
Succés contrari	Conjunt complementari	A^c, també $\overline{A}$
A i B	Intersecció	$A \cap B$
A o B	Unió	$A \cup B$
Successos incompatibles	Conjunts disjunts	$A \cap B = \varnothing$
Sistema complet de successos	Partició de Ω	$A_i \cap A_j = \varnothing$ $\cup A_i = \Omega$

Probabilitat del complementari de A: $P(A^c) = 1 - P(A)$.

Probabilitat de la unió: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

En un espai equiprobable: $P(A) = \dfrac{\text{nombre de casos favorables}}{\text{nombre de casos possibles}}$.

Probabilitat condicionada: $P(A/B) = \frac{P(A \cap B)}{P(B)}$.

Si A i B són independents, $P(A \cap B) = P(A)P(B)$.

Teorema de la probabilitat total

Si $A_1, A_2, \cdots A_n$ és un sistema complet de successos:

$$P(B) = P(B/A_1)P(A_1) + P(B/A_2)P(A_2) + \cdots + P(B/A_n)P(A_n).$$

Teorema de Bayes

Si $A_1, A_2, \cdots A_n$ és un sistema complet de successos:

$$P(A_i/B) = \frac{P(B/A_i)P(A_i)}{P(B/A_1)P(A_1) + P(B/A_2)P(A_2) + \cdots + P(B/A_n)P(A_n)}.$$

Variables aleatòries discretes

Valor mitjà o esperança : $E(X) = \sum_{i=1}^{n} a_i P(X = a_i)$.

Moment d'ordre 2: $E(X^2) = \sum_{i=1}^{n} a_i^2 P(X = a_i)$.

Moment d'ordre n: $E(X^n) = \sum_{i=1}^{n} a_i^n P(X = a_i)$.

Variància:

$Var(X) = E((X - E(X))^2) = \sum_{i=1}^{n} (a_i - E(X))^2 P(X = a_i) = E(X^2) - E(X)^2$.

Desviació típica: $\sigma = \sqrt{Var(X)}$.

Distribucions de variables aleatòries discretes més importants:

$X \sim$	k	$P(X = k)$	$E(X)$	$Var(X)$
$B(p)$	$\{0,1\}$	$P(X=1) = p$ $P(X=0) = 1 - p$	p	$p(1 - p)$
$Bin(n,p)$	$\{0,1,2,\ldots, n\}$	$P(X=k) = \binom{n}{k} p^k (1 - p)^{n-k}$	np	$np(1 - p)$
$Geom(p)$	$\{1,2,3,\ldots\}$	$P(X=k) = (1 - p)^{k-1} p$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
$Poiss(\lambda)$	$\{0,1,2,\ldots\}$	$P(X=k) = \frac{\lambda^k}{k!} e^{-\lambda}$	λ	λ

Funció de distribució: $F_X(x) = P(X \leq x\} \quad \forall x \in \mathbb{R}$.

Variables aleatòries contínues

Funció de distribució: $F_X(x) = P(X \leq x\} \quad \forall x \in \mathbb{R}$.

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

Funció de densitat: $f_X(x) = \frac{dF_X(x)}{dx} \quad \forall x \in \mathbb{R}$.

$P(a < X \leq b) = F_X(b) - F_X(a) = \int_a^b f_X(x)\,dx$.

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

Valor mitjà o esperança: $E(X) = \int_{-\infty}^{+\infty} x f(x)\,dx$.

Moment d'ordre 2: $E(X^2) = \int_{-\infty}^{+\infty} x^2 f(x)\,dx$.

Moment d'ordre n: $E(X^n) = \int_{-\infty}^{+\infty} x^n f(x)\, dx$.

Variància: $Var(X) = \int_{-\infty}^{+\infty} (x - E(X))^2 f(x)\, dx = E(X^2) - E(X)^2$.

Desviació típica: $\sigma = \sqrt{Var(X)}$.

Distribucions de variables aleatòries contínues més importants:

$X \sim$	Funció de densitat	$E(X)$	$Var(X)$
$U(a,b)$	$f(x) = \begin{cases} \frac{1}{b-a} & \text{si } x \in (a,b) \\ 0 & \text{altrament} \end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
$N(m,\sigma)$	$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}} \quad \forall x \in \mathbb{R}$	m	σ^2
$Exp(\lambda)$	$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x \geq 0 \\ 0 & \text{altrament} \end{cases}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$

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

Vectors aleatoris (X,Y) discrets

Probabilitat conjunta: $P(X = a_i \cap Y = b_j) = P(X = a_i, Y = b_j)$.

Probabilitat marginal de X: $P(X = a_i) = \sum_{j=1}^{m} P(X = a_i, Y = b_j)$.

Probabilitat marginal de Y: $P(Y = b_j) = \sum_{i=1}^{n} P(X = a_i, Y = b_j)$.

$P(X = a_i / Y = b_j) = \frac{P(X=a_i,\ Y=b_j)}{P(Y=b_j)}$.

X i Y són independents si i només si,

$P(X = a_i,\ Y = b_j) = P(X = a_i)\ P(Y = b_j) \quad \forall a_i, b_j$.

$E(XY) = \sum_{a_i, b_j} a_i, b_j\, P(X = a_i,\ Y = b_j)$.

$Cov(X,Y) = E((X - E(X))(Y - E(Y))) = E(XY) - E(X)E(Y)$.

$\rho = \frac{Cov(X,Y)}{\sigma_X\ \sigma_Y}$.

Vectors aleatoris (X,Y) continus

Distribució uniforme:

$$f_{XY}(x,y) = \begin{cases} \frac{1}{\text{àrea}(D)} & \text{si} \quad (x,y) \in D. \\ 0 & \text{altrament} \end{cases}$$

Densitats marginals de X i de Y:

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)\,dy, \qquad f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)\,dx.$$

Esperances de X i de Y:

$$E(X) = \int_{-\infty}^{+\infty} x\,f_X(x)\,dx, \qquad E(Y) = \int_{-\infty}^{+\infty} x\,f_Y(y)\,dy.$$

Moments d'ordre 2:

$$E(X^2) = \int_{-\infty}^{+\infty} x^2\,f_X(x)\,dx, \qquad E(Y^2) = \int_{-\infty}^{+\infty} y^2\,f_Y(y)\,dy.$$

Desviació típica de X i desviació típica de Y:

$$\sigma_X = \sqrt{E(X^2) - E(X)^2}, \qquad \sigma_Y = \sqrt{E(Y^2) - E(Y)^2}.$$

X i Y són independents si i només si:

$$f_{XY} = f_X(x) f_Y(y) \qquad \forall\, x, y.$$

Densitat de X condicionada a $Y = y$: $f(x/y) = \frac{f_{XY}(x,y)}{f_Y(y)}$.

Densitat de Y condicionada a $X = x$: $f(y/x) = \frac{f_{XY}(x,y)}{f_X(x)}$.

$E(XY) = \int_{-\infty}^{+\infty} x\,y\,f_{XY}(x, y)\,dx\,dy$.

$Cov(X,Y) = E((X - E(X))(Y - E(Y))) = E(XY) - E(X)E(Y)$.

$\rho = \frac{Cov(X,Y)}{\sigma_X\ \sigma_Y}$.