

PROBABILITAT I PROCESSOS ESTOCÀSTICS

Notes de Classe

J. M. Aroca

Departament de Matemàtica Aplicada IV,
Universitat Politècnica de Catalunya

e-mail: aroca@ma4.upc.edu
<http://www-ma4.upc.edu/~aroca>

Febrer de 2009

Índex

Notació	4
I Teoria de la probabilitat i variables aleatòries unidimensionals	7
1 Introducció a la teoria de la probabilitat	8
1.1 Experiment aleatori	8
1.2 Espai mostral	8
1.3 Esdeveniment aleatori	8
1.4 Àlgebra d'esdeveniments	9
1.5 Espai de probabilitat	10
1.6 Espais de probabilitat discrets	11
1.7 Espais de probabilitat continus	11
1.8 Esdeveniments independents	12
1.9 Probabilitat condicionada	12
1.10 Teorema de Bayes	13
1.11 Espai producte. Proves repetides	13
1.12 Significat de la probabilitat	14
2 Variables aleatòries	16
2.1 Variable aleatòria	16
2.2 Funció de distribució	17
2.3 Variables aleatòries discretes	18
2.4 Variables aleatòries contínues	20
2.5 Teorema de DeMoivre-Laplace	24
2.6 Funcions de densitat i distribució condicionades	24
Apèndix: Demostració del teorema de DeMoivre-Laplace	26
3 Funcions d'una variable aleatòria	27
3.1 Cas discret	27
3.2 Cas continu	28
3.3 Càlcul de la funció de densitat	30
4 Paràmetres estadístics	32

4.1	Esperança	32
4.2	Variància	33
4.3	Desviació estàndard	34
4.4	Variable aleatòria normalitzada	34
4.5	Paràmetres de les distribucions usuals	34
4.6	Moments d'una variable aleatòria	36
4.7	Desigualtat de Txebyshev	37
4.8	Llei dels grans nombres	37
4.9	Teorema del límit central	38
4.10	Probabilitat, esperança, mesura	38
II	Variables aleatòries multidimensionals i teoria de l'estimació	39
5	Variables aleatòries multidimensionals	40
5.1	Variable aleatòria bidimensional	40
5.2	Funció de distribució conjunta	40
5.3	Funcions de distribució marginals	41
5.4	Funció de densitat conjunta	41
5.5	Funcions de densitat marginals	42
5.6	Cas discret	43
5.7	Algunes distribucions usuals	44
5.8	Independència de variables aleatòries	45
5.9	Densitat condicionada	45
5.10	Esperança condicionada	46
5.11	Generalització al cas n -dimensional	47
6	Funcions de vàries variables aleatòries	49
6.1	Suma de variables aleatòries. Teorema de convolució	49
6.2	Altres funcions. Canvi de variables	51
7	Paràmetres estadístics de vàries v.a.'s	55
7.1	Esperança	55
7.2	Covariància	55
7.3	Coefficient de correlació	56
7.4	Moments conjunts	56
7.5	Ortogonalitat	57
7.6	Incorrelació	57
8	Estimació de variables aleatòries	59
8.1	Concepte i criteris d'estimació	59
8.2	Estimació per una constant	60

8.3	Estimació no lineal	60
8.4	Estimació lineal	60
8.5	Principi d'ortogonalitat	61
III	Processos estocàstics	63
9	Introducció als processos estocàstics	64
9.1	Definició de procés estocàstic	64
9.2	Funcions de distribució i densitat d'un PE	65
9.3	Valor mitjà, autocorrelació i autocovariància	65
9.4	Processos estocàstics normals	66
9.4.1	Variable gaussiana n -dimensional	66
9.4.2	Processos estocàstics gaussians	68
9.5	Processos estacionaris	68
10	Exemples de processos estocàstics	70
10.1	El procés de Poisson	70
10.2	Senyal telegràfic aleatori	71
10.3	Tren d'impulsos de Poisson	72
10.4	Distribució dels temps en el procés de Poisson	73
10.5	Senyals binaris	74
10.6	Oscil·lacions aleatòries	75
11	Processos estacionaris	77
11.1	Realitzacions d'un PE	77
11.2	Límit en mitjana quadràtica	77
11.3	Continuïtat estocàstica	77
11.4	Integrals estocàstiques	78
11.5	Ergodicitat	79
11.6	Condicions suficients d'ergodicitat	80
11.7	Espectre de potència d'un procés estacionari	81
11.8	Sistemes lineals	81
11.9	Valor mitjà, autocorrelació i espectre de potència d'un PE transformat linealment	82
IV	Taules	85
	Funció d'error	86
	Taula de variables aleatòries	88

V	Apèndixs	89
A	Lògica i conjunts	90
B	Combinatòria	92
C	Sèries	95
D	Integrals	97
E	Distribucions	99
F	Simulació numèrica de variables i processos aleatoris	101
VI	Problemes	104

Notació

♦ Final d'exemple.

DEM: Inici de demostració.

♣ Final de demostració.

$u(x)$ Funció de Heaviside. (1 si $x \geq 0$ i 0 si $x < 0$.)

$\delta(x)$ Delta de Dirac (veure apèndix sobre distribucions).

$(f * g)(x)$ Producte de convolució.

$f^*(x)$ Conjugació complexa.

Ω Espai mostral. (Possibles resultats d'un experiment aleatori.)

$\mathcal{P}(\Omega)$ Parts de Ω . (Conjunt dels subconjunts de Ω .)

\mathcal{F} Àlgebra d'esdeveniments.

A, B, C, \dots Esdeveniments. (Subconjunts de Ω .)

$P(A)$ Probabilitat de l'esdeveniment A .

$P(A | B)$ Probabilitat de A condicionada a B (És a dir, sabent que B s'ha produït.)

$P(A \text{ i } B), P(A \cap B), P(A, B)$ Probabilitat que es verifiquin els esdeveniments A i B .

$P(A \text{ o } B), P(A \cup B)$ Probabilitat que es verifiqui algun dels esdeveniments A o B .

X, Y, Z, \dots Variable aleatòries (En majúscules. Els valors numèrics es posen en minúscula. Així, per exemple, $X < x$ és l'esdeveniment: la variable aleatòria X és menor que el nombre x .)

$F_X(x)$ Funció de distribució de la variable aleatòria X .

$f_X(x)$ Funció de densitat de la variable aleatòria contínua X .

$P_X(x)$ Funció de probabilitat de la variable aleatòria discreta X .

$E[X], \overline{X}, m_X$ Esperança de la variable aleatòria X .

$V[X], \sigma_X^2$ Variància de la variable aleatòria X .

σ_X Desviació estàndard de la variable aleatòria X .

$F_{XY}(x, y)$ Funció de distribució conjunta de les variables aleatòries X i Y .

$f_{XY}(x, y)$ Funció de densitat conjunta de les variables aleatòries X i Y .

$f_X(x | y)$ Funció de densitat de la variable aleatòria X condicionada a $Y = y$. (De manera similar tenim $f_Y(y | x)$.)

$C[X, Y]$ Covariància de les variables aleatòries X i Y .

ρ Coeficient de correlació.

$m(t)$ Funció de valor mitjà d'un procés estocàstic.

$R(t_1, t_2)$ Funció d'autocorrelació d'un procés estocàstic.

$C(t_1, t_2)$ Funció d'autocovariància d'un procés estocàstic.

m_T Mitjana temporal d'un procés estocàstic.

$S(\omega)$ Espectre de potència d'un procés estocàstic.

Part I

Teoria de la probabilitat i variables aleatòries unidimensionals

Capítol 1

Introducció a la teoria de la probabilitat

1.1 Experiment aleatori

És un experiment \mathcal{E} que es pot repetir tantes vegades com es vulgui sempre en les mateixes condicions. Cada repetició de l'experiment s'anomena prova.

1.2 Espai mostral

És el conjunt Ω dels possibles resultats d'un experiment aleatori \mathcal{E} .

Exemple 1.1 \mathcal{E} = “Tirar un dau”, $\Omega = \{1, 2, 3, 4, 5, 6\}$.♦

Exemple 1.2 L'espai mostral és el que determina clarament la informació que extraïem de l'experiment. Si, per exemple, \mathcal{E} = “Tirar dos daus” hem d'especificar si podem distingir els dos daus una vegada tirats amb el que l'espai mostral és $\Omega_1 = \{(a, b) \mid 1 \leq a, b \leq 6\}$ que té 36 elements o si els daus son indistingibles i tenim $\Omega_2 = \{\{a, b\} \mid 1 \leq a, b \leq 6\}$ que té 21 elements. És a dir, en el segon cas no podem distingir, per exemple, les configuracions (1,2) i (2,1).♦

1.3 Esdeveniment aleatori

Donat un experiment \mathcal{E} un esdeveniment aleatori és una proposició lògica S tal que a l'efectuar una prova de \mathcal{E} podem determinar la seva veritat o falsedat. Considerem $\mathcal{S} = \{\text{conjunt d'esdeveniments } S\}$. Podem identificar un esdeveniment S amb el subconjunt de Ω format pels elements que fan que S sigui cert. D'aquesta manera identifiquem $\mathcal{S} = \mathcal{P}(\Omega)$. Per exemple, \mathcal{E} = “Tirar un dau”, S = “Que surti parell” = $\{2, 4, 6\}$. Com es veu cometrem l'abús de llenguatge de fer servir el mateix nom per la proposició i el subconjunt.

Aquesta identificació ens mou isomòrficament entre l'àlgebra de Boole de les proposicions lògiques (*i*, *o*, *no*) i l'àlgebra de Boole dels conjunts (unió, intersecció, complementari.) Les equivalències es resumeixen en la següent taula (veure l'apèndix A):

disjunció	$R \text{ o } S$	$R \cup S$	unió
conjunció	$R \text{ i } S$	$R \cap S$	intersecció
negació	no R	$\Omega - R$	complementari
implicació	$R \rightarrow S$	$R \subset S$	inclusió
equivalència	$R \leftrightarrow S$	$R = S$	igualtat

El complementari $\Omega - A$ el denotarem \overline{A} . L'esdeveniment \emptyset és sempre fals. Es pot pensar com $A \cap \overline{A}$, és a dir “A i no A” (contradicció). L'esdeveniment Ω és sempre cert. Es pot pensar com $A \cup \overline{A}$, és a dir “A o no A” (tautologia).

1.4 Àlgebra d'esdeveniments

És una família \mathcal{F} no buida de subconjunts de Ω tancada per la unió i el complementari. És a dir, $\mathcal{F} \subset \mathcal{P}(\Omega)$ verificant:

- (1) $\mathcal{F} \neq \emptyset$,
- (2) si $A \in \mathcal{F}$ llavors $\overline{A} \in \mathcal{F}$,
- (3) si $A, B \in \mathcal{F}$ llavors $A \cup B \in \mathcal{F}$.

D'aquí es desprenen les següents propietats:

- (4) $\emptyset, \Omega \in \mathcal{F}$,
- (5) si $A, B \in \mathcal{F}$ llavors $A \cap B \in \mathcal{F}$.

DEM:

(4) Per (1) existeix $A \in \mathcal{F}$. Per (2) $\overline{A} \in \mathcal{F}$. Per (3) $\Omega = A \cup \overline{A} \in \mathcal{F}$. Per (2) $\emptyset = \overline{\Omega} \in \mathcal{F}$.

(5) Siguin $A, B \in \mathcal{F}$ i recordem la propietat de les operacions entre conjunts $\overline{A \cap B} = \overline{A} \cup \overline{B}$. Llavors $A \cap B = \overline{\overline{A \cap B}} = \overline{\overline{A} \cup \overline{B}} \in \mathcal{F}$. ♣

Un exemple trivial d'àlgebra d'esdeveniments és considerar tot el conjunt $\mathcal{P}(\Omega)$. A vegades ens interessarà treballar amb conjunts més petits d'esdeveniments, ja sigui perquè només volem considerar uns esdeveniments determinats o perquè no podem assignar probabilitats quan $\mathcal{P}(\Omega)$ és molt complicat (cas continu). El que demanem és que el conjunt considerat sigui tancat sota les operacions lògiques *i*, *o* i *no*. Dit d'una altra manera, a l'igual que $\mathcal{P}(\Omega)$, \mathcal{F} ha de ser una àlgebra de Boole amb la unió, la intersecció i el complementari.

Exemple 1.3 En l'experiment "Tirar un dau" $\Omega = \{1, 2, 3, 4, 5, 6\}$. $\mathcal{P}(\Omega)$ té $2^6 = 64$ elements. És fàcil comprovar que $\mathcal{F} = \{\emptyset, \Omega, \{2, 4, 6\}, \{1, 3, 5\}\}$ és una àlgebra d'esdeveniments. Per tant els únics esdeveniments no trivials que considerem són "surt parell" i "surt senar". Així podem, per exemple, simular un experiment amb dos resultats equiprobables com el de tirar una moneda. ♦

La propietat (3) s'estén per inducció a la unió d'un nombre finit qualsevol d'esdeveniments. A vegades convé considerar unions numerables d'esdeveniments. En aquest cas diem que \mathcal{F} és una σ -àlgebra si en lloc de (3) es verifica

$$(3') \text{ si } A_i \in \mathcal{F}, i \in \mathbb{N} \text{ llavors } \bigcup_{i \in \mathbb{N}} A_i \in \mathcal{F}.$$

La propietat (5) passa a (5') posant interseccions numerables. Òbviament tota σ -àlgebra és una àlgebra però no a la inversa.

1.5 Espai de probabilitat

És una terna (Ω, \mathcal{F}, P) on Ω és un conjunt, \mathcal{F} una àlgebra (o σ -àlgebra) de Ω i P una funció probabilitat. És a dir, $P : \mathcal{F} \rightarrow \mathbb{R}$ verificant:

- (1) $P(\Omega) = 1$,
- (2) $P(A) \geq 0$ per a tot $A \in \mathcal{F}$,
- (3) si $A, B \in \mathcal{F}$ i $A \cap B = \emptyset$ llavors $P(A \cup B) = P(A) + P(B)$.

El parell (Ω, \mathcal{F}) s'anomena **espai probabilitzable**. P és una mesura (3) positiva (2) i normalitzada (1). Quan $A \cap B = \emptyset$ diem que A i B són **esdeveniments incompatibles**. En el cas que \mathcal{F} sigui una σ -àlgebra substituïm (3) per

$$(3') \text{ si } A_i \in \mathcal{F}, i \in \mathbb{N} \text{ i } A_i \cap A_j = \emptyset \text{ per a tot } i \neq j \text{ llavors } P\left(\bigcup_{i \in \mathbb{N}} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

on el sumatori s'entén com una sèrie de nombres reals positius. Les propietats (3) i (3') s'anomenen **additivitat** i **σ -additivitat** respectivament.

Com a conseqüència dels axiomes (1), (2) i (3) deduïm les següents propietats:

- (4) $P(\emptyset) = 0$,
- (5) $P(\overline{A}) = 1 - P(A)$ per a tot $A \in \mathcal{F}$,
- (6) $P(A) \leq 1$ per a tot $A \in \mathcal{F}$,
- (7) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ per a tot $A, B \in \mathcal{F}$.

DEM:

- (4) Tenim que $\emptyset \cup \emptyset = \emptyset$ i $\emptyset \cap \emptyset = \emptyset$. Utilitzant (3) $P(\emptyset) = P(\emptyset \cup \emptyset) = P(\emptyset) + P(\emptyset)$ d'on surt (4).
- (5) Com $A \cap \overline{A} = \emptyset$, $P(A) + P(\overline{A}) = P(A \cup \overline{A}) = P(\Omega) = 1$.
- (6) $P(A) = 1 - P(\overline{A}) \leq 1$ degut a (2).
- (7) Considerem les descomposicions en parts disjunes $A \cup B = A \cup (B \cap \overline{A})$, $B = (B \cap A) \cup (B \cap \overline{A})$. D'aquí $P(A \cup B) = P(A) + P(B \cap \overline{A})$ i $P(B) = P(B \cap A) + P(B \cap \overline{A})$. Restant les dues igualtats trobem $P(A \cup B) - P(B) = P(A) - P(B \cap A)$ que és (7). ♣

1.6 Espais de probabilitat discrets

Considerem $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ finit. Anomenem esdeveniments elementals als de la forma $\{\omega_i\}$, $i = 1, \dots, n$ i posem $p_i = P(\{\omega_i\})$. Aquests nombres verifiquen $0 \leq p_i \leq 1$. Com $\Omega = \bigcup_{i=1}^n \{\omega_i\}$ i els esdeveniments elementals són disjunts, $\sum_{i=1}^n p_i = 1$. Les probabilitats elementals p_i determinen la de

qualsevol esdeveniment. En efecte, si $A = \{\omega_{i_1}, \dots, \omega_{i_r}\}$ llavors $P(A) = P(\bigcup_{k=1}^r \{\omega_{i_k}\}) = \sum_{k=1}^r p_{i_k}$.

Encara que en aquest context és inevitable dir que p_i és la probabilitat del resultat ω_i , s'ha de tenir en compte que les probabilitats són sempre d'esdeveniments (subconjunts de Ω .) No és correcte assignar les probabilitats als elements de Ω i, si es fa, s'ha de tenir ben clar que la probabilitat de l'element ω_i vol dir la de l'esdeveniment $\{\omega_i\}$.

Un cas corrent és quan tots els esdeveniments elementals són equiprobables amb $p_i = \frac{1}{n} \forall i$. Llavors, si A és un esdeveniment de r elements $P(A) = \frac{r}{n}$ que és la **fórmula clàssica de les probabilitats**. L'expressió $\frac{r}{n}$ s'acostuma a llegir “nombre de casos favorables dividit pel nombre de casos possibles”. La situació d'esdeveniments elementals equiprobables sol anar associada a alguna simetria de l'experiment com en el cas de tirar una moneda o un dau simètrics.

Exemple 1.4 En l'experiment “Tirar dos daus” de l'exemple 1.2 quina és la probabilitat que la suma valgui 6? Utilitzant Ω_1 tots els esdeveniments són equiprobables amb probabilitat $\frac{1}{36}$. Com $A = \{(1, 5), (2, 4), (3, 3), (4, 2), (5, 1)\}$ és $P(A) = \frac{5}{36}$. Amb Ω_2 $A = \{\{1, 5\}, \{2, 4\}, \{3, 3\}\}$ però els esdeveniments elementals no són equiprobables ja que els parells diferents tenen probabilitat doble que els repetits. Ara hem de fer $P(A) = \frac{2}{36} + \frac{2}{36} + \frac{1}{36} = \frac{5}{36}$. ♦

Si Ω té un nombre infinit numerable d'elements val el mateix que en el cas finit considerant les sumes com a sèries quan els esdeveniments no siguin finits.

Exemple 1.5 Tirem una moneda fins que surt cara. Quina és la probabilitat que necessitem més de 5 tirades? Tenim que $\Omega = \{O, +O, ++O, +++O, \dots\}$. La probabilitat d'una configuració de cares i creus donada en una sèrie de n tirades és $\frac{1}{2^n}$. Comprovem que les probabilitats elementals sumen 1.

$$\sum_{n=1}^{\infty} \frac{1}{2^n} = \frac{1/2}{1 - 1/2} = 1. \text{ La probabilitat que volem calcular és } P(n > 5) = \sum_{n=6}^{\infty} \frac{1}{2^n} = \frac{1}{2^5} \sum_{n=1}^{\infty} \frac{1}{2^n} = \frac{1}{2^5}. \text{ ♦}$$

1.7 Espais de probabilitat continus

Considerem Ω amb cardinal infinit no numerable. Per exemple, quan $\Omega = \mathbb{R}$ o de manera més general quan Ω és un subconjunt de \mathbb{R}^n . Per poder definir funcions probabilitat cal considerar una σ -àlgebra apropiada. En el cas de \mathbb{R} definim \mathcal{B} com la σ -àlgebra generada per tots els intervals de la forma $[a, b]$ fent complementaris i unions numerables. \mathcal{B} s'anomena σ -àlgebra de Borel i els seus elements borelians.

\mathcal{B} conté intervals oberts, semirectes tancades i obertes, punts, conjunts numerables de punts, etc. Per exemple $[a, \infty) = \bigcup_{n \in \mathbb{N}} [a, a + n]$, $(a, b) = \bigcup_{n \in \mathbb{N}} [a + \frac{1}{n}, b - \frac{1}{n}]$. Per tant, si seleccionem un nombre real X a l'atzar té sentit plantejar $P(a \leq X \leq b)$, $P(X \text{ és racional})$, etc.

El cas de \mathbb{R}^n es tracta de manera similar generant la σ -àlgebra \mathcal{B}_n a partir dels conjunts de la forma $[a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$.

El significat precís d'àlgebra generada per una família \mathcal{C} de subconjunts és el següent. És fàcil demostrar que la intersecció de dues àlgebres és una àlgebra. La intersecció de totes les àlgebres que contenen \mathcal{C} ens dona l'àlgebra més petita \mathcal{F} que conté \mathcal{C} . \mathcal{F} s'anomena àlgebra generada per \mathcal{C} .

1.8 Esdeveniments independents

Diem que dos esdeveniments A i B són independents si $P(A \cap B) = P(A)P(B)$.

En la pràctica, solem saber que els esdeveniments són independents degut a la mecànica que els genera i utilitzem la relació anterior per calcular probabilitats. Per exemple, al tirar dos daus, el resultat d'un d'ells és independent del de l'altre. Així, $P(\text{"primer dau parell"} \text{ i "segon dau} = 2") = P(\text{"primer dau parell"})P(\text{"segon dau} = 2") = \frac{1}{12}$.

La independència de més de dos esdeveniments correspon a la factorització total (per exemple $P(A \cap B \cap C) = P(A)P(B)P(C)$) i, en general, no equival a la independència d'aquests esdeveniments dos a dos.

1.9 Probabilitat condicionada

El fet de tenir alguna informació sobre el resultat d'un experiment modifica les probabilitats. Al tirar un dau, $P(1) = \frac{1}{6}$. Ara bé, si sabem que el resultat ha estat parell, aquesta probabilitat seria 0 i si sabem que el resultat ha estat senar, aquesta probabilitat seria $\frac{1}{3}$.

Donats dos esdeveniments A i B amb $P(B) \neq 0$, la probabilitat de A condicionada a que s'ha verificat B és

$$P(A | B) = \frac{P(A \cap B)}{P(B)}. \quad (1.1)$$

En el cas que A i B siguin independents tenim que $P(A | B) = P(A)$.

A vegades és més fàcil conèixer la probabilitat condicionada i utilitzar-la per calcular $P(A \cap B) = P(A | B)P(B)$. És a dir, per que passin A i B ha de passar B i, sabent que ha passat B , ha de passar A .

Les probabilitats condicionades verifiquen les propietats de les probabilitats ordinàries. Per exemple, $P(\overline{A} | B) = 1 - P(A | B)$.

El fet de condicionar es pot veure com una *reducció de l'espai mostral*. Al saber que s'ha produït l'esdeveniment B restem en una situació on l'espai mostral passa de ser el conjunt Ω a un de més petit B . Així, qualsevol esdeveniment A queda reduït, en a aquest nou espai, a $A \cap B$. El factor $\frac{1}{P(B)}$ en (1.1) és una constant que normalitza les probabilitats en aquest nou espai (de la mateixa manera que $P(\Omega) = 1$ tenim que $P(B|B) = 1$).

Exemple 1.6 La pel·lícula “Misió improbable” va ser vista pel 3% dels barcelonins mentre que la segona part “Misió improbable II” ho va ser pel 2%. Sabem que el 60% dels que van veure la segona part havien vist la primera. Quina és la probabilitat que una persona triada a l'atzar hagi vist les dues pel·lícules? I si sabem que aquesta persona ha vist la primera part? I si sabem que aquesta persona n'ha vist alguna?

Denotem A_1, A_2 els esdeveniments corresponents a que aquesta persona hagi vist la primera i segona part de la pel·lícula. Llavors $P(A_1 \cap A_2) = P(A_1 | A_2)P(A_2) = 0,6 \cdot 0,02 = 0,012$.

El segon cas és $P(A_2 | A_1) = \frac{P(A_1 \cap A_2)}{P(A_1)} = \frac{0,012}{0,03} = 0,4$.

El tercer cas és $P(A_1 \cap A_2 \mid A_1 \cup A_2) = \frac{P((A_1 \cap A_2) \cap (A_1 \cup A_2))}{P(A_1 \cup A_2)} = \frac{P(A_1 \cap A_2)}{P(A_1 \cup A_2)}$. Com que $P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2) = 0,03 + 0,02 - 0,012 = 0,038$, és $P(A_1 \cap A_2 \mid A_1 \cup A_2) = \frac{0,012}{0,038} = 0,3157$. ♦

1.10 Teorema de Bayes

Considerem A_1, A_2, \dots, A_n una partició de Ω . És a dir, $A_i \in \mathcal{F}, i = 1, \dots, n, \bigcup_{i=1}^n A_i = \Omega, A_i \cap A_j = \emptyset \forall i \neq j$. Llavors val la fórmula de Bayes

$$P(A_i \mid B) = \frac{P(B \mid A_i)P(A_i)}{\sum_{k=1}^n P(B \mid A_k)P(A_k)}. \quad (1.2)$$

DEM: Com la intersecció és commutativa ($P(A_i \cap B) = P(B \cap A_i)$) tenim que $P(A_i \mid B)P(B) = P(B \mid A_i)P(A_i)$ d'on trobem $P(A_i \mid B) = \frac{P(B \mid A_i)P(A_i)}{P(B)}$ que en el cas que coneguem $P(B)$ és pot utilitzar directament. Si no, posem $B = B \cap \Omega = B \cap (A_1 \cup A_2 \cup \dots \cup A_n) = (B \cap A_1) \cup (B \cap A_2) \cup \dots \cup (B \cap A_n)$. Al ser els esdeveniments $B \cap A_i$ incompatibles $P(B) = P(B \cap A_1) + P(B \cap A_2) + \dots + P(B \cap A_n) = P(B \mid A_1)P(A_1) + P(B \mid A_2)P(A_2) + \dots + P(B \mid A_n)P(A_n)$ que és el denominador de la fórmula de Bayes i ens dóna el resultat conegut com a **fórmula de la probabilitat total**:

$$P(B) = \sum_{k=1}^n P(B \mid A_k)P(A_k). \quad (1.3)$$



Exemple 1.7 Tenim 6 urnes $U_k, k = 1, \dots, 6$ contenint U_k k boles blanques i $6 - k$ boles negres. Es tira un dau i s'extreu una bola a l'atzar de la urna corresponent al resultat del dau. Si la bola és blanca, quina és la probabilitat que el dau hagués tret 3?

Indiquem per D el resultat del dau i B l'esdeveniment “extreure bola blanca”. Llavors

$$P(D = 3 \mid B) = \frac{P(B \mid D = 3)P(D = 3)}{\sum_{k=1}^6 P(B \mid D = k)P(D = k)} = \frac{\frac{3}{6} \frac{1}{6}}{\sum_{k=1}^6 \frac{k}{6} \frac{1}{6}} = \frac{3}{1 + 2 + 3 + 4 + 5 + 6} = \frac{1}{7}. \quad \spadesuit$$

1.11 Espai producte. Proves repetides

Suposem que fem dos experiments independents \mathcal{E}_1 i \mathcal{E}_2 . Podem dir també que hem fet un experiment \mathcal{E} el resultat del qual consta dels resultats de cada experiment i escrivim $\mathcal{E} = \mathcal{E}_1 \times \mathcal{E}_2$. Si els dos experiments tenen associades estructures d'espai de probabilitat amb espais mostrals Ω_1 i Ω_2 i funcions probabilitat P_1 i P_2 , l'experiment \mathcal{E} té associat l'espai producte cartesià $\Omega = \Omega_1 \times \Omega_2$. Per esdeveniments de la forma $A = A_1 \times A_2$ tenim $P(A) = P_1(A_1)P_2(A_2)$ ja que els experiments 1 i 2 són independents. Com aquest tipus d'esdeveniments donen tots els possibles fent unions i complementaris, queda determinada la probabilitat P per \mathcal{E} .

Això ho podem estendre al cas de n experiments. En particular podem considerar que cada experiment consisteix en una prova d'un experiment \mathcal{E} . El resultat de les n proves constitueix l'experiment $\mathcal{E}^n = \mathcal{E} \times \dots \times \mathcal{E}$ (n vegades).

Exemple 1.8 Tirem un dau 10 vegades. Quina és la probabilitat d'obtenir el mateix resultat en tots?

L'esdeveniment no és producte cartesià d'esdeveniments individuals però si A = “mateix resultat en tots” i A_i = “tots donen i ” $i = 1, \dots, 6$ tenim que $A = \bigcap_{i=1}^6 A_i$ i que $A_i = \{i\} \times \dots \times \{i\}$. Així si P_1

és la probabilitat d'un resultat en una prova llavors $P(A_i) = P_1(i)^{10} = \frac{1}{6^{10}}$ i $P(A) = \sum_{i=1}^6 P(A_i) = 6 \cdot \frac{1}{6^{10}} = \frac{1}{6^9}$. ♦

1.12 Significat de la probabilitat

Fins ara hem descrit la probabilitat de forma axiomàtica. Un espai de probabilitat és essencialment una estructura matemàtica que ens assigna una mesura a un conjunt de proposicions lògiques. Per veure el significat que té aquesta mesura en un experiment real cal fer unes consideracions prèvies.

Distingim primer dos tipus de processos físics o naturals.

- **Deterministes.** Les condicions del sistema determinen completament el seu comportament. Si repetim l'experiment en les mateixes condicions sempre obtenim el mateix resultat. Per exemple, mesurar amb quina velocitat arriba a terra una pedra deixada caure des de 1m d'alçada.
- **Aleatoris.** Fixades les condicions, el sistema admet més d'un comportament possible. Al repetir l'experiment anem obtenint resultats diferents. Per exemple, tirar un dau.

En quina categoria cauen els fenòmens reals? De fet no ho sabem. Les descripcions físiques fonamentals són deterministes ja que les mecàniques newtoniana, relativista o quàntica estan regides per equacions diferencials amb el que les condicions inicials determinen unívocament l'evolució del sistema. Per molts científics el determinisme és una propietat que una descripció matemàtica fonamental de la natura hauria de contenir. Per una altra part, les descripcions físiques sempre són aproximades i tenen un caràcter provisional fins que n'aparegui una de millor. Es podria discutir també que vol dir "les mateixes condicions" ja que l'estat de l'univers no és mai el mateix. En general entenem que el resultat d'un experiment és afectat només per l'entorn immediat i podem negligir l'efecte dels objectes llunyans.

En l'ordre pràctic podem deixar de banda les qüestions filosòfiques ja que al marge de que la mecànica subjacent sigui determinista o no molts fenòmens reals presenten un caràcter clarament aleatori. Un cas en que passa això és quan el sistema és molt complex. En aquests casos serà impossible disposar d'informació completa sobre el sistema i haurem de recórrer a informació de tipus estadístic. Per exemple, pels sistemes macroscòpics formats per molts àtoms o molècules és impossible resoldre o tractar mínimament les equacions del moviment. Si en lloc d'això associem probabilitats als diferents estats del sistema arribem a una descripció satisfactòria (mecànica estadística). Paràmetres com la temperatura o l'entropia es poden interpretar llavors com variables que caracteritzen les funcions de probabilitat del sistema.

Per entendre com un sistema determinista pot aparentar un comportament aleatori pensem el següent experiment. Tenim una taula inclinada en la que posem un objecte en un punt determinat (condició inicial) i el deixem caure. Si la taula és llisa l'objecte baixarà seguint una línia recta. És trivial predir per on caurà si coneixem la seva posició inicial. Això és el que entenem en la pràctica per sistema determinista: aquell en que la condició inicial no només determina el comportament sinó que la relació condició inicial-comportament és prou senzilla.

Suposem ara que la taula no és llisa sinó que està plena de petits obstacles que desvien l'objecte al caure. Ara les trajectòries de caiguda són completament irregulars i és molt difícil predir el punt d'arribada a partir del punt de sortida. Notem que el sistema és encara determinista ja que l'objecte segueix trajectòries regides per les equacions de Newton però la complexitat del seu comportament fa més adequat considerar-lo com aleatori i descriure'l probabilísticament. Els sistemes d'aquesta mena s'anomenen *pseudoaleatoris*. Una tirada d'un dau n'és clarament un exemple ja que qualsevol variació per petita que sigui de les condicions inicials pot canviar el resultat.

Un altre exemple d'experiment pseudoaleatori és anar considerant com a resultats de l'experiment els successius dígitos decimals d'un nombre real entre 0 i 1. Si el nombre que els genera és, diguem, $\frac{11}{13}$ no tardarem en endevinar quin és ja que la regla que els va formant és molt simple. Però si ho fem amb el nombre π o e , tot i que existeix una regla generadora (per tant, determinista), la successió tindrà aparença purament aleatòria degut a la complexitat d'aquesta regla. Aquesta és la base de la generació de nombres "aleatoris" per part dels ordinadors que són sistemes absolutament deterministes.

Considerem ara un experiment (pseudo)aleatori i sigui A un enunciat lògic referit al resultat de l'experiment. Repetim N vegades el procés i A es verifica N_A vegades. Definim la **frequència relativa** de A com $f_A = \frac{N_A}{N}$. Llavors esperem que

$$\lim_{N \rightarrow \infty} \frac{N_A}{N} = P_A \quad (1.4)$$

on P_A és una constant que correspon a la probabilitat de A . S'ha d'entendre que el límit anterior no té cap sentit matemàtic. L'expressió diu que si fem moltes vegades l'experiment la freqüència f_A hauria de prendre valors molt propers a P_A que ve fixada per les condicions de l'experiment. Si volem axiomatitzar la probabilitat, li demanarem que tingui les propietats que tenen les freqüències relatives. De manera obvia

- (1) $0 \leq f_A \leq 1$, ja que $0 \leq N_A \leq N$,
- (2) $f_\Omega = 1$, ja que $N_\Omega = N$,
- (3) si A, B són incompatibles $f_{A \cup B} = f_A + f_B$, ja que $N_{A \cup B} = N_A + N_B$.

Aquestes propietats són precisament els axiomes que hem exigit a la funció probabilitat.

Finalment, cal tenir en compte les limitacions pràctiques en l'ús de probabilitats. No s'ha de confondre el fet que un procés sigui aleatori amb que tinguem incertesa total sobre el seu resultat. Quan tirem una moneda el resultat és incert però *sabem* que cada resultat té probabilitat $\frac{1}{2}$. A vegades el resultat d'un procés és incert i a més a més no tenim cap informació que ens permeti assignar probabilitats als possibles resultats. Seria incorrecte assignar probabilitats $\frac{1}{2}$ a un experiment amb dos resultats possibles pel fet que no sapiguem res sobre les seves circumstàncies. Quan la falta d'informació sobre el sistema es pot suplir amb un tractament probabilístic és per que aquest sistema té propietats que ho fan possible.

La definició d'experiment aleatori exigeix la possibilitat de repetir l'experiment un gran nombre de vegades sense variar les condicions. És perillós assignar probabilitats quan és impossible o molt difícil repetir l'experiment en idèntiques condicions. Per exemple, el resultat d'un partit de futbol o l'èxit o fracàs d'una missió espacial. En aquests casos les estimacions probabilístiques tenen un fort component subjectiu que fa que manquin de sentit.

Capítol 2

Variables aleatòries

2.1 Variable aleatòria

Considerem el conjunt \mathbb{R} . Ja s'ha descrit la σ -àlgebra \mathcal{B} dels borelians. En lloc de generar-la amb els intervals tancats com s'havia vist, la podem generar de manera equivalent amb les semirectes $(-\infty, a]$, $a \in \mathbb{R}$ a través també d'unions numerables i complementaris.

Donat (Ω, \mathcal{F}) espai probabilitzable, \mathcal{F} σ -àlgebra, diem que $f : \Omega \rightarrow \mathbb{R}$ és **mesurable** si $f^{-1}(A) \in \mathcal{F} \forall A \in \mathcal{B}$. És a dir, les antiimatges d'esdeveniments a \mathbb{R} (borelians) són esdeveniments de \mathcal{F} .

Una **variable aleatòria** és una aplicació $X : \Omega \rightarrow \mathbb{R}$ mesurable.

Essencialment, una variable aleatòria (v.a.) consisteix en assignar valors reals als resultats d'un experiment. Com, de manera natural, aquests resultats ja són moltes vegades expressats com a nombres reals es tendeix a confondre Ω amb la seva imatge sota l'aplicació X .

Exemple 2.1 \mathcal{E} = "Tirar una moneda", $\Omega = \{\text{cara}, \text{creu}\}$. $X(\text{cara})=1$, $X(\text{creu})=0$.♦

Exemple 2.2 \mathcal{E} = "Tirar un dau", $\Omega = \{1, 2, 3, 4, 5, 6\}$. $X(n) = n$, per $n = 1, \dots, 6$.♦

Exemple 2.3 \mathcal{E} = "Triar un instant del dia", $\Omega = [0, 24)$. $X(t) = t$, $\forall t \in \Omega$.♦

Exemple 2.4 Per un experiment donat hi ha infinitat de variables aleatòries. Per exemple, en l'experiment anterior $X(t) = 0$ si $0 \leq t < 12$ i $X(t) = 1$ si $12 \leq t \leq 24$ és una altra variable. Una altra variable és $X(t) = 1$, $\forall t$ (variable aleatòria constant).♦

Les variables aleatòries són els objectes que utilitzem quan fem un tractament probabilístic d'una situació real. En els casos pràctics l'espai mostral Ω pot tenir una enorme complexitat els detalls de la qual desconexim. Degut a això, les variables aleatòries rarament apareixen en forma funcional explícita $X(\omega)$. Això no serà important perquè podrem respondre a preguntes sobre els valors que pren X a partir de funcions que caracteritzen la variable. És convenient, però, no perdre de vista l'estructura global ($X : \Omega \rightarrow \mathbb{R}$) per evitar confusions i entendre bé els problemes.

Si tenim una probabilitat P definida a Ω , per a cada borelià A definim $P_X(A) = P(X^{-1}(A))$. Com és suficient conèixer les probabilitats dels generadors de l'àlgebra d'esdeveniments, estarem interessats en $P_X((-\infty, x])$. En el que segueix ometrem, si no hi ha confusió possible, el subíndex X en P_X .

2.2 Funció de distribució

Donat un espai de probabilitat (Ω, \mathcal{F}, P) amb variable aleatòria X definim la **funció de distribució** com

$$F_X(x) = P(X \leq x). \quad (2.1)$$

És, per tant, la probabilitat de l'interval $(-\infty, x]$. Si coneixem la funció F_X , tenim les probabilitats dels generadors de l'àlgebra \mathcal{B} i, per tant, queden determinades totes les probabilitats. En aquest sentit, F_X caracteritza completament la variable aleatòria X .

Immediatament trobem que donats $a < b$

$$P(a < X \leq b) = F_X(b) - F_X(a), \quad (2.2)$$

ja que $F_X(b) = P((-\infty, b]) = P((-\infty, a] \cup (a, b]) = P((-\infty, a]) + P((a, b]) = F_X(a) + P((a, b])$.

A més a més, F_X té les següents propietats:

- (1) $0 \leq F_X(x) \leq 1$,
- (2) F_X és una funció creixent,
- (3) $\lim_{x \rightarrow -\infty} F_X(x) = 0$, $\lim_{x \rightarrow \infty} F_X(x) = 1$,
- (4) F_X és una funció contínua per la dreta.

DEM:

(1) Pel fet de ser una probabilitat.

(2) Considerem $x_2 > x_1$. Llavors $F_X(x_2) = F_X(x_1) + P((x_1, x_2]) \geq F_X(x_1)$.

(3) El resultat és intuïtiu ja que $F_X(-\infty) = P(\emptyset) = 0$ i $F_X(\infty) = P(\mathbb{R}) = 1$. Per provar-ho, notem que $l_- = \lim_{x \rightarrow -\infty} F_X(x)$ i $l_+ = \lim_{x \rightarrow \infty} F_X(x)$ existeixen pel fet de ser F_X monòtona i fitada, i $0 \leq l_- \leq l_+ \leq 1$. Posem

$$\text{ara } 1 = P(\mathbb{R}) = P\left(\bigcup_{n=-\infty}^{\infty} (n, n+1]\right) = \sum_{n=-\infty}^{\infty} P((n, n+1]) = \sum_{n=-\infty}^{\infty} (F_X(n+1) - F_X(n)) = \lim_{M \rightarrow \infty} \sum_{n=-M}^{M-1} (F_X(n+1) - F_X(n)) = \lim_{M \rightarrow \infty} (F_X(M) - F_X(-M)) = l_+ - l_-.$$

Per tant, necessàriament $l_- = 0$ i $l_+ = 1$.

(4) Pel fet de ser F_X monòtona i fitada existeix $l = \lim_{t \rightarrow x+} F_X(t)$. La funció és contínua si $l = F_X(x)$. Per trobar el seu valor considerem una successió decreixent que tendeixi a x per la dreta. (per exemple $x_n = x + \frac{1}{n}$.) És a dir $x_n \rightarrow x$ i $x_{n-1} \geq x_n \geq x$ per a tot n . Llavors $l = \lim_{n \rightarrow \infty} F_X(x_n)$. Tenim que $F_X(x_1) - F_X(x) =$

$$P((x, x_1]) = P\left(\bigcup_{n=2}^{\infty} (x_n, x_{n-1}]\right) = \sum_{n=2}^{\infty} P((x_n, x_{n-1}]) = \sum_{n=2}^{\infty} (F_X(x_{n-1}) - F_X(x_n)) = \lim_{M \rightarrow \infty} \sum_{n=2}^M (F_X(x_{n-1}) - F_X(x_n)) = \lim_{M \rightarrow \infty} (F_X(x_1) - F_X(x_M)) = F_X(x_1) - l \text{ que ens dona } l = F_X(x). \clubsuit$$

Si F_X també és contínua per l'esquerra en el punt a tenim que

$$P(\{a\}) = \lim_{\epsilon \rightarrow 0^+} P(a - \epsilon < X \leq a) = \lim_{\epsilon \rightarrow 0^+} (F_X(a) - F_X(a - \epsilon)) = 0.$$

Per tant les discontinuïtats de F_X indiquen caràcter discret de la variable aleatòria i, com indica la fórmula anterior, la probabilitat d'un punt ve donada pel salt de F_X en aquest punt.

2.3 Variables aleatòries discretes

És quan X pren un conjunt finit o numerable de valors. Si aquests valors són, ordenats, x_1, x_2, \dots, x_n amb probabilitats p_1, p_2, \dots, p_n , F_X és una funció esglaonada que pren valors $0, p_1, p_1 + p_2, \dots, p_1 + \dots + p_{n-1}, 1$. Una expressió compacta per F_X la trobem mitjançant la funció de Heaviside $u(x)$.

$$F_X(x) = \sum_k p_k u(x - x_k). \quad (2.3)$$

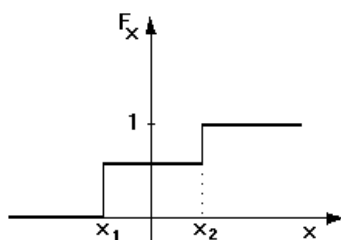


Figura 1: Funció de distribució d'una variable discreta.

Posarem $\Omega_X = \{x_1, x_2, \dots, x_n\}$. La funció en aquest conjunt definida per $P_X(x_k) = p_k$ s'anomena **funció de probabilitat** i es relaciona amb la de distribució per

$$P_X(x_k) = F_X(x_k) - F_X(x_{k-1}). \quad (2.4)$$

Les següents són algunes de les distribucions més corrents:

- **Bernoulli.**

$\Omega_X = \{0, 1\}$. $P_X(1) = p$, $P_X(0) = q = 1 - p$. Correspon per tant a un experiment amb dos resultats. Si A és un esdeveniment associat a un experiment qualsevol, podem construir una variable de Bernoulli assignant 1 a “ A es verifica” i 0 a “ A no es verifica.” Llavors $p = P(A)$.

- **Binomial.**

Repetim n vegades un experiment de Bernoulli. X és el nombre de 1's que obtenim. Així, $\Omega_X = \{0, 1, 2, \dots, n\}$ i

$$P_X(k) = \binom{n}{k} p^k q^{n-k}. \quad (2.5)$$

Exemple 2.5 Es tira un dau 10 vegades. Quina és la probabilitat que el 6 surti exactament 4 vegades?

Sigui X la variable que compta el nombre de vegades que surt el 6. X és binomial amb $n = 10$ i $p = \frac{1}{6}$. Llavors $P(X = 4) = \binom{10}{4} \left(\frac{1}{6}\right)^4 \left(\frac{5}{6}\right)^6 = 0,054$. ♦

- **Geomètrica.**

X és el nombre de vegades que hem de fer un experiment de Bernoulli fins que surt 1. Així, $\Omega_X = \{1, 2, \dots\}$ i

$$P_X(k) = q^{k-1} p. \quad (2.6)$$

Podem calcular la funció de distribució

$$F_X(n) = \sum_{k=1}^n P_X(k) = \sum_{k=1}^n q^{k-1} p = 1 - q^n. \quad (2.7)$$

Exemple 2.6 Es tira un dau fins que surt un 6. Quina és la probabilitat que calguin 5 o més tirades?

Sigui X la variable que compta el nombre de tirades que fem. X és geomètrica amb $p = \frac{1}{6}$. Llavors $P(X \geq 5) = 1 - P(X \leq 4) = 1 - F_X(4) = q^4 = (\frac{5}{6})^4 = 0,482$. ♦

- **Hipergeomètrica.**

Tenim una urna amb b boles blanques i $N - b$ boles negres. En traiem n . X és el nombre de boles blanques que traiem. $\Omega_X = \{0, 1, 2, \dots, n\}$ i

$$P_X(k) = \frac{\binom{b}{k} \binom{N-b}{n-k}}{\binom{N}{n}}. \quad (2.8)$$

- **Poisson.**

Fixem un interval de temps T . X és el nombre de vegades que es produeix un esdeveniment en aquest interval. Per exemple, comptar el nombre de trucades telefòniques rebudes durant 1h per una centraleta. La variable de Poisson es pot referir també a intervals espacials o d'una altra mena (el nombre de clots en un tram de carretera de 10 Km.)

$\Omega_X = \{0, 1, 2, \dots\}$. La variable és de Poisson si

$$P_X(k) = e^{-\alpha} \frac{\alpha^k}{k!}. \quad (2.9)$$

Podem justificar aquesta expressió i relacionar aquesta distribució amb la binomial. Per això, dividim l'interval T en n subinterval de longitud $\Delta t = \frac{T}{n}$. Considerem n molt gran i, per tant, Δt molt petit. Suposem que en cada subinterval l'esdeveniment pot produir-se com a màxim una vegada amb probabilitat $p = \lambda \Delta t$. Ara podem considerar que tenim un experiment binomial i fer el límit $n \rightarrow \infty$:

$$\begin{aligned} P_X(k) &= \binom{n}{k} p^k (1-p)^{n-k} = \binom{n}{k} \left(\frac{\lambda T}{n} \right)^k \left(1 - \frac{\lambda T}{n} \right)^{n-k} \\ &= \frac{1}{k!} \frac{n-1}{n} \cdot \frac{n-2}{n} \cdot \dots \cdot \frac{n-k+1}{n} (\lambda T)^k \frac{(1 - \frac{\lambda T}{n})^n}{(1 - \frac{\lambda T}{n})^k} \longrightarrow \frac{1}{k!} (\lambda T)^k e^{-\lambda T} \end{aligned}$$

que és la distribució de Poisson amb $\alpha = \lambda T$. La constant λ depèn només del tipus de fenomen així que podem relacionar els valors de α per intervals de diferent longitud.

Exemple 2.7 La probabilitat que hi hagi una o més connexions a un servidor durant un segon val 0,07. Quina és la probabilitat P que hi hagin cinc o més connexions durant un minut?

Sigui α_s el paràmetre de Poisson per $T = 1s$ i α_m el paràmetre de Poisson per $T = 60s$. Llavors $\alpha_m = 60\alpha_s$. Sabem que $0,07 = 1 - P_{\alpha_s}(0) = 1 - e^{-\alpha_s}$ d'on $\alpha_s = 0,0725$. Llavors $\alpha_m = 4,354$ i $P = 1 - P_{\alpha_m}(0) - P_{\alpha_m}(1) - P_{\alpha_m}(2) - P_{\alpha_m}(3) - P_{\alpha_m}(4) = 1 - e^{-\alpha_m} \left(1 + \alpha_m + \frac{\alpha_m^2}{2} + \frac{\alpha_m^3}{3!} + \frac{\alpha_m^4}{4!} \right) = 0,44$. ♦

Notem que hem demostrat la relació asimptòtica entre les distribucions binomial amb paràmetres n i p i de Poisson amb paràmetre $\alpha = np$ quan n és gran. Això és útil en alguns càlculs.

2.4 Variables aleatòries contínues

És quan F_X és contínua i derivable. Llavors existeix la funció

$$f_X(x) = \frac{dF_X(x)}{dx}. \quad (2.10)$$

f_X s'anomena **funció de densitat** de la v.a. X . Recordem que si F_X és contínua $P(\{a\}) = 0$. Per tant, la probabilitat d'un interval no varia incloquem o no els seus extrems.

f_X verifica les propietats:

- (1) $f_X \geq 0$,
- (2) $P(a \leq X \leq b) = \int_a^b f_X(x)dx$,
- (3) $F_X(x) = \int_{-\infty}^x f_X(x')dx'$,
- (4) $\int_{-\infty}^{\infty} f_X(x)dx = 1$.

DEM: (1) Ja que F_X és creixent. (2) $P(a \leq X \leq b) = F_X(b) - F_X(a) = \int_a^b F'_X(x)dx = \int_a^b f_X(x)dx$. (3) Fem $a \rightarrow -\infty$ en (2). (4) Fem $b \rightarrow \infty$ en (3). ♣

Una manera més acurada de definir variable aleatòria contínua és dir que és quan existeix una funció f_X tal que es verifica l'anterior propietat (3). Això implica (2.10) per derivació però de fet f_X pot no ser contínua amb el que F_X pot no ser derivable en alguns punts.

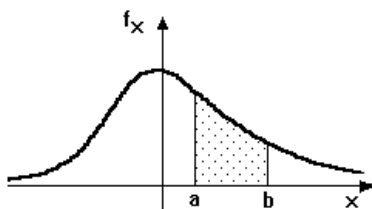


Figura 2: Probabilitat com àrea sota f_X .

Donada qualsevol funció $\varphi \geq 0$ contínua a trossos tal que $\int_{-\infty}^{\infty} \varphi(x)dx = A$ finit, la funció $A^{-1}\varphi(x)$ defineix una densitat de probabilitat. A vegades es parteix de la definició de la funció de densitat exigint (1),(2) i (4) per definir després la funció de distribució a través de (3). Notem que $f_X(x)$ no és una probabilitat i per tant pot prendre valors arbitràriament elevats. En particular, és possible que f_X valgui ∞ en algun punt. Aquestes divergències són admissibles mentre la funció sigui integrable. Una definició o manera d'expressar la funció de densitat és

$$f_X(x)dx = P(x \leq X \leq x + dx) \quad (2.11)$$

que simbolitza (2) a nivell infinitesimal.

Les variables discretes admeten una funció de densitat en sentit distribucional utilitzant la δ de Dirac:

$$f_X(x) = \sum_k p_k \delta(x - x_k). \quad (2.12)$$

Aquesta forma s'obté derivant la funció de distribució (2.3) i tenint en compte que $u' = \delta$.

Denotarem Ω_X el recorregut de la v.a. X . Aquest és el conjunt de \mathbb{R} que no conté cap interval on $f_X(x) = 0$. De fet, no hi ha cap inconvenient en afegir a Ω_X qualsevol regió on f_X s'anul·li. En la pràctica, Ω_X és el conjunt de valors que X pren de manera natural. (Si, per exemple, X representa una distància seria $\Omega_X = [0, \infty)$, etc.)

Les variables contínues més corrents són:

- **Uniforme.**

$$\Omega_X = [a, b].$$

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{si } a \leq x \leq b \\ 0 & \text{altrament} \end{cases} \quad (2.13)$$

$$F_X(x) = \begin{cases} 0 & \text{si } x \leq a \\ \frac{x-a}{b-a} & \text{si } a \leq x \leq b \\ 1 & \text{si } x \geq b \end{cases} \quad (2.14)$$

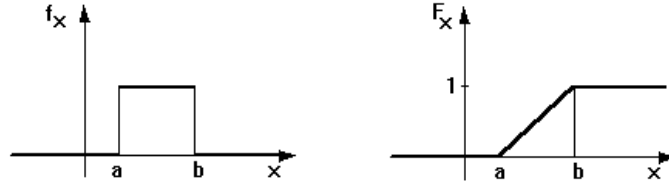


Figura 3: Densitat i distribució uniformes.

- **Exponencial.**

$$\Omega_X = [0, \infty), \lambda > 0.$$

$$f_X(x) = \begin{cases} 0 & \text{si } x < 0 \\ \lambda e^{-\lambda x} & \text{si } x \geq 0 \end{cases} \quad (2.15)$$

$$F_X(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - e^{-\lambda x} & \text{si } x \geq 0 \end{cases} \quad (2.16)$$

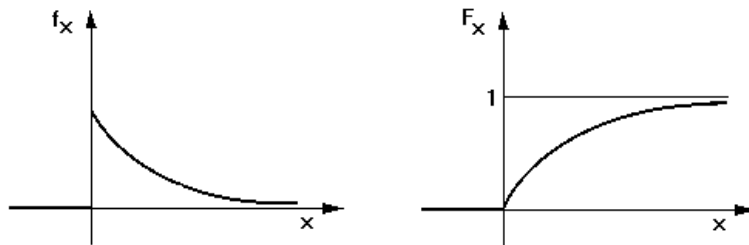


Figura 4: Densitat i distribució exponencials.

Una possible interpretació d'aquesta variable és que deixem transcórrer el temps des de 0 i X és l'instant en que es produeix per primera vegada un esdeveniment. Per exemple el moment en que rebem una trucada telefònica. Sota aquesta interpretació podem relacionar

la variable exponencial amb la geomètrica. Dividim l'interval $[0, x]$ en n subintervalls de longitud $\Delta t = \frac{x}{n}$. Considerem n molt gran i, per tant, Δt molt petit. Suposem que en cada subinterval l'esdeveniment pot produir-se amb probabilitat $p = \lambda \Delta t$. Ara podem considerar que tenim un experiment geomètric i, utilitzant (2.7),

$$1 - F_X(x) = P(X > x) = (1 - p)^n = \left(1 - \frac{\lambda x}{n}\right)^n \longrightarrow e^{-\lambda x}$$

el que ens dona la funció de distribució (2.16).

Si pensem X com el temps que passa entre esdeveniments en la distribució de Poisson tenim doncs que aquest és exponencial de paràmetre $\lambda = \frac{\alpha}{T}$.

Una característica de la distribució exponencial és la **propietat de no memòria**:

$$P(X > t + h \mid X > t) = P(X > h). \quad (2.17)$$

DEM: Com $P(X > x) = e^{-\lambda x}$, $P(X > t + h) = P(X > t)P(X > h)$. Per una altra part $P(X > t + h \mid X > t) = \frac{P(X > t + h)}{P(X > t)}$. ♣

Aquesta propietat vol dir que la probabilitat de produir-se l'esdeveniment en un instant donat és independent del nombre de vegades que s'hagi produït anteriorment.

En ocasions, s'utilitza la variable exponencial per indicar el temps de vida d'un dispositiu, organisme, etc. Tot i que en algun cas això és correcte (desintegració radioactiva, temps de vida d'una partícula inestable) en general és una primera aproximació que pot donar resultats paradoxals. Per exemple, suposem que el temps que viu una persona és una variable exponencial T . Si fem $t = 70, h = 10$ en (2.17) obtenim $P(T > 80 \mid T > 70) = P(T > 10)$. Això vol dir que una persona que als 70 anys està viva té la mateixa probabilitat de seguir viva als 80 que la que té una persona acabada de néixer de seguir viva als 10 anys. Això no pot ser correcte i el fet és que les persones són organismes que es van desgastant mentre que la propietat de no memòria descriuria un sistema sense desgast.

• Gaussiana.

També s'anomena normal. $\Omega_X = (-\infty, \infty)$, $m \in \mathbb{R}$, $\sigma > 0$.

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}. \quad (2.18)$$

$$F_X(x) = \frac{1}{2} \left(1 + \operatorname{erf} \left(\frac{x-m}{\sqrt{2}\sigma} \right) \right). \quad (2.19)$$

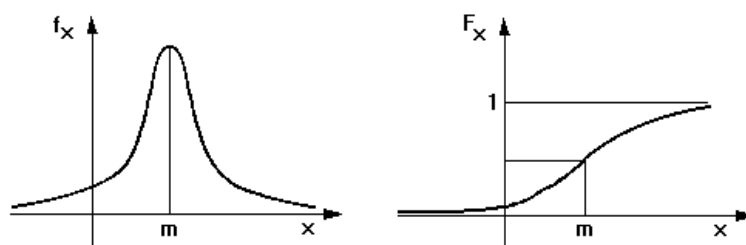


Figura 5: Densitat i distribució normals.

En 2.19 utilitzem la **funció d'error**, definida:

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt. \quad (2.20)$$

(Notem que $\operatorname{erf}(-x) = -\operatorname{erf}(x)$. Al final d'aquestes notes hi ha una taula amb els seus valors.)

DEM: En efecte, $F_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x e^{-\frac{(t-m)^2}{2\sigma^2}} dt$. Fent el canvi $z = \frac{t-m}{\sqrt{2}\sigma}$, és

$$\begin{aligned} F_X(x) &= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\frac{x-m}{\sqrt{2}\sigma}} e^{-z^2} dz = \frac{1}{\sqrt{\pi}} \left(\int_{-\infty}^0 e^{-z^2} dz + \int_0^{\frac{x-m}{\sqrt{2}\sigma}} e^{-z^2} dz \right) \\ &= \frac{1}{\sqrt{\pi}} \left(\frac{\sqrt{\pi}}{2} + \frac{\sqrt{\pi}}{2} \operatorname{erf}\left(\frac{x-m}{\sqrt{2}\sigma}\right) \right) = \frac{1}{2} \left(1 + \operatorname{erf}\left(\frac{x-m}{\sqrt{2}\sigma}\right) \right). \clubsuit \end{aligned}$$

Una funció gaussiana és simplement l'exponencial d'un polinomi de segon grau. La densitat gaussiana és la forma general d'una densitat amb aquest comportament. La seva gràfica mostra el típic aspecte de moltes distribucions que es mesuren experimentalment en estadística i s'anomena *campana de Gauss*.

- **Cauchy.**

$$\Omega_X = (-\infty, \infty), \alpha > 0$$

$$f_X(x) = \frac{\alpha/\pi}{x^2 + \alpha^2}. \quad (2.21)$$

$$F_X(x) = \frac{1}{2} + \frac{1}{\pi} \arctan \frac{x}{\alpha}. \quad (2.22)$$

Exemple 2.8 Sigui X una variable aleatòria i A l'esdeveniment $X^3 - 2X^2 - X + 2 < 0$. Calcular $P(A)$ pels casos en que X és (a) Uniforme en $[-2, 2]$, (b) Exponencial $\lambda = 3$, (c) Normal $m = 0, \sigma = 1$ i (d) Cauchy $\alpha = 1$.

Primer posem l'esdeveniment en termes d'interval: $A = J_1 \cup J_2$ on $J_1 = (-\infty, -1)$, $J_2 = (1, 2)$. Així $P(A) = P(J_1) + P(J_2)$.

(a) És fàcil verificar que, per una variable uniforme en I la probabilitat d'un interval J contingut en I val $\frac{\text{longitud}(J)}{\text{longitud}(I)}$. Així $P(J_1) = P(J_2) = \frac{1}{4}$ d'on $P(A) = \frac{1}{2}$.

(b) És $F(x) = 1 - e^{-3x}$. $P(J_1) = 0$ i $P(J_2) = F(2) - F(1) = e^{-3} - e^{-3 \cdot 2} = 0,047$.

(c) És $F(x) = \frac{1 + \operatorname{erf}(x/\sqrt{2})}{2}$. $P(J_1) = F(-1) = \frac{1 - \operatorname{erf}(1/\sqrt{2})}{2} = 0,1586$, $P(J_2) = F(2) - F(1) = \frac{\operatorname{erf}(2/\sqrt{2}) - \operatorname{erf}(1/\sqrt{2})}{2} = 0,1359$, $P(A) = 0,294$.

(d) És $F(x) = \frac{1}{2} + \frac{1}{\pi} \arctan x$. $P(J_1) = F(-1) = \frac{1}{4}$, $P(J_2) = F(2) - F(1) = \frac{\arctan 2 - \arctan 1}{\pi} = 0,1024$, $P(A) = 0,352$. ♦

Algunes variables tenen comportament continu i discret alhora. S'anomenen **variables mixtes**. Es donen quan F_X té discontinuïtats i no és constant a trossos.

Exemple 2.9 Tirem una moneda. Si surt cara X val 2. Si surt creu X es tria uniformement en l'interval $[0, 1]$. Llavors (veure (2.28))

$$f_X(x) = \frac{1}{2}(u(x) - u(x-1)) + \frac{1}{2}\delta(x-2). \blacklozenge$$

Quan una v.a. discreta concentra les seves probabilitats en valors molt elevats sol ser més pràctic tractar-la com si fos contínua. Això passa, per exemple, a l'estudiar la estadística de les poblacions humanes. Alguns resultats estableixen relacions útils entre distribucions discretes i contínues, com ja hem vist al cas exponencial-geomètrica o com el següent resultat.

2.5 Teorema de DeMoivre-Laplace

Aquest teorema, que demostrarem al final del capítol, estableix una relació entre la distribució binomial i la normal. Diu que si $npq \gg 1$

$$\binom{n}{k} p^k q^{n-k} \simeq \frac{1}{\sqrt{2\pi npq}} e^{-\frac{(k-np)^2}{2npq}}. \quad (2.23)$$

Més precisament, el quocient dels dos costats tendeix a 1 quan $npq \rightarrow \infty$. El resultat ens diu que per n molt gran (en la pràctica $n > 10$) la distribució binomial de paràmetres n, p és correspon a la gaussiana de paràmetres $m = np, \sigma^2 = npq$.

La correspondència entre la binomial i la normal s'utilitza per calcular probabilitats associades a la binomial. En lloc de fer els sumatoris pels quals no tenim cap fórmula compacta, avaluem les probabilitats utilitzant 2.19. Si X és la variable binomial i \hat{X} la seva aproximació normal, identifiquem $X = k$ amb $k - \frac{1}{2} < \hat{X} \leq k + \frac{1}{2}$ de manera que $P(k_1 \leq X \leq k_2) \simeq P(k_1 - \frac{1}{2} < \hat{X} \leq k_2 + \frac{1}{2})$, $P(X \geq k) \simeq P(\hat{X} > k - \frac{1}{2})$ i $P(X \leq k) \simeq P(\hat{X} \leq k + \frac{1}{2})$.

Exemple 2.10 Un monitor té 10×15 píxels. Cada píxel s'il·lumina amb probabilitat $p = 0,4$. Quina és la probabilitat P que hi hagin més de 80 píxels il·luminats?

Com no tenim una expressió compacta per la funció de distribució binomial, seria

$$P = \sum_{k=81}^{150} \binom{150}{k} 0,4^k 0,6^{150-k}.$$

En lloc de fer la suma utilitzem l'aproximació normal amb $m = 150 \cdot 0,4 = 60$ i $\sigma^2 = 150 \cdot 0,4 \cdot 0,6 = 36$. $P = 1 - F(80,5) = \frac{1 - \text{erf}(\frac{24159}{2})}{2} = 3,17 \cdot 10^{-4}$ (el resultat exacte és $3,64 \cdot 10^{-4}$).♦

En aquest exemple veiem l'aplicació de la distribució normal a la descripció del soroll. Si aquest soroll està produït per la suma de petites fluctuacions aleatòries independents, pren forma gaussiana i s'anomena soroll blanc.

2.6 Funcions de densitat i distribució condicionades

Com qualsevol probabilitat, la funció de distribució condicionada a que l'esdeveniment B s'ha produït és

$$F(x | B) = P(X \leq x | B) = \frac{P(X \leq x \text{ i } B)}{P(B)}. \quad (2.24)$$

En el cas de distribucions contínues la funció de densitat condicionada és

$$f(x | B) = \frac{dF(x | B)}{dx}. \quad (2.25)$$

Aquestes funcions tenen les mateixes propietats que les funcions sense condicionar. En particular, F va de 0 a 1 i f està normalitzada a 1.

Tenim llavors la forma contínua del teorema de Bayes:

$$f(x | B) = \frac{P(B | X = x) f(x)}{\int_{-\infty}^{\infty} P(B | X = x') f(x') dx'}. \quad (2.26)$$

DEM: Per la formula de Bayes habitual

$$P(x \leq X \leq x + \Delta x | B) P(B) = P(B | x \leq X \leq x + \Delta x) P(x \leq X \leq x + \Delta x),$$

és a dir,

$$(F(x + \Delta x | B) - F(x | B))P(B) = P(B | x \leq X \leq x + \Delta x)(F(x + \Delta x) - F(x)).$$

Si dividim els dos costats per Δx i fem el limit $\Delta x \rightarrow 0$ trobem

$$f(x | B)P(B) = P(B | X = x)f(x).$$

Integrant els dos costats entre $-\infty$ i ∞ obtenim el denominador en (2.26):

$$P(B) = \int_{-\infty}^{\infty} P(B | X = x)f(x)dx \quad (2.27)$$

que substituït en l'anterior expressió és el que volíem demostrar. ♣

La fórmula (2.27) és la versió contínua de la fórmula de la probabilitat total. Al marge de la fórmula de Bayes, és pot utilitzar per calcular la probabilitat en casos on l'esdeveniment ve condicionat per alguna variable contínua.

Exemple 2.11 Tenim una col·lecció gran de monedes. La probabilitat de treure cara varia segons la moneda, de manera que si en triem una a l'atzar la probabilitat p que tregui cara es pot considerar una variable uniforme en $[0, 1]$. Extraïem una moneda i la tirem tres vegades. Quina és la probabilitat que surtin exactament dues cares?

Segui X la probabilitat de treure cara de la moneda triada i N el nombre de cares que surten. $P(N = k | X = p) = \binom{3}{k}p^k(1-p)^{3-k}$ i $f_X(p) = 1$ per a $0 \leq p \leq 1$. Llavors:

$$P(N = 2) = \int_{-\infty}^{\infty} P(N = 2 | X = p)f_X(p)dp = \int_0^1 \binom{3}{2}p^2(1-p) \cdot 1dp = 3 \int_0^1 (p^2 - p^3)dp = \frac{1}{4}. \blacklozenge$$

Una nova versió de la fórmula de la probabilitat total (en aquest cas, de la densitat total) s'obté considerant una partició de l'espai mostral $\{A_i\}_{i=1, \dots, n}$. Utilitzant (1.3) tenim que $F(x) = \sum_{i=1}^n F(x | A_i)P(A_i)$ d'on, derivant, traiem

$$f(x) = \sum_{i=1}^n f(x | A_i)P(A_i). \quad (2.28)$$

Exemple 2.12 Un sistema disposa de dos processadors \mathcal{P}_1 i \mathcal{P}_2 on el temps que dura un cert servei és una variable exponencial de paràmetre $\lambda = 2$ i $\lambda = 1$ respectivament. Als usuaris se'ls assigna un dels dos processadors aleatòriament, de manera que la probabilitat de triar \mathcal{P}_1 val $\frac{1}{3}$. Quina és la densitat del temps de servei? Si T és el temps de servei, per $t > 0$:

$$f_T(t) = f_T(t | \mathcal{P}_1)P(\mathcal{P}_1) + f_T(t | \mathcal{P}_2)P(\mathcal{P}_2) = 2e^{-2t}\frac{1}{3} + e^{-t}\frac{2}{3} = \frac{2}{3}(e^{-2t} + e^{-t}). \blacklozenge$$

La fórmula de Bayes adquireix una forma més definida si l'esdeveniment B és de la forma $X \in A$ on A és un borelià de \mathbb{R} . Si X és una variable amb densitat $f(x)$:

$$f(x | X \in A) = \begin{cases} \frac{1}{P(A)}f(x) & \text{si } x \in A \\ 0 & \text{altrement} \end{cases} \quad (2.29)$$

ja que $P(X \in A | X = x)$ val 1 si $x \in A$ i 0 en cas contrari.

Apèndix: Demostració del teorema de DeMoivre-Laplace

Volem una expressió asimptòtica de $p_n(k) = \binom{n}{k} p^k q^{n-k}$. Per això fem el canvi de variable $k = np + \sqrt{npq}x$. Considerarem x finit de manera que el límit $n \rightarrow \infty$ implica també $k \rightarrow \infty$. El símbol \simeq vol dir que el quocient dels dos costats tendeix a 1.

Notem les relacions

$$\frac{k}{n} = p + \frac{\sqrt{pq}}{\sqrt{n}}x, \quad 1 - \frac{k}{n} = q - \frac{\sqrt{pq}}{\sqrt{n}}x.$$

Ara tenim que, fent servir Stirling ($N! \simeq \sqrt{2\pi N} N^N e^{-N}$ quan $N \rightarrow \infty$)

$$\begin{aligned} p_n(k) &= \frac{n!}{k!(n-k)!} p^k q^{n-k} \simeq \frac{\sqrt{2\pi n} (\frac{n}{e})^n}{\sqrt{2\pi k} (\frac{k}{e})^k \sqrt{2\pi(n-k)} (\frac{n-k}{e})^{n-k}} p^k q^{n-k} \\ &= \frac{1}{\sqrt{2\pi k} \sqrt{1 - \frac{k}{n}}} \frac{n^n}{k^k (n-k)^{n-k}} p^k q^{n-k} = \frac{1}{\sqrt{2\pi k} \sqrt{1 - \frac{k}{n}}} \frac{1}{(\frac{k}{np})^k (\frac{1-k/n}{q})^{n-k}} \end{aligned}$$

La primera part es comporta asimptòticament

$$\frac{1}{\sqrt{2\pi k} \sqrt{1 - \frac{k}{n}}} \simeq \frac{1}{\sqrt{2\pi np} \sqrt{q}} = \frac{1}{\sqrt{2\pi npq}}$$

que és el coeficient de l'exponencial en (2.23).

Per la segona part notem primer les següents relacions ($N \rightarrow \infty$):

$$\left(1 + \frac{a}{N}\right)^{Nb} \simeq e^{ab}$$

$$\left(1 + \frac{a}{\sqrt{N}}\right)^{Nb} = \exp\{Nb \log\left(1 + \frac{a}{\sqrt{N}}\right)\} \simeq \exp\left\{Nb \left(\frac{a}{\sqrt{N}} - \frac{a^2}{2N}\right)\right\} = e^{ab\sqrt{N} - \frac{a^2 b}{2}}$$

on hem fet servir el desenvolupament de Taylor $\log(1+x) = x - \frac{x^2}{2} + o(x^2)$.

Ara és

$$\begin{aligned} \frac{1}{(\frac{k}{np})^k (\frac{1-k/n}{q})^{n-k}} &= \frac{1}{(1 + \sqrt{\frac{p}{q}} \frac{x}{\sqrt{n}})^{np + \sqrt{npq}x} (1 - \sqrt{\frac{p}{q}} \frac{x}{\sqrt{n}})^{nq - \sqrt{npq}x}} \\ &\simeq \frac{1}{e^{\sqrt{pq}x\sqrt{n} - \frac{q}{2}x^2} e^{qx^2} e^{-\sqrt{pq}x\sqrt{n} - \frac{p}{2}x^2} e^{px^2}} = e^{-x^2} = e^{-\frac{(k-np)^2}{2npq}}. \clubsuit \end{aligned}$$

Capítol 3

Funcions d'una variable aleatòria

Considerem una variable aleatòria X amb funció de densitat f_X . Si tenim $Y = g(X)$ on g és una certa funció, resulta que Y és una nova variable aleatòria. (Recordem que $X: \Omega \rightarrow \mathbb{R}$. Ara tenim $\Omega \xrightarrow{X} \mathbb{R} \xrightarrow{g} \mathbb{R}$ i en termes de composició de funcions $Y = g \circ X$.) Ens plantejem trobar les funcions de distribució i densitat de Y a partir de les de X .

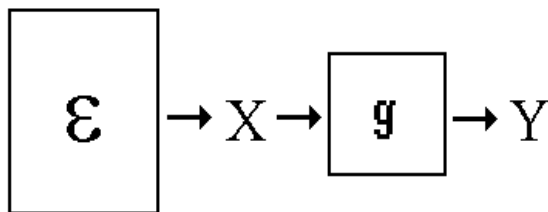


Figura 6: Transformació d'una v.a.

La funció g es pot considerar sota dos punts de vista. Per un costat podem pensar que el resultat de l'experiment X pateix algun tipus de procés que el transforma en $g(X)$ que és el que nosaltres observem. Per exemple, podem tenir un senyal electric aleatori del qual el nostre instrumental només ens permet calcular-ne el signe. Llavors, el senyal X pot prendre qualsevol valor real però nosaltres mesurem $g(X) = \text{signe}(X)$ que només val 1 i -1 . També podria ser que detectéssim el senyal després d'algun tipus de processat amb el que $g(X)$ seria una funció més complicada. En aquest exemple es veu que al fer la mesura final es pot perdre informació respecte al resultat inicial X . Això passa quan la funció g no és injectiva sobre el conjunt Ω_X .

Per una altra part, pot ser que disposem directament de X però preferim treballar amb una altra variable $Y = g(X)$. Si considerem que estem fent un canvi de variable, g ha de ser injectiva sobre Ω_X ja que hem de poder invertir la relació entre X i Y .

3.1 Cas discret

X pren els diferents valors x_i . Y pren els valors $g(x_i)$. Podria ser que un valor de Y correspongués a més d'un valor de X . Per tant tenim

$$P_Y(y) = P_X(g^{-1}(y)). \quad (3.1)$$

Exemple 3.1 Un circuit té tres estats de sortida $\Omega_X = \{-1, 0, 1\}$ en els que s'hi troba amb probabilitats $P_X(-1) = P_X(0) = P_X(1) = \frac{1}{3}$. Quina és la funció de probabilitat de $Y = X^2$?

Tenim $\Omega_Y = \{0, 1\}$. $P_Y(0) = P_X(g^{-1}(0)) = P_X(0) = \frac{1}{3}$ i $P_Y(1) = P_X(g^{-1}(1)) = P_X(\{-1, 1\}) = \frac{2}{3}$. ♦

3.2 Cas continu

Considerem la v.a. $Y = g(X)$. Definim els conjunts $\phi_y = \{x \mid g(x) \leq y\}$. Llavors

$$F_Y(y) = P(Y \leq y) = P(g(X) \leq y) = P_X(\phi_y). \quad (3.2)$$

Per que aquesta probabilitat estigui definida g no pot ser qualsevol funció. Cal que

- (1) g estigui definida pels valors que pot prendre X . És a dir, $\Omega_X \subset \text{dom}(g)$.
- (2) ϕ_y sigui un borelià per a tot y . En aquest cas es diu que g és una funció de Baire.
- (3) $P(g(X) = \pm\infty) = 0$. El conjunt de singularitats infinites ha de ser de mesura 0 en el sentit de la probabilitat. Això passa sempre que X és contínua i g val $\pm\infty$ com a molt en un conjunt numerable de punts.

Totes les funcions que utilitzem en la pràctica verifiquen aquestes condicions. Per calcular F_Y de manera efectiva hem d'estudiar en cada cas la forma de g . Una vegada obtinguda F_Y podem calcular f_Y per derivació.

Partirem de la representació gràfica de g . Recorrem l'eix de les y 's i per cada valor observem la forma de ϕ_y . Aquest conjunt tindrà la forma de unió de alguns intervals de les x 's el que ens permet escriure la probabilitat combinant la funció F_X en alguns punts. Considerem diferents tipus de comportament:

- g monòtona.

Si és creixent, $\phi_y = (-\infty, x]$ i, per tant, $F_Y(y) = F_X(x)$.

Si és decreixent, $\phi_y = [x, \infty)$ i, per tant, $F_Y(y) = 1 - F_X(x)$.

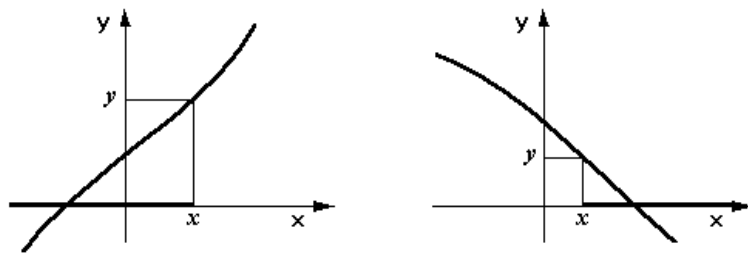


Figura 7: g monòtona.

- g no és monòtona.

En el primer exemple $\phi_y = [x_1, x_2]$ i, per tant, $F_Y(y) = F_X(x_2) - F_X(x_1)$.

En el segon exemple $\phi_y = (-\infty, x_1] \cup [x_2, x_3]$ i, per tant, $F_Y(y) = F_X(x_3) - F_X(x_2) + F_X(x_1)$.

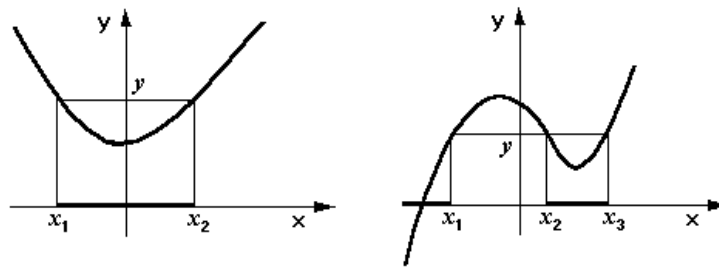


Figura 8: g no monòtona.

- g és constant en un tros.

L'interval $[a, b]$ que té probabilitat finita va a parar a un valor únic. Hi ha llavors un punt y_1 tal que $P(\{y_1\}) \neq 0$. La variable que resulta té comportament discret (mixte).

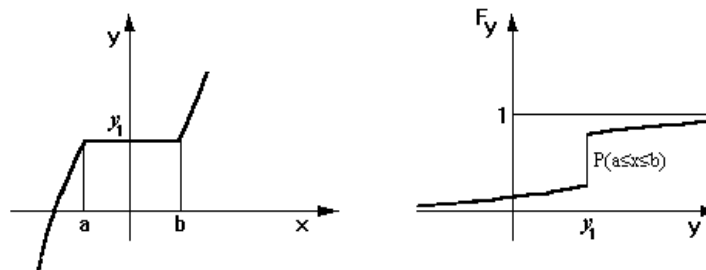


Figura 9: g constant.

- g és discontinua.

Al recórrer l'interval $c \leq y \leq d$, ϕ_y resta invariant. Llavors F_Y resulta constant sobre l'interval $[c, d]$.

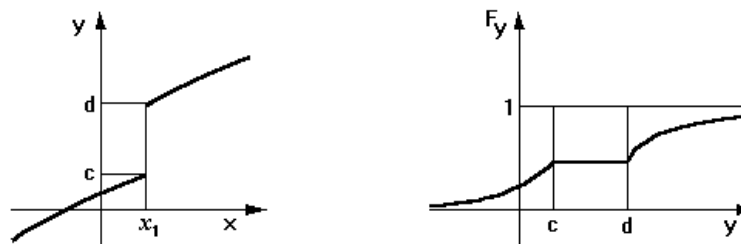
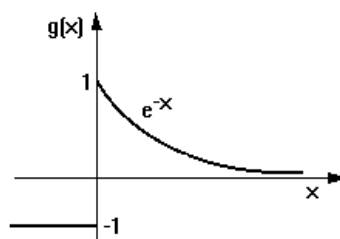


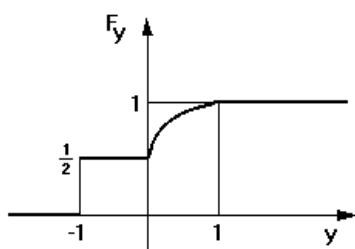
Figura 10: g discontinua.

Exemple 3.2 Sigui X una v.a. de Cauchy amb $F_X(x) = \frac{1}{2} + \frac{1}{\pi} \arctan x$. Considerem $Y = g(X)$ on g és la funció de la figura.



Tenim el següent:

- $F_Y(y) = 0$ per $y < -1$.
- En $y = -1$, $F_Y(y)$ salta $P_X(g^{-1}(-1)) = P_X((-\infty, 0]) = F_X(0) = \frac{1}{2}$.
- $F_Y(y)$ es manté constant en $-1 < y < 0$.
- Per $0 < y < 1$ tenim $x = -\ln y$ i $F_Y(y) = P_X(\phi_y) = P_X((-\infty, 0] \cup (x, \infty]) = P_X((-\infty, 0]) + 1 - P_X((-\infty, x]) = F_X(0) + 1 - F_X(-\ln y) = 1 + \frac{1}{\pi} \arctan(\ln y)$.
- Per $y \geq 1$, $F_Y(y) = 1$.



El resultat és una variable mixta amb funció de densitat

$$f_Y(y) = F'_Y(y) = \frac{1}{2}\delta(y+1) + \frac{1}{\pi y(1+(\ln y)^2)}(u(y) - u(y-1)). \blacklozenge$$

3.3 Càlcul de la funció de densitat

Quan g és derivable a trossos i no és constant en cap tros podem obtenir directament f_Y sense calcular prèviament la funció de distribució.

L'equació $y = g(x)$ té cap, una o vàries solucions $x_i(y)$. Llavors

$$f_Y(y) = \sum_i f_X(x_i(y)) \frac{1}{|g'(x_i(y))|}. \quad (3.3)$$

En l'anterior equació posem 0 quan no hi han solucions.

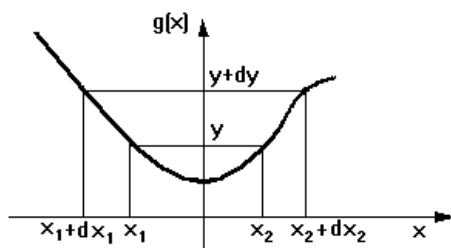
DEM: Per definició

$$f_Y(y) = \lim_{\Delta y \rightarrow 0} \frac{P(y \leq Y \leq y + \Delta y)}{\Delta y}.$$

Per tant, considerem $\Delta y > 0$ i prou petit per que tinguem

$$P(y \leq Y \leq y + \Delta y) = \sum_i P(X \text{ entre } x_i \text{ i } x_i + \Delta x_i)$$

on x_i són els valors tals que $g(x_i) = y$. Notem que Δx_i pot ser negatiu.



Ara calculem

$$f_Y(y) = \lim_{\Delta y \rightarrow 0} \sum_i \frac{P(X \text{ entre } x_i \text{ i } x_i + \Delta x_i)}{\Delta y}$$

$$= \lim_{\Delta y \rightarrow 0} \sum_i \frac{P(X \text{ entre } x_i \text{ i } x_i + \Delta x_i)}{|\Delta x_i|} \frac{1}{\frac{\Delta y}{|\Delta x_i|}} = \sum_i f_X(x_i) \frac{1}{\left| \frac{dy}{dx} \right|_{x_i}}. \clubsuit$$

Exemple 3.3 X és gaussiana amb $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$. $Y = X^2$.

L'equació $y = x^2$ no té solució per y negativa. Per y positiva té dues solucions $x_1(y) = -\sqrt{y}$ i $x_2(y) = \sqrt{y}$. Com $g'(x) = 2x$, $|g'(x_1)| = |g'(x_2)| = 2\sqrt{y}$. Llavors $f_Y(y) = 0$ si $y < 0$ i

$$f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y}{2}} \frac{1}{2\sqrt{y}} + \frac{1}{\sqrt{2\pi}} e^{-\frac{y}{2}} \frac{1}{2\sqrt{y}} = \frac{e^{-\frac{y}{2}}}{\sqrt{2\pi y}}$$

si $y \geq 0$. ♦

Exemple 3.4 X és uniforme en $[0, 2\pi]$. $Y = \sin X$.

$f_X(x) = \frac{1}{2\pi}$ i es veu immediatament que $\Omega_Y = [-1, 1]$. Per $-1 < y < 1$, l'equació $y = \sin x$ té dues solucions en les quals $|g'(x_1)| = |g'(x_2)| = \cos x_1 = \sqrt{1 - y^2}$. Llavors

$$f_Y(y) = \frac{1}{\pi \sqrt{1 - y^2}}. \spadesuit$$

Capítol 4

Paràmetres estadístics

En el que segueix, X designa una variable aleatòria discreta amb funció de probabilitat P_X o contínua amb funció de densitat f_X .

4.1 Esperança

L'esperança de la v.a. contínua X és el nombre

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx. \quad (4.1)$$

En el cas discret és

$$E[X] = \sum_k x_k P_X(x_k). \quad (4.2)$$

$E[X]$ és una mitjana dels valors que pot prendre X ponderats amb les seves probabilitats.

$E[X]$ també s'anomena valor esperat, valor mitjà o mitjana de la variable X . Altres notacions emprades són m_X o \overline{X} .

Les seves propietats bàsiques són:

- (1) Si $Y = g(X)$, $E[Y] = E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx$ (teorema de l'esperança),
- (2) $E[1] = 1$,
- (3) $E[\alpha g(X) + \beta h(X)] = \alpha E[g(X)] + \beta E[h(X)]$, per $\alpha, \beta \in \mathbb{R}$,
- (4) Si $g \geq 0$, $E[g(X)] \geq 0$.

DEM:

- (1) Considerem una partició de l'eix real amb punts y_i separats per la distància $\Delta y_i = y_{i+1} - y_i$. $E[Y]$ s'obté considerant

$$S = \sum_i y_i P(y_i \leq Y \leq y_i + \Delta y_i)$$

i prenent el límit quan les particions són cada vegada més fines. Si designem $x_i^{(k)}$ el conjunt de punts tals que $g(x_i^{(k)}) = y_i$, tenim que

$$S = \sum_i \sum_k g(x_i^{(k)}) P(X \text{ entre } x_i^{(k)} \text{ i } x_i^{(k)} + \Delta x_i^{(k)})$$

i aquests formen una partició del recorregut de la variable X . Podem estendre aquesta partició a tot l'eix afegint parts que tenen probabilitat 0. En el cas de variables contínues, al fer el límit trobem

$$E[Y] = \int_{-\infty}^{\infty} y f_Y(y) dy = \int_{-\infty}^{\infty} g(x) f_X(x) dx.$$

El procediment amb que hem construït aquestes integrals consistent en fer la partició no en els valors de la variable sinó en els de la funció dóna lloc al que s'anomena integral de Lebesgue. Aquesta té propietats que la fan una eina poderosa en l'anàlisi. A efectes pràctics, en les aplicacions resulta suficient la integral de Riemann ja que els dos tipus d'integració coincideixen en les funcions i conjunts d'integració habituals.

Aquesta propietat és simplement la regla de canvi de variable per integrals. Això es veu fàcilment en el cas que g és monòtona creixent amb recorregut (a, b) . Llavors f_Y és nul·la fora d'aquest interval i, fent el canvi $y = g(x)$,

$$E[Y] = \int_{-\infty}^{\infty} y f_Y(y) dy = \int_a^b y f_Y(y) dy = \int_{-\infty}^{\infty} g(x) \frac{f_X(x)}{g'(x)} \cdot g'(x) dx = \int_{-\infty}^{\infty} g(x) f_X(x) dx.$$

(2) Posem $g(X) = 1$ en l'anterior propietat.

(3) La integral és lineal respecte a les combinacions de funcions. Això és, $\int (\alpha g + \beta h) = \alpha \int g + \beta \int h$.

(4) La integral d'una funció positiva és positiva. ♣

Per interpretar el significat del nombre $E[X]$ considerem el cas d'una variable discreta finita amb valors x_i i probabilitats p_i , $i = 1, \dots, n$. Si fem N vegades l'experiment obtenim els resultats $r^{(a)}$, $a = 1, \dots, N$ on cada valor correspon a algun dels x_i . Als N resultat, x_i apareixerà un nombre de vegades N_i corresponent a una freqüència relativa que es pot aproximar per la seva probabilitat quan N és gran. Si fem la mitjana aritmètica dels resultats

$$\frac{1}{N} \sum_{a=1}^N r^{(a)} = \frac{1}{N} \sum_{i=1}^n x_i N_i = \sum_{i=1}^n x_i \frac{N_i}{N} \longrightarrow \sum_{i=1}^n x_i p_i = E[X].$$

Notem també que mentre la funció de densitat sempre és integrable a tot \mathbb{R} , l'expressió (4.1) pot no existir. En aquest cas, si la funció integrada en (4.1) és positiva direm que $E[X] = \infty$. Si aquesta funció conté àrees positives i negatives infinites direm que l'esperança no existeix. Si fem repeticions de l'experiment que ens genera X i anem calculant les mitjanes aritmètiques dels resultats obtinguts ens anirem acostant al valor $E[X]$. Quan l'esperança és infinita aquestes mitjanes divergiran cap a infinit. Quan l'esperança no existeix mostraran oscil·lacions irregulars.

4.2 Variància

$$V[X] = E[(X - \bar{X})^2]. \quad (4.3)$$

Les seves propietats bàsiques són:

- (1) $V[X] \geq 0$,
- (2) $V[X] = E[X^2] - \bar{X}^2$.
- (3) $V[\alpha X] = \alpha^2 V[X]$ per $\alpha \in \mathbb{R}$.
- (4) $V[X + \beta] = V[X]$ per $\beta \in \mathbb{R}$.

DEM:

- (1) Per la propietat (4) de l'esperança.
- (2) $V[X] = E[X^2 - 2\bar{X}X + \bar{X}^2] = E[X^2] - 2\bar{X}E[X] + \bar{X}^2 E[1] = E[X^2] - \bar{X}^2$.
- (3) Per la propietat (3) de l'esperança, $E[\alpha X] = \alpha E[X]$ i $E[(\alpha X)^2] = E[\alpha^2 X^2] = \alpha^2 E[X^2]$. Utilitzant la propietat (2), queda demostrat.

$$(4) \quad V[X + \beta] = E[(X + \beta - \overline{(X + \beta)})^2] = E[(X - \overline{X})^2]. \spadesuit$$

L'esperança indica generalment al voltant d'on es concentren els valors amb probabilitat alta. La variància mesura l'amplada d'aquesta regió. El quadrat a l'expressió (4.3) es posa per mesurar sempre desviacions positives. El següent paràmetre mesura el mateix però té comportament "lineal".

4.3 Desviació estàndard

$$Std[X] = V[X]^{\frac{1}{2}}. \quad (4.4)$$

Verifica $Std[\alpha X] = |\alpha| Std[X]$, per $\alpha \in \mathbb{R}$. $Std[X]$ també es coneix com desviació típica.

En el que segueix denotarem $m = E[X]$, $\sigma^2 = V[X]$ i $\sigma = Std[X]$.

4.4 Variable aleatòria normalitzada

Donada la v.a. X , fent un desplaçament i un canvi d'escala, definim la variable normalitzada

$$\tilde{X} = \frac{X - m}{\sigma}. \quad (4.5)$$

Aquesta variable verifica $E[\tilde{X}] = 0$, $V[\tilde{X}] = 1$ i $Std[\tilde{X}] = 1$.

La variable normalitzada també s'anomena tipificada o estandarditzada.

Notem que una variable aleatòria pot tenir dimensions físiques com, per exemple, longitud, freqüència o energia. Com les probabilitats són adimensionals, la funció de distribució també ho és. En canvi, si X té dimensió L la funció de densitat té dimensions L^{-1} , l'esperança L , la variància L^2 i la desviació estàndard L . Les variables normalitzades no tenen dimensions.

4.5 Paràmetres de les distribucions usals

- **Bernoulli.**

$$m = p, \quad \sigma^2 = pq, \quad \sigma = \sqrt{pq}.$$

DEM:

$$m = E[X] = 0 \cdot q + 1 \cdot p = p.$$

$$\sigma^2 = V[X] = (0 - p)^2 \cdot q + (1 - p)^2 \cdot p = pq. \spadesuit$$

- **Binomial.**

$$m = np, \quad \sigma^2 = npq, \quad \sigma = \sqrt{npq}.$$

DEM: En el que segueix considerem p i q variables independents i posem $q = 1 - p$ només al final. Una derivació afecta tot el que té a la dreta.

$$m = \sum_{k=0}^n k \cdot \binom{n}{k} p^k q^{n-k} = p \frac{d}{dp} (p + q)^n = pn(p + q)^{n-1} = np.$$

$$E[X^2] = \sum_{k=0}^n k^2 \cdot \binom{n}{k} p^k q^{n-k} = p \frac{d}{dp} p \frac{d}{dp} (p + q)^n$$

$$= pn(p + q)^{n-1} + n(n-1)p^2(p + q)^{n-2} = n^2 p^2 + npq.$$

$$\sigma^2 = (n^2 p^2 + npq) - (np)^2 = npq. \clubsuit$$

• **Geomètrica.**

$$m = \frac{1}{p}, \sigma^2 = \frac{q}{p^2}, \sigma = \frac{\sqrt{q}}{p}.$$

DEM:

$$m = \sum_{n=1}^{\infty} n \cdot q^{n-1} p = \frac{d}{dq} \sum_{n=0}^{\infty} q^n p = \frac{d}{dq} \frac{p}{1-q} = \frac{p}{(1-q)^2} = \frac{1}{p}.$$

$$E[X^2] = \sum_{n=1}^{\infty} n^2 \cdot q^{n-1} p = \frac{d}{dq} q \frac{d}{dq} \sum_{n=0}^{\infty} q^n p = \frac{1+q}{(1-q)^3} p = \frac{1+q}{p^2}.$$

$$\sigma^2 = \frac{1+q}{p^2} - \frac{1}{p^2} = \frac{q}{p^2}. \clubsuit$$

• **Hipergeomètrica.**

$$m = \frac{nb}{N}, \sigma^2 = \frac{nb(N-b)(N-n)}{N^2(N-1)}.$$

DEM: Considerem les igualtats

$$\left(\frac{d}{dx}(1+x)^A\right)(1+x)^B = A(1+x)^{A+B-1},$$

$$\left(\frac{d}{dx}x \frac{d}{dx}(1+x)^A\right)(1+x)^B = A(1+x)^{A+B-1} + A(A-1)x(1+x)^{A+B-2}.$$

Si utilitzem el binomi de Newton en totes les potències, trobem les relacions combinatòries

$$\sum_{k+l=n} k \binom{A}{k} \binom{B}{l} = A \binom{A+B-1}{n-1},$$

$$\sum_{k+l=n} k^2 \binom{A}{k} \binom{B}{l} = A \binom{A+B-1}{n-1} + A(A-1) \binom{A+B-2}{n-2}.$$

Aplicant aquest resultat trobem

$$E[X] = \sum_k k \frac{\binom{b}{k} \binom{N-b}{n-k}}{\binom{N}{n}} = \frac{1}{\binom{N}{n}} b \binom{N-1}{n-1} = b \frac{n}{N}.$$

$$\begin{aligned} E[X^2] &= \sum_k k^2 \frac{\binom{b}{k} \binom{N-b}{n-k}}{\binom{N}{n}} \\ &= \frac{1}{\binom{N}{n}} \left(b \binom{N-1}{n-1} + b(b-1) \binom{N-2}{n-2} \right) = \frac{bn}{N} + b(b-1) \frac{n(n-1)}{N(N-1)}. \\ \sigma^2 &= \left(\frac{bn}{N} + b(b-1) \frac{n(n-1)}{N(N-1)} \right) - b^2 \frac{n^2}{N^2} = \frac{nb(N-b)(N-n)}{N^2(N-1)}. \clubsuit \end{aligned}$$

• **Poisson.**

$$m = \alpha, \sigma^2 = \alpha, \sigma = \sqrt{\alpha}.$$

DEM:

$$m = \sum_{k=0}^{\infty} k \cdot e^{-\alpha} \frac{\alpha^k}{k!} = e^{-\alpha} \alpha \frac{d}{d\alpha} \sum_{k=0}^{\infty} \frac{\alpha^k}{k!} = e^{-\alpha} \alpha \frac{d}{d\alpha} e^{\alpha} = \alpha.$$

$$E[X^2] = \sum_{k=0}^{\infty} k^2 \cdot e^{-\alpha} \frac{\alpha^k}{k!} = e^{-\alpha} \alpha \frac{d}{d\alpha} \alpha \frac{d}{d\alpha} e^{\alpha} = e^{-\alpha} \alpha (e^{\alpha} + \alpha e^{\alpha}) = \alpha^2 + \alpha.$$

$$\sigma^2 = (\alpha^2 + \alpha) - \alpha^2 = \alpha. \clubsuit$$

- **Uniforme.**

$$m = \frac{a+b}{2}, \sigma^2 = \frac{(b-a)^2}{12}, \sigma = \frac{(b-a)}{\sqrt{12}}.$$

DEM:

$$m = \int_a^b x \cdot \frac{1}{b-a} dx = \frac{a+b}{2}. \quad E[X^2] = \int_a^b x^2 \cdot \frac{1}{b-a} dx = \frac{b^2 + ab + a^2b}{3}.$$

$$\sigma^2 = \frac{b^2 + ab + a^2b}{3} - \frac{(a+b)^2}{4} = \frac{(a-b)^2}{12}. \clubsuit$$

- **Exponencial.**

$$m = \frac{1}{\lambda}, \sigma^2 = \frac{1}{\lambda^2}, \sigma = \frac{1}{\lambda}.$$

DEM:

$$m = \int_0^\infty x \cdot \lambda e^{-\lambda x} dx = \frac{1}{\lambda} \int_0^\infty z \cdot e^{-z} dz = \frac{1}{\lambda}.$$

$$E[X^2] = \int_0^\infty x^2 \cdot \lambda e^{-\lambda x} dx = \frac{1}{\lambda^2} \int_0^\infty z^2 \cdot e^{-z} dz = \frac{2}{\lambda^2}.$$

(A les integrals hem fet el canvi $z = \lambda x$.)

$$\sigma^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}. \clubsuit$$

- **Gaussiana.**

Els paràmetres m i σ^2 ja apareixen a la funció de densitat normal.

DEM:

$$E[X] = \int_{-\infty}^\infty x \cdot \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty (\sigma z + m) e^{-\frac{z^2}{2}} dz = m.$$

$$E[X^2] = \int_{-\infty}^\infty x^2 \cdot \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty (\sigma z + m)^2 e^{-\frac{z^2}{2}} dz$$

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

(A les integrals hem fet el canvi $z = \frac{x-m}{\sigma}$. S'ha fet servir que la integral d'una funció ímparella en un interval simètric és 0.)

$$V[X] = (\sigma^2 + m^2) - m^2 = \sigma^2. \clubsuit$$

- **Cauchy.**

Els paràmetres m , σ^2 no existeixen ja que la integral $\int_{-\infty}^\infty \frac{x}{x^2+1} dx$ no existeix.

4.6 Moments d'una variable aleatòria

El **moment** n -èsim és

$$m_n = E[X^n] = \int_{-\infty}^\infty x^n f_X(x) dx, \quad (4.6)$$

o, en el cas discret

$$m_n = \sum_k x_k^n P_X(x_k), \quad (4.7)$$

on $n = 0, 1, \dots$. Notem que $m_0 = 1$, $m_1 = E[X]$ i $V[X] = m_2 - m_1^2$.

Els **moments centrals** són

$$\mu_n = E[(X - \bar{X})^n] = \int_{-\infty}^{\infty} (x - \bar{X})^n f_X(x) dx, \quad (4.8)$$

on $n = 0, 1, \dots$. Notem que $\mu_0 = 1$, $\mu_1 = 0$ i $\mu_2 = V[X]$.

Exemple 4.1 Calcular els moments d'una variable exponencial.

$$m_n = E[X^n] = \int_0^{\infty} x^n \lambda e^{-\lambda x} dx = \frac{1}{\lambda^n} \int_0^{\infty} z^n e^{-z} dz = \frac{n!}{\lambda^n}. \blacklozenge$$

Exemple 4.2 Calcular els moments centrals d'una variable normal.

$$\mu_n = E[(X - m)^n] = \int_{-\infty}^{\infty} (x - m)^n \cdot \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}} dx = \frac{(\sqrt{2}\sigma)^n}{\sqrt{\pi}} \int_{-\infty}^{\infty} z^n e^{-z^2} dz$$

(canvi $z = \frac{x-m}{\sqrt{2}\sigma}$). Ara veiem que

$$\mu_n = \begin{cases} 0 & \text{si } n \text{ és senar} \\ \sigma^n \frac{n!}{2^{\frac{n}{2}} (\frac{n}{2})!} & \text{si } n \text{ és parell} \end{cases}$$

(veure a l'apèndix les fórmules (D.12) i (D.14).) \blacklozenge

4.7 Desigualtat de Txebyshev

Per a tota v.a. X amb esperança i variància finites:

$$\forall a > 0 \quad P(|X - \bar{X}| \geq a) \leq \frac{\sigma^2}{a^2}. \quad (4.9)$$

Si posem $a = k\sigma$ podem expressar-ho de la forma

$$P(|X - \bar{X}| \geq k\sigma) \leq \frac{1}{k^2}. \quad (4.10)$$

DEM:

$$\begin{aligned} \sigma^2 &= \int_{-\infty}^{\infty} (x - \bar{X})^2 f_X(x) dx \geq \int_{|x - \bar{X}| \geq a} (x - \bar{X})^2 f_X(x) dx \\ &\geq a^2 \int_{|x - \bar{X}| \geq a} f_X(x) dx = a^2 P(|X - \bar{X}| \geq a). \clubsuit \end{aligned}$$

L'anterior resultat caracteritza σ com a mesura de la dispersió.

4.8 Llei dels grans nombres

La següent és la llei dels grans nombres en forma dèbil per un experiment de Bernoulli.

Repetim n vegades un experiment de Bernoulli amb probabilitat d'èxit p . Considerem la variable aleatòria X donada pel nombre total d'èxits. Llavors,

$$\forall \epsilon > 0 \quad \lim_{n \rightarrow \infty} P\left(\left|\frac{X}{n} - p\right| \leq \epsilon\right) = 1. \quad (4.11)$$

DEM: X és una variable binomial d'esperança np i variància npq . Utilitzant la desigualtat de Txebychev tenim que

$$P\left(\left|\frac{X}{n} - p\right| \leq \epsilon\right) = 1 - P(|X - np| > n\epsilon) \geq 1 - \frac{npq}{n^2\epsilon^2} = 1 - \frac{pq}{n\epsilon^2}.$$

Fent el límit $n \rightarrow \infty$ trobem (4.11). ♣

La llei dels grans nombres ens diu que en algun sentit hem obtingut (1.4). Només cal considerar que l'experiment de Bernoulli correspon a veure si es verifica l'esdeveniment A . Llavors $X = N_A$, $n = N$ i $\frac{X}{n}$ és la freqüència relativa.

Diem que $\frac{X}{n}$ tendeix a p en probabilitat si passa (4.11). Amb aquesta noció de convergència veiem, doncs, que les freqüències relatives d'un esdeveniment tendeixen a la seva probabilitat.

4.9 Teorema del límit central

Siguin X_1, X_2, X_3, \dots mesures independents corresponents a una variable aleatòria X d'esperança m i variància σ . La mitjana

$$Y_n = \frac{X_1 + X_2 + \dots + X_n}{n}$$

és tal que $\sqrt{n}Y_n$ tendeix, per $n \rightarrow \infty$, a una variable normal d'esperança m i variància σ .

4.10 Probabilitat, esperança, mesura

Els dos aspectes bàsics en problemes de probabilitat són l'obtenció de la probabilitat d'un esdeveniment i el càlcul de l'esperança d'una v.a. Els desenvolupaments teòrics afavoreixen la probabilitat com a concepte bàsic i deixen l'esperança com a concepte subsidiari.

En realitat el càlcul de valors esperats és l'element principal en moltes aplicacions. De fet la probabilitat és un cas particular de l'esperança on la variable és la de Bernoulli associada a l'esdeveniment. L'objecte fonamental és la mesura que tinguem. Aquesta mesura ens dona els valors de les probabilitats i de les esperances segons quina cosa mesurem.

El fet que en moltes aplicacions només importi calcular mitjanes es deu a que en aquests casos es poden negligir les fluctuacions i es pot suposar que les variables prenen valors constants. Aquest és el cas en sistemes grans (poblacions humanes, matèria macroscòpica) on l'elevat nombre de components permet aplicar la llei dels grans nombres.

Part II

Variables aleatòries multidimensionals i teoria de l'estimació

Capítol 5

Variables aleatòries multidimensionals

Anàlogament al cas unidimensional, una variable aleatòria n -dimensional és una aplicació $X: \Omega \rightarrow \mathbb{R}^n$ mesurable. És a dir, les antiimatges de borelians de \mathbb{R}^n són esdeveniments de la σ -àlgebra que tinguem a Ω . Es comprova que si l'aplicació és $X = (X_1, X_2, \dots, X_n)$ llavors X_i , $i = 1, \dots, n$ són variables aleatòries unidimensionals. De la mateixa manera, donades vèries variables aleatòries unidimensionals, l'aplicació sobre \mathbb{R}^n donada per $X = (X_1, X_2, \dots, X_n)$ és una v.a. n -dimensional. Les variables aleatòries multidimensionals també s'anomenen vectors aleatoris.

Essencialment, una v.a. multidimensional consisteix en el resultat d'un experiment expressat com una col·lecció de nombres reals. Per exemple, seleccionar una persona a l'atzar i mesurar la seva alçada i pes o mesurar les tres coordenades espacials de la posició d'un electró en un àtom.

En el que segueix, fixarem $n = 2$ per simplicitat. L'extensió a n arbitrari és directa i la resumim al final.

5.1 Variable aleatòria bidimensional

És una aplicació $\Omega \rightarrow \mathbb{R}^2$ que expressarem (X, Y) . Com X i Y són v.a.'s direm que tenim una distribució conjunta de probabilitat. L'àlgebra \mathcal{B}_2 es pot generar a partir de rectangles $[a_1, b_1] \times [a_2, b_2]$ o, de manera alternativa i com ara ens interessa, amb regions $(-\infty, a] \times (-\infty, b]$. Les probabilitats d'aquestes regions les determinen totes.

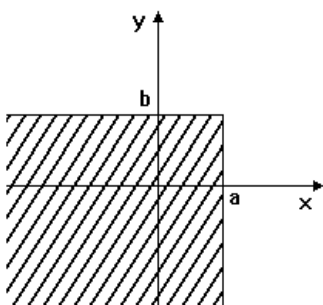


Figura 11: Regió $(-\infty, a] \times (-\infty, b]$.

5.2 Funció de distribució conjunta

La funció de distribució conjunta de dues v.a.'s és la funció

$$F_{XY}(x, y) = P(X \leq x \text{ i } Y \leq y). \quad (5.1)$$

És, per tant, la probabilitat de la regió $(-\infty, x] \times (-\infty, y]$.

Les seves propietats són:

- (1) $0 \leq F_{XY}(x, y) \leq 1$,
- (2) F_{XY} és una funció creixent en x i y , és a dir, si $x_1 \leq x_2$ i $y_1 \leq y_2$ llavors $F_{XY}(x_1, y_1) \leq F_{XY}(x_2, y_2)$,
- (3) $\lim_{x \rightarrow -\infty} F_{XY}(x, y) = 0$, $\lim_{y \rightarrow -\infty} F_{XY}(x, y) = 0$, $\lim_{x, y \rightarrow \infty} F_{XY}(x, y) = 1$,
- (4) F_{XY} és una funció contínua per la dreta segons la x i la y ,
- (5) $P(x_1 < X \leq x_2 \text{ i } y_1 < Y \leq y_2) = F_{XY}(x_2, y_2) - F_{XY}(x_2, y_1) - F_{XY}(x_1, y_2) + F_{XY}(x_1, y_1)$ (on $x_1 < x_2$ i $y_1 < y_2$).

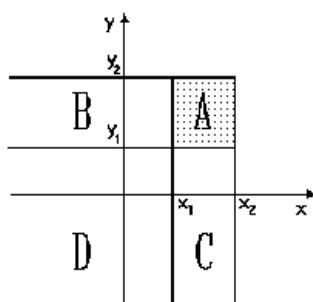


Figura 12

DEM:

(1), (2), (3) i (4) es demostren igual que en el cas unidimensional.

- (5) Considerem les regions disjunts $A = (x_1, x_2] \times (y_1, y_2]$, $B = (-\infty, x_1] \times (y_1, y_2]$, $C = (x_1, x_2] \times (-\infty, y_1]$, $D = (-\infty, x_1] \times (-\infty, y_1]$. Llavors $F_{XY}(x_2, y_2) = P(A) + P(B) + P(C) + P(D)$, $F_{XY}(x_1, y_2) = P(B) + P(D)$, $F_{XY}(x_2, y_1) = P(C) + P(D)$, $F_{XY}(x_1, y_1) = P(D)$. D'aquí eliminem $P(A)$ i obtenim la propietat. ♣

5.3 Funcions de distribució marginals

Són les funcions

$$F_X(x) = F_{XY}(x, \infty), \quad F_Y(y) = F_{XY}(\infty, y). \quad (5.2)$$

Són les funcions de distribució de les v.a.'s X i Y respectivament.

5.4 Funció de densitat conjunta

Si existeix la derivada

$$f_{XY}(x, y) = \frac{\partial^2 F_{XY}(x, y)}{\partial x \partial y} \quad (5.3)$$

es diu que X i Y són **conjuntament contínues** i f_{XY} s'anomena **funció de densitat conjunta** de les variables X i Y .

Si expressem la propietat (5) de la funció de distribució per un rectangle infinitesimal trobem l'expressió simbòlica

$$f_{XY}(x, y) dx dy = P(x \leq X \leq x + dx, y \leq Y \leq y + dy). \quad (5.4)$$

DEM: Fem Taylor fins a segon ordre:

$$\begin{aligned}
P(x \leq X \leq x + \Delta x, y \leq Y \leq y + \Delta y) &= \\
F_{XY}(x + \Delta x, y + \Delta y) - F_{XY}(x + \Delta x, y) - F_{XY}(x, y + \Delta y) + F_{XY}(x, y) \\
&\simeq F_{XY}(x, y) + \frac{\partial}{\partial x} F_{XY}(x, y) \Delta x + \frac{\partial}{\partial y} F_{XY}(x, y) \Delta y \\
&+ \frac{1}{2} \frac{\partial^2}{\partial x^2} F_{XY}(x, y) \Delta x^2 + \frac{1}{2} \frac{\partial^2}{\partial y^2} F_{XY}(x, y) \Delta y^2 + \frac{\partial^2}{\partial x \partial y} F_{XY}(x, y) \Delta x \Delta y \\
&- (F_{XY}(x, y) + \frac{\partial}{\partial x} F_{XY}(x, y) \Delta x + \frac{1}{2} \frac{\partial^2}{\partial x^2} F_{XY}(x, y) \Delta x^2) \\
&- (F_{XY}(x, y) + \frac{\partial}{\partial y} F_{XY}(x, y) \Delta y + \frac{1}{2} \frac{\partial^2}{\partial y^2} F_{XY}(x, y) \Delta y^2) + F_{XY}(x, y) \\
&= \frac{\partial^2}{\partial x \partial y} F_{XY}(x, y) \Delta x \Delta y. \clubsuit
\end{aligned}$$

f_{XY} verifica les propietats:

- (1) $f_{XY} \geq 0$,
- (2) $P((X, Y) \in A) = \int \int_A f_{XY}(x, y) dx dy$, per a tot borelià $A \subset \mathbb{R}^2$.
- (3) $F_{XY}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f_{XY}(x', y') dx' dy'$,
- (4) $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(x, y) dx dy = 1$.

DEM: (1) es segueix de la propietat (1) de F_{XY} . (3) i (4) s'obtenen posant en (2) $A = (-\infty, x] \times (-\infty, y]$ i $A = \mathbb{R}^2$ respectivament. Per demostrar (2) podem dividir A en rectangles cada vegada més petits i aplicar (5.4). Això limita el tipus de conjunt A que podem considerar però és suficient en la pràctica i correspon a la integral de Riemann. Per considerar conjunts més generals cal passar a la integral de Lebesgue. \clubsuit

5.5 Funcions de densitat marginals

Són les funcions de densitat de les variables unidimensionals X i Y :

$$f_X(x) = \frac{dF_X(x)}{dx}, \quad f_Y(y) = \frac{dF_Y(y)}{dy}, \quad (5.5)$$

o, de manera alternativa

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y) dy, \quad f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y) dx. \quad (5.6)$$

DEM: De (5.2) i la propietat (3) de la densitat conjunta obtenim

$$\begin{aligned}
F_X(x) &= \int_{-\infty}^x \left(\int_{-\infty}^y f_{XY}(x', y) dx' \right) dy. \\
f_X(x) &= \frac{dF_X(x)}{dx} = \frac{d}{dx} \int_{-\infty}^{\infty} \left(\int_{-\infty}^x f_{XY}(x', y) dx' \right) dy = \int_{-\infty}^{\infty} f_{XY}(x, y) dy. \clubsuit
\end{aligned}$$

Exemple 5.1 Una variable bidimensional (X, Y) té densitat conjunta

$$f(x, y) = \begin{cases} 6x & \text{si } x \geq 0, y \geq 0, x + 1 \leq 1, \\ 0 & \text{en cas contrari.} \end{cases}$$

Comprovem que $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1$:

$$\int_0^1 \int_0^{1-x} 6x dx dy = 6 \int_0^1 x \left(\int_0^{1-x} dy \right) dx = 6 \int_0^1 (x - x^2) dx = 6 \left(\frac{1}{2} - \frac{1}{3} \right) = 1.$$

Calculem ara les densitats marginals:

$$f_X(x) = \int_0^{1-x} 6x dy = 6(x - x^2), \quad 0 < x < 1.$$

$$f_Y(y) = \int_0^{1-y} 6x dx = 3(1 - y)^2, \quad 0 < y < 1. \blacklozenge$$

5.6 Cas discret

Partim de X i Y v.a. discretes que prenen un nombre finit o numerable de valors x_i, y_j . La funció de probabilitat conjunta és

$$P_{XY}(x_i, y_j) = P(X = x_i \text{ i } Y = y_j). \quad (5.7)$$

Les funcions de probabilitat marginals són

$$P_X(x_i) = \sum_{y_j} P_{XY}(x_i, y_j), \quad P_Y(y_j) = \sum_{x_i} P_{XY}(x_i, y_j). \quad (5.8)$$

Com es comprova fàcilment, les funcions de densitat marginals i les funcions de probabilitat marginals estan normalitzades a 1 i constitueixen funcions de densitat i probabilitat unidimensionals. En un context bidimensional l'esdeveniment $\{x\}$ vol dir realment $\{x\} \times \mathbb{R}$. En el cas discret l'esdeveniment $\{x_i\}$ vol dir realment $\{x_i\} \times \{y_1, \dots, y_n\}$. Els subíndex XY, X , etc. en les funcions de distribució i densitat conjunts i marginals es poden ometre si es sobreenten en les v.a.'s i la notació és prou clara.

Exemple 5.2 Una variable bidimensional (X, Y) és tal que X només pren els valors 1, 2, 3 i Y els valors 0, 1. La seva funció de probabilitat conjunta és representa en la taula següent:

$Y \backslash X$	1	2	3
0	0,19	0,25	0,11
1	0,13	0,15	0,17

És a dir, $P_{XY}(1, 0) = 0,19$, etc. Les sis probabilitats de la taula sumen 1, el que mostra que P_{XY} està ben normalitzada. Per calcular les marginals hem d'anar sumant els valors en les files i columnes de la taula. Llavors, la funció de probabilitat marginal de X és $P_X(1) = 0,32$, $P_X(2) = 0,3$ i $P_X(3) = 0,28$. Per la Y , $P_Y(0) = 0,55$, $P_Y(1) = 0,45$. \blacklozenge

	X = 1	X = 2	X = 3	
Y = 0	0,19	0,25	0,11	0,55
Y = 1	0,13	0,15	0,17	0,45
	0,32	0,3	0,28	

5.7 Algunes distribucions usuals

• Trinomial

És una generalització de la binomial. En un experiment considerem tres esdeveniments A_1 , A_2 i A_3 que formen una partició de Ω . Les seves probabilitats respectives p_1 , p_2 i p_3 verifiquen $p_1 + p_2 + p_3 = 1$. Repetim n vegades l'experiment i considerem les tres v.a. X_i = nombre de vegades que passa A_i , $i = 1, 2, 3$. Com $X_1 + X_2 + X_3 = n$, en el fons tenim dues v.a.'s discretes. Si indiquem per n_i el valor que pren la variable X_i llavors és:

$$P(n_1, n_2, n_3) = \frac{n!}{n_1!n_2!n_3!} p_1^{n_1} p_2^{n_2} p_3^{n_3}. \quad (5.9)$$

on, implícitament, $n_1 + n_2 + n_3 = n$. També denotarem aquesta expressió per $P(n_1, n_2)$.

En efecte, el nombre de maneres en que podem situar les n_1 , n_2 , n_3 ocurrències entre les n és triar n_1 de les n i n_2 de les $n - n_1$ restants. La resta queden fixades i $n_3 = n - n_1 - n_2$.

$$\binom{n}{n_1} \binom{n - n_1}{n_2} = \frac{n!}{n_1!n_2!n_3!}$$

La probabilitat marginal de X_1 és

$$\begin{aligned} P(n_1) &= \sum_{n_2} P(n_1, n_2) = \binom{n}{n_1} p_1^{n_1} \sum_{n_2=0}^{n-n_1} \binom{n-n_1}{n_2} p_2^{n_2} p_3^{n-n_1-n_2} \\ &= \binom{n}{n_1} p_1^{n_1} (p_2 + p_3)^{n-n_1} = \binom{n}{n_1} p_1^{n_1} (1 - p_1)^{n-n_1}. \end{aligned}$$

que és de tipus binomial. De manera anàloga es defineix la distribució **multinomial** de la qual les marginals són multinomials d'ordre menor.

Exemple 5.3 Es llença una parella de daus 4 vegades. Quina és la probabilitat que la suma dels dos daus valgui 7 i 8 exactament una vegada?

Indiquem X_1, X_2, X_3 el nombre de vegades que la suma val 7, 8 i qualsevol valor diferent de 7 i 8, respectivament. llavors tenim una v.a. trinomial amb $n = 4$, $p_1 = \frac{6}{36}$, $p_2 = \frac{5}{36}$, $p_3 = \frac{25}{36}$.

$$P(X_1 = 1, X_2 = 1, X_3 = 2) = \frac{4!}{1! \cdot 1! \cdot 2!} \cdot \frac{6}{36} \cdot \frac{5}{36} \cdot \left(\frac{25}{36}\right)^2 = 0,1339. \blacklozenge$$

• Uniforme.

La distribució uniforme sobre el domini $D \subset \mathbb{R}^2$ ve donada per la funció de densitat:

$$f(x, y) = \begin{cases} K & \text{si } (x, y) \in D, \\ 0 & \text{altrament.} \end{cases} \quad (5.10)$$

on K és una constant. La condició de normalització ens dona $1 = K \int \int_D dx dy$, d'on $K = \frac{1}{\text{area}(D)}$.

Les distribucions marginals d'una distribució uniforme *no* són uniformes en general.

Exemple 5.4 Considerem una distribució uniforme sobre un cercle de radi a centrat a l'origen de coordenades. En aquest cas $D = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq a^2\}$. $f(x, y) = \frac{1}{\pi a^2}$ sobre D . La densitat marginal de X és

$$f_X(x) = \int_{-\sqrt{a^2-x^2}}^{\sqrt{a^2-x^2}} \frac{1}{\pi a^2} dy = \frac{2}{\pi a^2} \sqrt{a^2 - x^2}$$

en l'interval $[-a, a]$.

Considerem les variables radi $r = \sqrt{x^2 + y^2}$ i angle $\theta = \arctan \frac{y}{x}$, on θ pren valors sobre un període complet. Podem calcular les funcions de densitat d'aquestes variables de la següent manera. Considerem $0 \leq r < r + \Delta r \leq a$. Llavors

$$P(r < \sqrt{x^2 + y^2} \leq r + \Delta r) = \frac{1}{\pi a^2} (\pi(r + \Delta r)^2 - \pi r^2) = \frac{2r\Delta r + (\Delta r)^2}{a^2}.$$

En el límit $\Delta r \rightarrow 0$ deduïm la funció de densitat $f_R(r) = 2\frac{r}{a^2}$ per $0 \leq r \leq a$. Anàlogament trobem

$$P(\theta < \arctan \frac{y}{x} \leq \theta + \Delta\theta) = \frac{1}{\pi a^2} \left(\pi a^2 \frac{\Delta\theta}{2\pi} \right) = \frac{\Delta\theta}{2\pi}$$

que ens dona la densitat angular $f_\Theta(\theta) = \frac{1}{2\pi}$. ♦

• Normal bidimensional.

La distribució gaussiana es generalitza considerant l'exponencial d'un polinomi de segon grau en les variables x i y .

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left\{\frac{(x-m_1)^2}{\sigma_1^2} - 2\rho\frac{(x-m_1)}{\sigma_1}\frac{(y-m_2)}{\sigma_2} + \frac{(y-m_2)^2}{\sigma_2^2}\right\}} \quad (5.11)$$

per $-\infty < x, y < \infty$. Els paràmetres són les esperances $-\infty < m_1, m_2 < \infty$, les desviacions típiques $\sigma_1, \sigma_2 > 0$ i la correlació $-1 < \rho < 1$.

5.8 Independència de variables aleatòries

Es diu que les v.a. X i Y són *independents* si per a tot parell A i B de borelians de \mathbb{R}

$$P(X \in A \text{ i } Y \in B) = P(X \in A)P(Y \in B). \quad (5.12)$$

Prenent $A = (-\infty, x]$ i $B = (-\infty, y]$ tenim que $F_{XY}(x, y) = F_X(x)F_Y(y)$. Fent les derivades respecte a x i y trobem que si X i Y són independents

$$f_{XY}(x, y) = f_X(x)f_Y(y). \quad (5.13)$$

Es veu fàcilment a l'integrar aquesta expressió sobre la regió $A \times B$ que si (5.13) es verifica llavors les variables X i Y són independents.

Quan X i Y són independents es diu que l'espai de probabilitat bidimensional és producte dels espais unidimensionals de X i de Y . L'operació (5.13) de multiplicar dues funcions d'una variable per obtenir-ne una de dues variables s'anomena producte tensorial o directe de les funcions f_X i f_Y . Cal tenir en compte que la situació d'independència és particular. En general, el conèixer les distribucions marginals de X i Y *no* determina la seva distribució conjunta.

5.9 Densitat condicionada

La funció de distribució de X condicionada a $Y = y$ és

$$F_X(x | y) = \frac{\int_{-\infty}^x f_{XY}(x', y) dx'}{f_Y(y)}. \quad (5.14)$$

Derivant respecte a x s'obté la funció de densitat condicionada

$$f_X(x | y) = \frac{f_{XY}(x, y)}{f_Y(y)}. \quad (5.15)$$

Les anteriors expressions només es defineixen quan $f_Y(y) \neq 0$.

DEM:

$$F_X(x | y) = \lim_{h \rightarrow 0} \frac{P(X \leq x \text{ i } y \leq Y \leq y + h)}{P(y \leq Y \leq y + h)} = \lim_{h \rightarrow 0} \frac{F_{XY}(x, y + h) - F_{XY}(x, y)}{F_Y(y + h) - F_Y(y)} = \frac{\frac{\partial F_{XY}(x, y)}{\partial y}}{\frac{\partial F_Y(y)}{\partial y}}.$$

La derivada del numerador es calcula amb la propietat (3) de la funció de densitat. ♣

Anàlogament podem definir $f_Y(y | x)$. Les funcions de densitat condicionades estan normalitzades a 1. Notem que la funció marginal $f_X(x)$ ens dona les probabilitats d'esdeveniments referits a X quan no sabem res sobre el resultat de la variable Y . La funció condicionada $f_X(x | y)$ ens dona aquestes probabilitats quan sabem que el resultat de la variable Y és y .

Si X i Y són independents, substituint (5.13) en (5.15) tenim, com és d'esperar, que $f_X(x | y) = f_X(x)$.

5.10 Esperança condicionada

L'esperança de X condicionada a que Y val y és

$$E[X | y] = \int_{-\infty}^{\infty} x f_X(x | y) dx. \quad (5.16)$$

Aquesta esperança verifica

$$E[X] = \int_{-\infty}^{\infty} E[X | y] f_Y(y) dy. \quad (5.17)$$

DEM:

$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty} x f_X(x) dx = \int_{-\infty}^{\infty} x \left(\int_{-\infty}^{\infty} f_{XY}(x, y) dy \right) dx \\ &= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} x \frac{f_{XY}(x, y)}{f_Y(y)} dx \right) f_Y(y) dy = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} x f_X(x | y) dx \right) f_Y(y) dy. \end{aligned}$$

I el terme del parèntesi és $E[X | y]$. ♣

(5.17) s'interpreta de la següent manera: per fer l'esperança de X podem fer primer l'esperança de X condicionada a Y . Això és una v.a. (funció de Y) i l'esperança de X és l'esperança d'aquesta variable. És a dir,

$$E[X] = E[E[X | Y]]. \quad (5.18)$$

Exemple 5.5 Per la variable bidimensional de l'exemple 5.1, calculem $E[X | y]$.

Necessitem primer la densitat condicionada

$$f_X(x | y) = \frac{f(x, y)}{f_Y(y)} = \frac{2x}{(1-y)^2}, \quad 0 < x < 1-y.$$

Ara podem calcular

$$E[X | y] = \int_0^{1-y} x \frac{2x}{(1-y)^2} dx = \frac{2}{(1-y)^2} \int_0^{1-y} x^2 dx = \frac{2}{3}(1-y). \blacklozenge$$

Exemple 5.6 Generem una v.a. X uniforme a $[3, 4]$ i després una variable Y exponencial de paràmetre $\lambda = X$. Calcular la densitat conjunta $f(x, y)$ i les esperances de X i de Y .

Tenim una variable bidimensional (X, Y) de la que coneixem la distribució marginal de X i la distribució de Y condicionada a X . Aquest és sempre el cas quan generem les dues variables fent

primer una i després, a partir del resultat d'aquesta, la segona. Així, el que coneixem per l'enunciat és $f(x) = 1$ per $3 \leq x \leq 4$ i $f(y | x) = xe^{-xy}$ per $y > 0$. La forma de (5.15) que ens interessa és $f(x, y) = f(y | x)f(x)$. Llavors $f(x, y) = xe^{-xy}$ per $3 \leq x \leq 4, y > 0$.

També tenim que $E[X] = \frac{3+4}{2} = \frac{7}{2}$ i $E[Y] = E[E[Y | X]] = E[\frac{1}{X}] = \int_3^4 x^{-1} \cdot 1 dx = \ln \frac{4}{3}$. ♦

Exemple 5.7 Calculem les **distribucions marginals i condicionades de la variable normal bidimensional** (5.11).

$$f_X(x) = \frac{e^{-\frac{1}{2(1-\rho^2)} \frac{(x-m_1)^2}{\sigma_1^2}}}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2(1-\rho^2)} \left\{ \frac{(y-m_2)^2}{\sigma_2^2} - 2\rho \frac{(x-m_1)}{\sigma_1} \frac{(y-m_2)}{\sigma_2} \right\}} dy$$

(fent el canvi $y \rightarrow z = \frac{(y-m_2)}{\sigma_2} - \rho \frac{(x-m_1)}{\sigma_1}$)

$$\begin{aligned} &= \frac{e^{-\frac{(x-m_1)^2}{2(1-\rho^2)\sigma_1^2}}}{2\pi\sigma_1\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2(1-\rho^2)} \{z^2 - \rho^2 \frac{(x-m_1)^2}{\sigma_1^2}\}} dz \\ &= \frac{e^{-\frac{(x-m_1)^2}{2(1-\rho^2)\sigma_1^2}}}{2\pi\sigma_1\sqrt{1-\rho^2}} e^{\frac{\rho^2(x-m_1)^2}{2(1-\rho^2)\sigma_1^2}} \sqrt{2\pi(1-\rho^2)} = \frac{e^{-\frac{(x-m_1)^2}{2\sigma_1^2}}}{\sqrt{2\pi}\sigma_1}. \end{aligned}$$

Per tant, fent un càlcul idèntic per f_Y veiem que les distribucions marginals d'una normal bidimensional són normals unidimensionals d'esperances m_1 i m_2 i variàncies σ_1^2 i σ_2^2 respectivament. A partir d'això tenim que la densitat condicionada

$$\begin{aligned} f_X(x | y) &= \frac{f(x, y)}{f_Y(y)} \\ &= \frac{1}{\sqrt{2\pi}\sigma_1\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)} \left\{ \frac{(x-m_1)^2}{\sigma_1^2} - 2\rho \frac{(x-m_1)}{\sigma_1} \frac{(y-m_2)}{\sigma_2} + \rho^2 \frac{(y-m_2)^2}{\sigma_2^2} \right\}} \\ &= \frac{1}{\sqrt{2\pi}\sigma_1\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)} \left(\frac{(x-m_1)}{\sigma_1} - \rho \frac{(y-m_2)}{\sigma_2} \right)^2}. \end{aligned}$$

Novament resulta ser una distribució gaussiana, ara amb variància $\sigma^2 = \sigma_1^2(1-\rho^2)$ i esperança $m = E[X | y] = m_1 + \rho \frac{\sigma_1}{\sigma_2}(y - m_2)$.

Notem que X i Y són independents si i només si $\rho = 0$. ♦

5.11 Generalització al cas n -dimensional

Considerem X_1, X_2, \dots, X_n v.a.'s unidimensionals o, de manera equivalent, una aplicació mesurable $\Omega \rightarrow \mathbb{R}^n$.

La funció de distribució conjunta és

$$F_{X_1 X_2 \dots X_n}(x_1, x_2, \dots, x_n) = P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n). \quad (5.19)$$

És, per tant, la probabilitat de la regió $(-\infty, x_1] \times (-\infty, x_2] \times \dots \times (-\infty, x_n]$.

Diem que les n variables són conjuntament contínues si existeix la funció de densitat

$$f_{X_1 X_2 \dots X_n}(x_1, x_2, \dots, x_n) = \frac{\partial^n F_{X_1 X_2 \dots X_n}(x_1, x_2, \dots, x_n)}{\partial x_1 \partial x_2 \dots \partial x_n}. \quad (5.20)$$

En aquest cas

$$F_{X_1 \dots X_n}(x_1, \dots, x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f_{X_1 \dots X_n}(x_1, \dots, x_n) dx_1 \cdots dx_n. \quad (5.21)$$

Si A és una regió de \mathbb{R}^n

$$P((X_1, X_2, \dots, X_n) \in A) = \int \cdots \int_A f_{X_1 X_2 \dots X_n}(x_1, x_2, \dots, x_n) dx_1 dx_2 \cdots dx_n. \quad (5.22)$$

Les distribucions marginals s'obtenen integrant a tot \mathbb{R} algunes de les n variables. Per exemple, si $k < n$

$$f_{X_1 \dots X_k}(x_1, \dots, x_k) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{X_1 \dots X_n}(x_1, \dots, x_n) dx_{k+1} \cdots dx_n. \quad (5.23)$$

Les condicionals s'obtenen dividint la funció de densitat conjunta per la funció de densitat marginal de la condició. Per exemple,

$$f_{X_1 X_2 \dots X_k}(x_1, x_2, \dots, x_k \mid x_{k+1}, \dots, x_n) = \frac{f_{X_1 X_2 \dots X_n}(x_1, x_2, \dots, x_n)}{f_{X_{k+1} \dots X_n}(x_{k+1}, \dots, x_n)}. \quad (5.24)$$

Les variables X_1, X_2, \dots, X_n són independents si

$$f_{X_1 X_2 \dots X_n}(x_1, x_2, \dots, x_n) = f_{X_1}(x_1) f_{X_2}(x_2) \cdots f_{X_n}(x_n). \quad (5.25)$$

En el cas discret existeixen definicions anàlogues per les funcions de probabilitat.

Capítol 6

Funcions de vàries variables aleatòries

Considerem una variable aleatòria bidimensional contínua (X, Y) amb funció de densitat conjunta f_{XY} . Una funció $Z = g(X, Y)$ defineix una nova variable aleatòria que volem estudiar.

Definim les regions

$$\phi_z = \{(x, y) \in \mathbb{R}^2 \mid g(x, y) \leq z\}. \quad (6.1)$$

Llavors, la funció de distribució $F_Z(z) = P(Z \leq z)$ és

$$F_Z(z) = \int \int_{\phi_z} f_{XY}(x, y) dx dy. \quad (6.2)$$

A partir d'aquí obtenim la funció de densitat $f_Z(z) = \frac{dF_Z(z)}{dz}$.

6.1 Suma de variables aleatòries. Teorema de convolució

Partim de X i Y v.a.'s independents i posem $Z = X + Y$. Llavors, si $*$ designa el producte de convolució,

$$f_Z = f_X * f_Y. \quad (6.3)$$

DEM:

$$\begin{aligned} F_Z(z) &= \int \int_{x+y \leq z} f_{XY}(x, y) dx dy = \int \int_{x+y \leq z} f_X(x) f_Y(y) dx dy \\ &= \int_{-\infty}^{\infty} f_X(x) \left(\int_{-\infty}^{z-x} f_Y(y) dy \right) dx. \end{aligned}$$

Derivant respecte al paràmetre z trobem

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x) dx = (f_X * f_Y)(z). \clubsuit$$

Exemple 6.1 Suma de variables exponencials independents. X, Y són variables aleatòries exponencials de paràmetre λ_1, λ_2 respectivament. $Z = X + Y$.

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{\infty} f_X(x) f_Y(z-x) dx = \int_0^z \lambda_1 e^{-\lambda_1 x} \lambda_2 e^{-\lambda_2(z-x)} dx. \\ &= \lambda_1 \lambda_2 e^{-\lambda_2 z} \int_0^z e^{-(\lambda_1 - \lambda_2)x} dx = \frac{\lambda_1 \lambda_2}{\lambda_1 - \lambda_2} (e^{-\lambda_2 z} - e^{-\lambda_1 z}). \end{aligned}$$

Els límits d'integració $0, z$ apareixen degut a que la densitat exponencial és nul·la per argument negatiu ($x < 0, z - x < 0$). Igualment, la densitat anterior val 0 per $z < 0$.

El resultat és singular quan $\lambda_1 = \lambda_2$. En aquest cas recalculem i trobem que (per $z > 0$)

$$f_Z(z) = \lambda^2 z e^{-\lambda z}. \blacklozenge$$

Exemple 6.2 Suma de variables de Poisson independents. En el cas discret tenim que, si $Z = g(X, Y)$,

$$P_Z(z_i) = \sum_{g(x_k, y_l)=z_i} P(x_k, y_l). \quad (6.4)$$

Si X i Y són variables de Poisson independents de paràmetres α_1 i α_2 respectivament i $Z = X + Y$, llavors $\Omega_Z = \Omega_X = \Omega_Y = \{0, 1, \dots\}$ i

$$\begin{aligned} P_Z(n) &= \sum_{k+l=n} e^{-\alpha_1} \frac{\alpha_1^k}{k!} e^{-\alpha_2} \frac{\alpha_2^l}{l!} = e^{-(\alpha_1+\alpha_2)} \sum_{k=0}^n \frac{1}{k!(n-k)!} \alpha_1^k \alpha_2^{n-k} \\ &= e^{-(\alpha_1+\alpha_2)} \frac{(\alpha_1 + \alpha_2)^n}{n!}. \end{aligned}$$

Per tant, la suma de variables de Poisson independents de paràmetres α_1 i α_2 és una variable de Poisson de paràmetre $\alpha_1 + \alpha_2$.

Aquest fet resulta intuïtiu ja que si les variables inicials són el nombre d'esdeveniments en intervals temporals T_1 i T_2 , llavors la seva suma representa el nombre d'esdeveniments en un interval $T_1 + T_2$ (recordem $\alpha = \lambda T$ amb λ dependent només del tipus de fenomen). ♦

Exemple 6.3 Suma de variables normals independents. Prenem

$$f_X(x) = \frac{e^{-\frac{(x-m_1)^2}{2\sigma_1^2}}}{\sqrt{2\pi}\sigma_1}, \quad f_Y(y) = \frac{e^{-\frac{(y-m_2)^2}{2\sigma_2^2}}}{\sqrt{2\pi}\sigma_2}.$$

Llavors, utilitzant (6.3),

$$\begin{aligned} f_{X+Y}(z) &= \int_{-\infty}^{\infty} \frac{e^{-\frac{(x-m_1)^2}{2\sigma_1^2}}}{\sqrt{2\pi}\sigma_1} \frac{e^{-\frac{(z-x-m_2)^2}{2\sigma_2^2}}}{\sqrt{2\pi}\sigma_2} dx \\ &= \frac{1}{2\pi\sigma_1\sigma_2} \int_{-\infty}^{\infty} dx e^{-\frac{1}{2}\left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\right)x^2 + \left(\frac{m_1}{\sigma_1^2} + \frac{z-m_2}{\sigma_2^2}\right)x - \left(\frac{m_1^2}{2\sigma_1^2} + \frac{(z-m_2)^2}{2\sigma_2^2}\right)}. \end{aligned}$$

Considerant la integral gaussiana general (D.5) $\int_{-\infty}^{\infty} dx e^{-\frac{a}{2}x^2 + bx + c} = \sqrt{\frac{2\pi}{a}} e^{\frac{b^2}{2a} + c}$ i simplificant les expressions que surten, arribem a

$$f_{X+Y}(z) = \frac{1}{\sqrt{2\pi(\sigma_1^2 + \sigma_2^2)}} e^{-\frac{(z-(m_1+m_2))^2}{2(\sigma_1^2 + \sigma_2^2)}}.$$

Resulta, per tant, que $X + Y$ és una variable gaussiana d'esperança $m = m_1 + m_2$ i variància $\sigma^2 = \sigma_1^2 + \sigma_2^2$. Donades constants α i β , si X i Y són normals, αX i βY també ho són. Llavors $\alpha X + \beta Y$ és una v.a. normal amb esperança $\alpha m_1 + \beta m_2$ i variància $\alpha^2 \sigma_1^2 + \beta^2 \sigma_2^2$. ♦

Exemple 6.4 L'alçada dels homes i de les dones són v.a.'s normals amb esperances 175cm i 160cm, i desviacions típiques 10cm i 9cm respectivament. Es tria un home i una dona a l'atzar. Quina és la probabilitat que l'home sigui més alt que la dona?

Disseminem H i D aquestes variables. $Z = H - D$ és una variable normal d'esperança $175 - 160 = 15$ cm i variància $10^2 + 9^2 = 181$. La probabilitat que volem és $P(Z \geq 0) = 1 - F_Z(0) = 1 - \frac{1}{2}(1 + \operatorname{erf}\left(-\frac{15}{\sqrt{2}\sqrt{181}}\right)) = 0,5 + 0,5 \operatorname{erf}(0,788) = 0,8675$. ♦

6.2 Altres funcions. Canvi de variables

Considerem ara la distribució conjunta de dues variables Z i T construïdes a partir de X i Y .

$$\begin{cases} Z = g(X, Y) \\ T = h(X, Y) \end{cases} \quad (6.5)$$

Quan la transformació és injectiva diem que estem fent un canvi de variable. En aquest cas pensem que tenim la mateixa distribució de probabilitat expressada en variables diferents.

Recordem la definició del determinant jacobí

$$\frac{\partial(x, y)}{\partial(z, t)} = \begin{vmatrix} \frac{\partial x}{\partial z} & \frac{\partial x}{\partial t} \\ \frac{\partial y}{\partial z} & \frac{\partial y}{\partial t} \end{vmatrix}. \quad (6.6)$$

Per calcular aquesta expressió s'ha d'invertir (6.5) prèviament. També es pot calcular considerant que

$$\frac{\partial(x, y)}{\partial(z, t)} = \left(\frac{\partial(z, t)}{\partial(x, y)} \right)^{-1}. \quad (6.7)$$

En qualsevol cas voldrem expressar-ho en termes de les variables z i t . El seu valor absolut ens dóna el canvi de l'element de mesura.

$$dxdy = \left| \frac{\partial(x, y)}{\partial(z, t)} \right| dzdt. \quad (6.8)$$

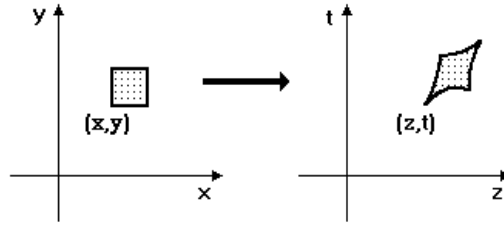


Figura 13: Canvi de l'element de mesura.

Quan la transformació és injectiva f_{ZT} és nul·la en els punts que no són de la imatge. Pels punts de la imatge considerem la probabilitat de la regió infinitesimal indicada en la figura.

$$P(\square) = f_{XY}(x, y)dxdy = f_{ZT}(z, t)dzdt.$$

Utilitzant (6.8) trobem

$$f_{ZT}(z, t) = f_{XY}(x(z, t), y(z, t)) \left| \frac{\partial(x, y)}{\partial(z, t)} \right|. \quad (6.9)$$

Si la transformació no fos injectiva en aquesta expressió hauríem de sumar el terme de la dreta sobre tots els punts tals que $(x_i, y_i) \longrightarrow (z, t)$.

A vegades es consideren canvis de variable que són singulars en algun punt. Això no és cap inconvenient ja que trobem probabilitats integrant sobre regions finites i aquest resultat no depèn del que li passi a la densitat en un punt o corba aïllada.

Exemple 6.5

$$\begin{cases} Z = X + Y \\ T = X - Y \end{cases}$$

És una transformació bijectiva de \mathbb{R}^2 (lineal, no singular). La inversa és $x = \frac{z+t}{2}, y = \frac{z-t}{2}$.

$$\frac{\partial(x, y)}{\partial(z, t)} = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{vmatrix} = -\frac{1}{2}.$$

Per tant,

$$f_{ZT}(z, t) = f_{XY}\left(\frac{z+t}{2}, \frac{z-t}{2}\right) \frac{1}{2}. \blacklozenge$$

Exemple 6.6 Coordenades polars. Són $R = \sqrt{X^2 + Y^2}$ i $\Theta = \arctan \frac{Y}{X}$. Varien $0 \leq r < \infty$ i $0 \leq \theta \leq 2\pi$.

$$\begin{cases} x = r \cos \theta \\ y = r \sin \theta \end{cases} \quad (6.10)$$

$$\frac{\partial(x, y)}{\partial(r, \theta)} = \begin{vmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{vmatrix} = r.$$

Per tant,

$$f(r, \theta) = f_{XY}(r \cos \theta, r \sin \theta) r. \blacklozenge$$

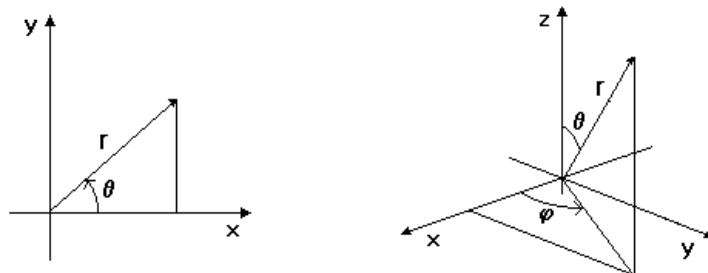


Figura 14: Coordenades polars i esfèriques.

Exemple 6.7 Coordenades esfèriques. Partim d'una variable tridimensional (X, Y, Z) i fem el canvi

$$\begin{cases} x = r \sin \theta \cos \varphi \\ y = r \sin \theta \sin \varphi \\ z = r \cos \theta \end{cases} \quad (6.11)$$

$$0 \leq r < \infty, 0 \leq \theta \leq \pi, 0 \leq \varphi \leq 2\pi.$$

$$\frac{\partial(x, y, z)}{\partial(r, \theta, \varphi)} = \begin{vmatrix} \sin \theta \cos \varphi & r \cos \theta \cos \varphi & -r \sin \theta \sin \varphi \\ \sin \theta \sin \varphi & r \cos \theta \sin \varphi & r \sin \theta \cos \varphi \\ \cos \theta & -r \sin \theta & 0 \end{vmatrix} = r^2 \sin \theta.$$

Per tant, la densitat de la v.a. (R, Θ, Φ) és

$$f(r, \theta, \varphi) = f_{XYZ}(r \sin \theta \cos \varphi, r \sin \theta \sin \varphi, r \cos \theta) r^2 \sin \theta.$$

Apliquem això a la distribució uniforme en una esfera de radi a . Sobre aquesta esfera $f(x, y, z) = \frac{1}{\frac{4}{3}\pi a^3}$. En coordenades polars tenim

$$f(r, \theta, \varphi) = \frac{r^2 \sin \theta}{\frac{4}{3}\pi a^3}$$

per $0 \leq r \leq a$ Calculem ara la densitat marginal de R :

$$f(r) = \int_0^{2\pi} \int_0^\pi \frac{r^2 \sin \theta}{\frac{4}{3}\pi a^3} d\varphi d\theta = 3 \frac{r^2}{a^3}.$$

Aquesta densitat es pot calcular també de la següent manera. La funció de distribució de R és $F(r) = P(R \leq r) = \frac{r^3}{a^3}$, quocient entre els volums de l'esfera de radi r i de tota l'esfera. Per tant $f(r) = dF(r)/dr = 3 \frac{r^2}{a^3}$. ♦

Exemple 6.8 Coordenades cilíndriques. En el cas tridimensional és el canvi

$$\begin{cases} x = r \cos \varphi \\ y = r \sin \varphi \\ z = z \end{cases} \quad (6.12)$$

$$0 \leq r < \infty, 0 \leq \varphi \leq 2\pi, -\infty < z < \infty.$$

$$\frac{\partial(x, y, z)}{\partial(r, \varphi, z)} = \begin{vmatrix} \cos \varphi & -r \sin \varphi & 0 \\ \sin \varphi & r \cos \varphi & 0 \\ 0 & 0 & 1 \end{vmatrix} = r. \quad \blacklozenge$$

Els canvis de variable es fan segons les simetries que tingui la distribució que estem utilitzant. Quan es fa un canvi cal vigilar no només com varia la forma funcional de la densitat sinó també com es transformen les possibles regions de frontera que tinguem. El canvi que hem fet en la uniforme sobre una esfera no hagués servit de res en la uniforme sobre un cub ja que la frontera del cub te una expressió complicada en esfèriques. En aquest cas les variables cartesianes són les més naturals.

Algun tipus de funcions pot requerir un tractament una mica particular com es veu en el següent exemple.

Exemple 6.9 Funcions màxim i mínim

Donades dues variables contínues independents X i Y es tracta de trobar les propietats marginals i conjunts de les variables $U = \min(X, Y)$ i $V = \max(X, Y)$.

Comencem per les marginals.

$$\begin{aligned} F_U(u) &= P(U \leq u) = P(\min(X, Y) \leq u) = P(X \leq u \text{ o } Y \leq u) = \\ &= P(X \leq u) + P(Y \leq u) - P(X \leq u \text{ i } Y \leq u) = F_X(u) + F_Y(u) - F_X(u)F_Y(u) \\ F_V(v) &= P(V \leq v) = P(\max(X, Y) \leq v) = P(X \leq v \text{ i } Y \leq v) \\ &= P(X \leq v)P(Y \leq v) = F_X(v)F_Y(v) \end{aligned}$$

Les densitats marginals de U i de V s'obtenen derivant les anteriors expressions.

La funció de distribució conjunta és

$$F_{UV}(u, v) = P(U \leq u, V \leq v) = P(\min(X, Y) \leq u, \max(X, Y) \leq v)$$

que val 0 si $u > v$. Si $u < v$, notem que $\max(X, Y) \leq v$ correspon a la regió $(-\infty, v] \times (-\infty, v]$ i $\min(X, Y) \leq u$ a $\mathbb{R}^2 - (u, \infty) \times (u, \infty)$ de manera que

$$F_{UV}(u, v) = P((-\infty, v] \times (-\infty, v]) - P((u, v] \times (u, v])$$

$$\begin{aligned}
&= F_{XY}(v, v) - (F_{XY}(v, v) - F_{XY}(u, v) - F_{XY}(v, u) + F_{XY}(u, u)) \\
&= F_{XY}(u, v) + F_{XY}(v, u) - F_{XY}(u, u).
\end{aligned}$$

Derivant arribem a

$$f_{UV}(u, v) = \begin{cases} f_{XY}(u, v) + f_{XY}(v, u) & \text{si } u < v \\ 0 & \text{si } u > v \end{cases}$$

El resultat anterior val encara que X, Y no siguin independents. Si ho són, $f_{UV}(u, v) = f_X(u)f_Y(v) + f_X(v)f_Y(u)$ per $u < v$. ♦

Capítol 7

Paràmetres estadístics de vàries v.a.'s

7.1 Esperança

Generalitzen el teorema de l'esperança unidimensional. Donades dues v.a.'s X i Y amb funció de densitat conjunta $f(x, y)$ considerem la nova v.a. $Z = g(X, Y)$. Llavors l'esperança de Z verifica

$$E[Z] = E[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f(x, y) dx dy. \quad (7.1)$$

DEM:

Considerem una partició de \mathbb{R} en segments de longitud Δz_i . Aquesta porta associada una partició de \mathbb{R}^2 en regions $\Delta D_{z_i} = \phi_{z_i+\Delta z_i} - \phi_{z_i}$ (veure (6.1)) i tenim que

$$P(z_i \leq Z \leq z_i + \Delta z_i) = \int \int_{\Delta D_{z_i}} f(x, y) dx dy.$$

Llavors

$$\begin{aligned} E[Z] &= \int_{-\infty}^{\infty} z f_Z(z) dz = \lim \sum_{z_i} z_i P(z_i \leq Z \leq z_i + \Delta z_i) \\ &= \lim \sum_{\Delta D_{z_i}} z_i \int \int_{\Delta D_{z_i}} f(x, y) dx dy = \lim \sum_{\Delta D_{z_i}} \int \int_{\Delta D_{z_i}} g(x, y) f(x, y) dx dy \\ &= \int \int_{\mathbb{R}^2} g(x, y) f(x, y) dx dy. \clubsuit \end{aligned}$$

De (7.1) veiem que si $\alpha, \beta \in \mathbb{R}$

$$E[\alpha g(X, Y) + \beta h(X, Y)] = \alpha E[g(X, Y)] + \beta E[h(X, Y)]. \quad (7.2)$$

En particular

$$E[X + Y] = E[X] + E[Y]. \quad (7.3)$$

7.2 Covariància

La covariància de les variables X i Y és el nombre:

$$C[X, Y] = E[(X - \overline{X})(Y - \overline{Y})]. \quad (7.4)$$

Desenvolupant el producte i utilitzant (7.2) trobem que

$$C[X, Y] = E[XY] - \overline{X}\overline{Y}. \quad (7.5)$$

7.3 Coeficient de correlació

$$\rho = \frac{C[X, Y]}{\sigma_X \sigma_Y}. \quad (7.6)$$

Verifica

$$|\rho| \leq 1. \quad (7.7)$$

DEM:

Sigui λ un paràmetre real. Llavors

$$E[(\lambda(X - \bar{X}) + (Y - \bar{Y}))^2] = \lambda^2 \sigma_X^2 + 2\lambda C[X, Y] + \sigma_Y^2 \geq 0$$

Com el polinomi en λ és sempre positiu, el discriminant $C[X, Y]^2 - \sigma_X^2 \sigma_Y^2 \leq 0$. Per tant, $|C[X, Y]| \leq \sigma_X \sigma_Y$ d'on surt (7.7). (Aquesta demostració es fa de la desigualtat de Cauchy-Schwartz per productes escalars.) ♣

Exemple 7.1 Matriu de covariàncies.

Calcular els paràmetres de d'una variable bidimensional uniforme en la regió $\mathcal{R} = \{(x, y) \in \mathbb{R}^2 | x \geq 0, y \geq 0, 2x + y \leq 2\}$.

Com la regió és un triangle d'àrea 1, $f_{XY}(x, y) = 1$ en \mathcal{R} . Començarem per calcular les densitats marginals:

$$f_X(x) = \int_0^{2-2x} 1 dy = 2(1-x), \quad 0 \leq x \leq 1, \quad f_Y(y) = \int_0^{\frac{2-y}{2}} 1 dx = \frac{2-y}{2}, \quad 0 \leq y \leq 2.$$

Ara és fàcil calcular tots els moments de X i de Y

$$m_{X,n} = \int_0^1 x^n \cdot 2(1-x) dx = \frac{2}{(n+1)(n+2)}, \quad m_{Y,n} = \int_0^2 y^n \cdot \frac{2-y}{2} dy = \frac{2^{n+1}}{(n+1)(n+2)}.$$

D'aquí traiem $m_X = \frac{1}{3}$, $\sigma_X^2 = \frac{1}{6} - (\frac{1}{3})^2 = \frac{1}{18}$, $m_Y = \frac{2}{3}$, $\sigma_Y^2 = \frac{2}{3} - (\frac{2}{3})^2 = \frac{2}{9}$. Calculem també

$$E[XY] = \int_0^1 \int_0^{2-2x} xy \cdot 1 dy dx = 2 \int_0^1 x \cdot (1-x)^2 dx = \frac{1}{6}$$

d'on $C[X, Y] = \frac{1}{6} - \frac{1}{3} \cdot \frac{2}{3} = -\frac{1}{18}$, $\rho = -\frac{1}{2}$.

Per qualsevol variable multidimensional (X_1, X_2, \dots, X_n) els paràmetres de primer i segon grau es poden organitzar en forma matricial amb un *vector d'esperances* $\mathbf{m} = (m_{X_1}, m_{X_2}, \dots, m_{X_n})$ i una *matriu de covariàncies* \mathbf{C} definida per $C_{i,j} = C[X_i, X_j]$.

En el nostre cas $\mathbf{m} = (m_X, m_Y) = (\frac{1}{3}, \frac{2}{3})$ i

$$\mathbf{C} = \begin{pmatrix} \sigma_X^2 & \rho \sigma_X \sigma_Y \\ \rho \sigma_X \sigma_Y & \sigma_Y^2 \end{pmatrix} = \begin{pmatrix} \frac{1}{18} & -\frac{1}{18} \\ -\frac{1}{18} & \frac{2}{9} \end{pmatrix}. \quad \blacklozenge$$

7.4 Moments conjunts

Els **moments conjunts** de les variables X i Y són

$$m_{jk} = E[X^j Y^k] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^j y^k f(x, y) dx dy. \quad (7.8)$$

Els primers són $m_{00} = 1$, $m_{10} = \bar{X}$, $m_{01} = \bar{Y}$, etc.

Els **moments centrals** són

$$\mu_{jk} = E[(X - \bar{X})^j (Y - \bar{Y})^k] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - m_X)^j (y - m_Y)^k f(x, y) dx dy. \quad (7.9)$$

Els primers són $\mu_{00} = 1$, $\mu_{10} = \mu_{01} = 0$, $\mu_{11} = C[X, Y]$, $\mu_{20} = \sigma_X^2$, $\mu_{02} = \sigma_Y^2$, etc.

7.5 Ortogonalitat

Diem que X i Y són variables aleatòries **ortogonals** si $E[XY] = 0$. En aquest cas escrivim $X \perp Y$.

7.6 Incorrelació

Diem que X i Y són variables aleatòries **incorrelades** si $\rho = 0$. És equivalent a dir $C[X, Y] = 0$ o $E[XY] = \overline{X}\overline{Y}$.

Es donen les següents relacions

- X i Y incorrelades $\Leftrightarrow (X - \overline{X}) \perp (Y - \overline{Y})$.
- X i Y incorrelades $\Leftrightarrow V[X + Y] = V[X] + V[Y]$.
- X i Y independents $\Rightarrow X$ i Y incorrelades.

DEM:

La primera és evident a partir de (7.4). La segona prové de la relació general

$$V[X + Y] = E[(X + Y - (\overline{X} + \overline{Y}))^2] = E[(X - \overline{X})^2 + (Y - \overline{Y})^2 + 2(X - \overline{X})(Y - \overline{Y})] = V[X] + V[Y] + 2C[X, Y].$$

Per la tercera, si X i Y són independents $f(x, y) = f_X(x)f_Y(y)$ i

$$\begin{aligned} E[XY] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyf(x, y)dxdy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyf_X(x)f_Y(y)dxdy \\ &= \int_{-\infty}^{\infty} xf_X(x)dx \int_{-\infty}^{\infty} yf_Y(y)dy = \overline{X}\overline{Y}. \clubsuit \end{aligned}$$

Amb una demostració anàloga veiem que en el cas de variables independents $E[g(X)h(Y)] = E[g(X)]E[h(Y)]$. Per variables incorrelades això només es verifica en general per g i h iguals a la funció identitat. Un cas especial en són les variables gaussianes per les que ja s'ha vist que si són incorrelades llavors són independents.

Exemple 7.2 Ja s'han vist alguns resultats sobre variables gaussianes. Completam ara el seu estudi.

Primer, demostrem que, efectivament, la constant ρ que apareix en la densitat és el coeficient de correlació. Per evitar calcular integrals bidimensionals utilitzem els resultats de l'exemple 5.7:

$$E[X | y] = m_1 + \rho \frac{\sigma_1}{\sigma_2}(y - m_2). \quad (7.10)$$

$$\begin{aligned} E[XY] &= E[E[XY | Y]] = E[E[X | Y]Y] = E[(m_1 + \rho \frac{\sigma_1}{\sigma_2}(Y - m_2))Y] \\ &= (m_1 - \rho \frac{\sigma_1}{\sigma_2}m_2)E[Y] + \rho \frac{\sigma_1}{\sigma_2}E[Y^2] = (m_1 - \rho \frac{\sigma_1}{\sigma_2}m_2)m_2 + \rho \frac{\sigma_1}{\sigma_2}(\sigma_2^2 + m_2^2) = m_1m_2 + \rho\sigma_1\sigma_2 \end{aligned}$$

d'on s'obté immediatament el resultat.

Ara, trobem una parametrització convenient per una variable bidimensional gaussiana arbitrària. És immediat veure que si X és normal, qualsevol variable de la forma $aX + b$ és normal. Si X i Y són variables gaussianes amb esperances m_X, m_Y i variàncies σ_X, σ_Y , llavors les variables $\frac{X-m_X}{\sigma_X}, \frac{Y-m_Y}{\sigma_Y}$ són normals amb esperança 0 i variància 1. Considerem les noves variables

$$\begin{cases} Z = \frac{X-m_X}{\sigma_X} + \frac{Y-m_Y}{\sigma_Y} \\ T = \frac{X-m_X}{\sigma_X} - \frac{Y-m_Y}{\sigma_Y} \end{cases}$$

que són normals d'esperança 0 i verifiquen

$$V[Z] = E[Z^2] = E\left[\left(\frac{X - m_X}{\sigma_X}\right)^2\right] + E\left[\left(\frac{Y - m_Y}{\sigma_Y}\right)^2\right] + 2E\left[\frac{(X - m_X)(Y - m_Y)}{\sigma_X \sigma_Y}\right] = 2(1 + \rho),$$

$$V[T] = 2(1 - \rho),$$

$$E[ZT] = E\left[\left(\frac{X - m_X}{\sigma_X}\right)^2\right] - E\left[\left(\frac{Y - m_Y}{\sigma_Y}\right)^2\right] = 0,$$

d'on veiem que són incorrelades. Així, unes coordenades especialment útils per treballar amb la variable bidimensional (X, Y) són

$$\begin{cases} U = \frac{1}{\sqrt{2(1+\rho)}} \left(\frac{X - m_X}{\sigma_X} + \frac{Y - m_Y}{\sigma_Y} \right) \\ V = \frac{1}{\sqrt{2(1-\rho)}} \left(\frac{X - m_X}{\sigma_X} - \frac{Y - m_Y}{\sigma_Y} \right) \end{cases}$$

U, V són variables normals estandarditzades independents.

Notem que

$$f_{UV}(u, v) = f_U(u)f_V(v) = \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}} \frac{e^{-\frac{v^2}{2}}}{\sqrt{2\pi}} = \frac{1}{2\pi} e^{-\frac{u^2+v^2}{2}}.$$

A l'efectuar el pas a les variables X, Y es recupera la densitat normal (5.11).♦

Capítol 8

Estimació de variables aleatòries

8.1 Concepte i criteris d'estimació

Considerem un experiment que té associada una variable aleatòria Y . Si repetim un cert nombre N de vegades aquest experiment obtenim els resultats Y_1, Y_2, \dots, Y_N . Estimar la variable Y vol dir fer una predicció que s'acosti el màxim possible a aquests resultats.

Suposem que la predicció resulta ser c_1, c_2, \dots, c_N . L'error comès vindrà donat per algun tipus de “distància” $d(Y_i, c_i)$. Globalment posarem $\epsilon = \frac{1}{N} \sum_i d(Y_i, c_i)$. Les prediccions s'haurien de triar de manera que ϵ fos el més petit possible.

Es poden considerar diferents tipus d'error com, per exemple, $d(Y_i, c_i) = |Y_i - c_i|$. Quan elegim $d(Y_i, c_i) = (Y_i - c_i)^2$ diem que estem fent l'**estimació en mitjana quadràtica**. Aquesta és la que considerarem en el que segueix.

Donat que les condicions de l'experiment no varien, les variables Y_i són distribuïdes idènticament a la variable Y . Llavors, si els elements c_i són nombres determinats, aquests hauran de ser sempre el mateix, $c_i = c$ per a tot i .

Per una altra part, pot ser que el nostre experiment tingui associades altres variables aleatòries X_1, X_2, \dots, X_r de les quals Y no és independent. En aquest cas és convenient estimar Y amb una variable aleatòria de la forma $c(X_1, X_2, \dots, X_r)$. c_i seria el resultat i -èsim d'aquesta variable.

Com l'error ϵ depèn de variables aleatòries, el seu valor no està fixat. La magnitud que finalment ens dona la bondat de l'estimació és l'**error quadràtic mig**:

$$\bar{\epsilon} = E[\epsilon] = \frac{1}{N} \sum_i E[(Y_i - c_i)^2] = E[(Y - c)^2]. \quad (8.1)$$

El procediment a seguir per estimar la v.a. Y és fixar un tipus d'estimació c i minimitzar $\bar{\epsilon}$ per determinar del tot c . L'error associat a l'estimació serà finalment $\bar{\epsilon}_{min}$. c s'anomena **estimador**.

Exemple 8.1 L'experiment consisteix en seleccionar una persona a l'atzar. La variable Y que ens interessa estimar és l'alçada d'aquesta persona. La seqüència de valors Y_i consisteix en mesurar les alçades de varies persones triades a l'atzar, això és, fer un mostreig estadístic. S'entén que no coneixem la distribució que segueix Y però notem que al fer-ne un mostreig podem avaluar amb prou precisió els seus paràmetres com l'esperança o la desviació típica. Les variables X_i podrien correspondre al sexe, edat, pes, color dels ulls, etc. La idea és avaluar l'alçada de la persona coneixent, per exemple, el seu pes. ♦

Darrera l'estimació hi han dos possibles objectius. Un és fer realment una predicció d'un resultat abans que aquest es produeixi. Per exemple, predir cotitzacions a la borsa. En aquest cas les variables X_i serien els indicadors econòmics que si que coneixem. L'altre objectiu, relacionat

amb l'anterior, és establir relacions entre magnituds aleatòries. És clar que no hi ha una relació funcional fixa entre el pes i l'alçada de les persones. El màxim que podem aconseguir és tenir una bona estimació d'una magnitud coneguda l'altra. A vegades el problema és trobar variables apropiades per fer l'estimació. Per exemple, en medicina quan es vol establir l'origen o els factors de risc d'una malaltia.

8.2 Estimació per una constant

Estimem Y per c constant. La millor estimació és $c = E[Y]$ i $\bar{\epsilon}_{min} = V[Y]$.

DEM:

$$\begin{aligned}\bar{\epsilon} &= E[(Y - c)^2] = c^2 - 2E[Y]c + E[Y^2] = (c - E[Y])^2 + E[Y^2] - E[Y]^2 \\ &= (c - E[Y])^2 + V[Y],\end{aligned}$$

que es minimitza quan $c - E[Y] = 0$. ♣

8.3 Estimació no lineal

Estimem Y per $c(X)$ funció a determinar de la v.a. X . La millor estimació és $c(x) = E[Y | x]$.

DEM:

$$\begin{aligned}\bar{\epsilon} &= \int_{\mathbb{R}^2} (y - c(x))^2 f(x, y) dx dy = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} (y - c(x))^2 f(y | x) dy \right) f(x) dx \\ &= \int_{-\infty}^{\infty} (c(x)^2 - 2E[Y | x]c(x) + E[Y^2 | x]) f(x) dx \\ &= \int_{-\infty}^{\infty} ((c(x) - E[Y | x])^2 + V[Y | x]) f(x) dx.\end{aligned}$$

Com $f(x) \geq 0$ la integral es minimitza si $c(x) - E[Y | x] = 0$. L'error mínim val $\int_{-\infty}^{\infty} V[Y | x] f(x) dx$. ♣

La funció $c(x) = E[Y | x]$ s'anomena **corba de regressió**. Notem que si intentem mesurar una relació fixada entre dues variables però el que s'obté és una v.a. bidimensional degut als errors experimentals, la corba de regressió seria el que prendríem com relació entre les dues variables. En aquest cas la corba de regressió fa tornar (regressar) els punts que s'han mesurat al seu lloc correcte.

Hi han dos casos extrems de l'estimació no lineal:

- X i Y són independents. Llavors $E[Y | x] = E[Y]$ i $c(x) = E[Y]$ que és una constant.
- $Y = g(X)$, és a dir, hi ha la màxima dependència possible. En aquest cas la funció de densitat conjunta es pot expressar $f(x, y) = f(x)\delta(y - g(x))$ i s'obté $E[Y | x] = g(x)$. Per tant, $c(x) = g(x)$ i $\bar{\epsilon}_{min} = 0$.

En el primer cas, la informació donada per X és irrellevant per estimar Y . Per exemple, tractar d'estimar l'alçada mitjançant el color dels ulls. En el segon, X determina completament Y . Per exemple, estimar l'edat a partir de l'any de naixement. En la pràctica, ens interessen els casos intermedis on X dona informació rellevant sobre Y però no la determina.

8.4 Estimació lineal

Estimem Y per $aX + b$ on a i b són constants. La millor estimació ve donada per

$$a = \rho \frac{\sigma_Y}{\sigma_X}, \quad b = m_Y - \rho \frac{\sigma_Y}{\sigma_X} m_X, \quad \bar{\epsilon}_{min} = \sigma_Y^2 (1 - \rho^2). \quad (8.2)$$

La recta $y = ax + b$ s'anomena recta de regressió.

DEM: $\bar{\epsilon} = E[(Y - (aX + b))^2] = E[((Y - aX) - b)^2]$. Així b ha de ser la millor estimació constant de $Y - aX$. Per tant $b = E[Y - aX] = m_Y - am_X$. Ara tenim

$$\begin{aligned}\bar{\epsilon} &= E[((Y - m_Y) - a(X - m_X))^2] = a^2 \sigma_X^2 - 2aC[X, Y] + \sigma_Y^2 \\ &= a^2 \sigma_X^2 - 2a\rho\sigma_X\sigma_Y + \sigma_Y^2 = (a\sigma_X - \rho\sigma_Y)^2 + \sigma_Y^2(1 - \rho^2),\end{aligned}$$

que es minimitza quan $a\sigma_X - \rho\sigma_Y = 0$. ♣

Un es pot preguntar per que fer l'estimació lineal si la no lineal dóna un error més petit. Un motiu és que l'estimació no lineal exigeix conèixer la distribució conjunta de X i Y mentre que la lineal fa servir només els paràmetres més bàsics. També cal notar que en el cas de variables normals l'estimació no lineal coincideix amb la lineal (veure (7.10)).

8.5 Principi d'ortogonalitat

Per un experiment fixat, el conjunt de variables aleatòries que podem considerar formen un espai vectorial ja que la suma de v.a.'s i el producte d'una v.a. per un nombre real són noves variables aleatòries. Donades Z i T v.a.'s podem definir l'operació

$$\langle Z, T \rangle = E[ZT]. \quad (8.3)$$

Aquesta operació és bilineal, simètrica, definida positiva i no degenerada. Constitueix, per tant, un producte escalar en aquest espai vectorial. De fet, aquesta construcció es incorrecta ja que una variable aleatòria pot no tenir esperança. Ens restringirem, per tant, a variables per les quals les operacions considerades existeixen. Això vol dir treballar en un espai vectorial més petit.

Si tota la informació generada per un experiment queda ben codificada en una variable n -dimensional (X_1, X_2, \dots, X_n) , l'espai de variables aleatòries es pot considerar definit com un espai de funcions $Z = g(X_1, X_2, \dots, X_n)$. La restricció que considerem són les de "quadrat mesurable", és a dir, aquelles per les que existeix $E[Z^2]$.

El producte escalar (8.3) ens defineix una norma $\|Z\| = \langle Z, Z \rangle^{1/2}$. L'error quadràtic que consideràvem és precisament $\bar{\epsilon} = \|Y - c(X)\|^2$. Per tant, podem usar tècniques dels espais euclidians quan $c(X)$ correspon a alguna varietat lineal.

La forma general de l'estimació lineal és

$$c = a_1 X_1 + a_2 X_2 + \dots + a_n X_n, \quad (8.4)$$

on a_i són constants i X_i són variables aleatòries. Quan considerem només una v.a. fonamental a part de Y l'estimació lineal pren la forma

$$c(X) = a_1 g_1(X) + a_2 g_2(X) + \dots + a_n g_n(X), \quad (8.5)$$

on g_i són funcions fixades.

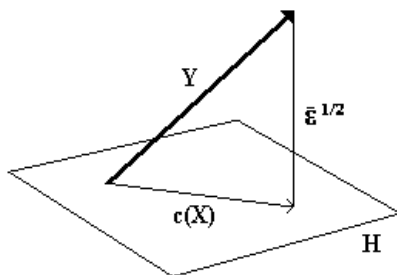


Figura 15: Estimació lineal.

Així, $c(X)$ descriu un hiperplà \mathcal{H} , la millor estimació és l'element d'aquest hiperplà més proper a Y i l'error és la distància (al quadrat) entre Y i \mathcal{H} .

Com se sap, la millor estimació ve donada per la projecció de Y sobre \mathcal{H} . Si aquesta projecció és $c(X)$ tindrem que $Y - c(X)$ és ortogonal a \mathcal{H} . Per tant, $(Y - c(X)) \perp g_i(X)$, $i = 1, \dots, n$. Hem arribat així al principi d'ortogonalitat ("l'error és ortogonal a les dades")

$$\langle g_i(X), Y - c(X) \rangle = 0, \quad i = 1, \dots, n. \quad (8.6)$$

Aquest és un sistema d'equacions lineals que podem escriure

$$E[g_i(X)c(X)] = E[g_i(X)Y], \quad i = 1, \dots, n. \quad (8.7)$$

La solució d'aquest sistema ens dona els valors de a_i .

Exemple 8.2 Tornem a deduir l'expressió de la recta de regressió. Posem $g_1(X) = X$, $g_2(X) = 1$. L'estimador és $c(X) = a_1X + a_2$. El principi d'ortogonalitat produeix les equacions

$$\begin{cases} \langle X, a_1X + a_2 \rangle = \langle X, Y \rangle \\ \langle 1, a_1X + a_2 \rangle = \langle 1, Y \rangle \end{cases}$$

és a dir,

$$\begin{cases} a_1E[X^2] + a_2E[X] = E[XY] \\ a_1E[X] + a_2 = E[Y] \end{cases}$$

De la segona equació traïem $a_2 = E[Y] - a_1E[X]$ i

$$a_1 = \frac{E[XY] - E[X]E[Y]}{E[X^2] - E[X]^2} = \frac{C[X, Y]}{\sigma_X^2} = \rho \frac{\sigma_Y}{\sigma_X},$$

que són els resultats (8.2).

Recordem que la projecció es pot calcular també a partir d'una base ortonormal. En aquest exemple podem prendre $X_1 = \frac{X - m_X}{\sigma_X}$, $X_2 = 1$ (notem que X_1 és la X normalitzada) amb el que la projecció és

$$c(X) = \langle X_1, Y \rangle X_1 + \langle X_2, Y \rangle X_2. \blacklozenge$$

Exemple 8.3 L'estimació de la forma $c(X) = aX$ s'anomena estimació homogènia. Per determinar a tenim l'equació $\langle X, aX \rangle = \langle X, Y \rangle$ d'on

$$a = \frac{E[XY]}{E[X^2]}. \quad (8.8)$$

De manera més general se sol anomenar **estimació no homogènia** a la de tipus $c = a_1X_1 + a_2X_2 + \dots + a_nX_n + b$ on b és una constant, si bé no és més que un cas de l'estimació lineal. \blacklozenge

Finalment cal notar que per calcular l'error mínim

$$\bar{\epsilon}_{min} = E[(Y - c)^2] = E[(Y - c)Y] - E[(Y - c)c] = E[(Y - c)Y]$$

ja que la segona esperança és nul·la (pel mateix principi d'ortogonalitat, $(Y - c) \perp c$). Notem també que $\bar{\epsilon}_{min} = E[Y^2] - E[c^2]$ (Pitàgores).

Part III

Processos estocàstics

Capítol 9

Introducció als processos estocàstics

9.1 Definició de procés estocàstic

Parlem de variables aleatòries quan el resultat d'un experiment aleatori és un o més nombres reals (v.a. unidimensional o multidimensional). Un **procés estocàstic** és quan el resultat de l'experiment és una funció.

Anomenem $X(t)$ a aquesta funció. És corrent, donat el tipus d'aplicacions que es consideren, interpretar t com el temps. Si t varia sobre tot un interval real (potser indefinit) diem que el procés és **a temps continu**. Si només considerem un conjunt numerable de valors t_i , la funció constitueix una seqüència $X_i = X(t_i)$ i diem que el procés és **a temps discret**.

Un procés a temps discret apareix quan es quantitza el temps en un procés a temps continu. En aquest cas, el temps entre valors consecutius de t_i acostuma a ser un valor Δt petit que s'anomena interval de mostreig. El valor X_i sol ser la mitjana del senyal continu $X(t)$ al voltant de t_i . En el que segueix, els processos són a temps continu mentre no es digui el contrari.

Fixat un valor t , $X(t)$ és una variable unidimensional ordinària. Com generalitzar el concepte de funció de densitat al cas en que els valors de la variable són funcions és complicat, s'acostuma a pensar un procés estocàstic com una v.a. que depèn d'un índex continu.

En moltes aplicacions el procés $X(t)$ representa algun tipus de senyal elèctric i, per tant, es representa més còmodament amb un nombre complex. Si $X(t)$ i $Y(t)$ son processos reals, diem que $Z(t) = X(t) + jY(t)$ és un procés estocàstic complex.

Un procés estocàstic (PE) també s'anomena procés aleatori.

Exemple 9.1 Moviment brownià. Al 1828 el botànic escocès Robert Brown, estudiant partícules de pol·len en suspensió en aigua, observà que aquestes feien un moviment molt irregular. Posteriorment es va veure que aquest moviment també el feia qualsevol tipus de partícula inorgànica i, per tant, no era degut al caràcter “viu” de la matèria. L'explicació del fenomen va haver d'esperar al 1905. A. Einstein va donar una explicació satisfactòria amb un model estadístic de les col·lisions amb la partícula de les molècules d'aigua en agitació tèrmica. El valor mitjà de la força realitzada sobre el gra en suspensió val zero però les fluctuacions provoquen un petit impuls instantani que dóna lloc al moviment irregular. La trajectòria de la partícula, $(X(t), Y(t), Z(t))$ és un procés estocàstic (en \mathbb{R}^3). ♦

Exemple 9.2 El consum elèctric d'una gran ciutat al llarg d'un dia és un procés $C(t)$, $0 \leq t \leq 24$. El seu caràcter aleatori prové del fet que és la suma de moltes petites contribucions aleatòries (els consums individuals). Des del punt de vista de la central que proveeix aquesta energia apareix un problema natural d'estimació, important per dimensionar correctament la producció: en un instant t_1 , que valdrà $C(t)$ per $t > t_1$ coneguts els seus valors per $t < t_1$? ♦

Exemple 9.3 Una manera senzilla d'obtenir un PE és prendre una funció ordinària i introduir-li paràmetres aleatoris. Per exemple, suposem que V i ω són v.a.'s unidimensionals. Llavors $X(t) = V \cos \omega t$ és un PE. ♦

Exemple 9.4 X és una v.a. uniforme en $[0, 1]$. El seu desenvolupament decimal $X = 0, b_1 b_2 b_3 \dots$ és un procés a temps discret b_i . ♦

Hem de pensar que, en general, les funcions resultants en un PE són molt irregulars. En aquest sentit l'exemple 9.3 resulta artificial. Encara que per fer càlculs il·lustratius considerarem funcions regulars que contenen paràmetres aleatoris com exemple de PE, les mostres més genuïnes venen donades pels tipus dels exemples 9.1 i 9.2.

9.2 Funcions de distribució i densitat d'un PE

Considerem un PE $X(t)$. Fixats els valors t_1, t_2, \dots, t_n tenim una variable aleatòria n -dimensional $(X(t_1), X(t_2), \dots, X(t_n))$ que anomenem **mostra** del procés. El procés queda totalment descrit si coneixem la distribució estadística d'aquestes variables per a tots els valors t_1, t_2, \dots, t_n i per a tot $n \geq 1$. (En rigor, això és cert només pels processos anomenats separables. Els exemples d'utilitat ho són tots.)

Les mostres d'un PE, i, per tant, tot el procés, queden caracteritzades per les següents funcions.

Les funcions de distribució d'ordre n :

$$F(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = P(X(t_1) \leq x_1, X(t_2) \leq x_2, \dots, X(t_n) \leq x_n). \quad (9.1)$$

Les funcions de densitat d'ordre n :

$$f(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = \frac{\partial^n F(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n)}{\partial x_1 \partial x_2 \dots \partial x_n}. \quad (9.2)$$

Són les funcions de distribució i densitat conjuntes de les v.a.'s $X(t_1), X(t_2), \dots, X(t_n)$. Les propietats que involucren les funcions d'ordre màxim n es diu que són d'ordre n .

La jerarquia de funcions (9.2) està lligada per relacions de consistència que estableixen que unes són les marginals d'altres d'ordre superior.

$$\begin{aligned} \int_{-\infty}^{\infty} f(x; t) dx &= 1, \\ f(x_1; t_1) &= \int_{-\infty}^{\infty} f(x_1, x_2; t_1, t_2) dx_2, \\ &\text{etc.} \end{aligned} \quad (9.3)$$

Les funcions (9.1) i (9.2) són simètriques sota permutacions dels índexs. Per tant, és suficient conèixer-les per $t_1 < t_2 < \dots < t_n$.

9.3 Valor mitjà, autocorrelació i autocovariància

Són propietats de primer i segon ordre.

Valor mitjà:

$$m(t) = E[X(t)] = \int_{-\infty}^{\infty} x f(x; t) dx. \quad (9.4)$$

Autocorrelació:

$$R(t_1, t_2) = E[X(t_1)X(t_2)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 f(x_1, x_2; t_1, t_2) dx_1 dx_2. \quad (9.5)$$

En el cas complex l'autocorrelació es defineix com $E[Z(t_1)Z^*(t_2)]$.

Autocovariància:

$$C(t_1, t_2) = R(t_1, t_2) - m(t_1)m(t_2). \quad (9.6)$$

Notem que C i R són funcions simètriques sota intercanvi $t_1 \leftrightarrow t_2$. Si $C(t_1, t_2) = 0$, sempre que $t_1 \neq t_2$, es diu que el procés és **incorelat** o de **soroll blanc**.

La **potència** mitjana és la funció

$$R(t, t) = E[X(t)^2]. \quad (9.7)$$

Exemple 9.5 Considerem el procés $X(t) = V \cos(\omega t + \phi)$ on ω està fixat, V és una v.a. exponencial de paràmetre $\lambda = \frac{1}{V_0}$ i ϕ és una v.a. uniforme en $[0, 2\pi]$, V i ϕ independents.

El seu valor mitjà és

$$m(t) = E[V \cos(\omega t + \phi)] = E[V]E[\cos(\omega t + \phi)] = V_0 \int_0^{2\pi} \cos(\omega t + \phi) \frac{d\phi}{2\pi} = 0.$$

L'autocorrelació és

$$\begin{aligned} R(t_1, t_2) &= E[V \cos(\omega t_1 + \phi) V \cos(\omega t_2 + \phi)] = E[V^2] E[\cos(\omega t_1 + \phi) \cos(\omega t_2 + \phi)] \\ &= 2V_0^2 \int_0^{2\pi} \cos(\omega t_1 + \phi) \cos(\omega t_2 + \phi) \frac{d\phi}{2\pi} \\ &= 2V_0^2 \int_0^{2\pi} \frac{1}{2} [\cos(\omega(t_1 + t_2) + 2\phi) + \cos(\omega(t_2 - t_1))] \frac{d\phi}{2\pi} \\ &= V_0^2 \cos(\omega(t_2 - t_1)). \end{aligned}$$

La potència mitjana és $R(t, t) = V_0^2$. L'autocovariància coincideix amb $R(t_1, t_2)$. ♦

9.4 Processos estocàstics normals

9.4.1 Variable gaussiana n -dimensional

Diem que la v.a. n -dimensional (X_1, X_2, \dots, X_n) és gaussiana o normal si té la següent funció de densitat:

$$f(x_1, x_2, \dots, x_n) = \frac{|A|^{\frac{1}{2}}}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2}(x-m)^T A (x-m)} \quad (9.8)$$

Notem que és l'exponencial d'un polinomi de segon grau en les variables x_1, x_2, \dots, x_n . Com notació hem posat $x = (x_1, x_2, \dots, x_n)$, $m = (m_1, m_2, \dots, m_n)$ i A és una matriu $n \times n$ (de fet representem els vectors com matrius columna de manera que entenem l'exponent com un producte matricial que dona un resultat escalar). $|A|$ és el determinant de la matriu A .

També es diu que les variables X_1, X_2, \dots, X_n són conjuntament normals. m i A són paràmetres de la distribució gaussiana. Determinem a continuació el seu significat. Per això és convenient fer un canvi de variable que simplifiqui l'expressió (9.8).

Existeixen v.a.'s Y_1, Y_2, \dots, Y_n gaussianes independents amb esperança 0 i variància 1, relacionades amb les X_i a través d'una transformació lineal no homògena.

DEM: Com l'exponent de (9.8) és $-\frac{1}{2} \sum_{i,j} A_{ij}(x_i - m_i)(x_j - m_j)$ podem prendre la matriu A simètrica. Perquè la densitat sigui normalitzable A ha de ser definida positiva. Com se sap de la teoria d'operadors en espais euclidians, tota matriu simètrica té tots els seus valors propis reals i diagonalitza en una base ortonormal. Així, existeix una matriu diagonal $A_d = \text{diag}(a_1, a_2, \dots, a_n)$ i una matriu ortogonal de canvi de base P ($P^{-1} = P^T$) tals que

$$A_d = PAP^T.$$

Els valors propis són $a_i > 0$ ja que A és definida positiva. Així, podem considerar la matriu $A_d^{\frac{1}{2}} = \text{diag}(\sqrt{a_1}, \sqrt{a_2}, \dots, \sqrt{a_n})$. Definint $B = A_d^{\frac{1}{2}}P$ tenim que

$$A = P^T A_d P = P^T A_d^{\frac{1}{2}} A_d^{\frac{1}{2}} P = (A_d^{\frac{1}{2}} P)^T A_d^{\frac{1}{2}} P = B^T B,$$

i, per tant, $|A| = |B^T B| = |B^T| |B| = |B|^2$ d'on $|B| = \pm |A|^{\frac{1}{2}}$.

Fem ara el canvi de variables $x \rightarrow y = B(x - m)$. És a dir, $x = m + B^{-1}y$. El seu determinant jacobí és $\frac{\partial(x)}{\partial(y)} = |B^{-1}| = |A|^{-\frac{1}{2}}$. Com $(x - m)^T A (x - m) = (B(x - m))^T B(x - m) = y^T y$, la funció de densitat de les noves variables resulta

$$f(y_1, \dots, y_n) = \frac{|A|^{\frac{1}{2}}}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2} y^T y} |A|^{-\frac{1}{2}} = \prod_{i=1}^n \frac{e^{-\frac{y_i^2}{2}}}{\sqrt{2\pi}}.$$

És a dir, les variables Y_i són normals, independents, totes amb esperança 0 i variància 1. ♣

A partir de l'anterior resultat és demostra que m és el vector d'esperances, i donada la matriu de covariàncies $C_{ij} = C[X_i, X_j]$, A és la seva inversa:

$$m_i = E[X_i], \quad A = C^{-1}. \quad (9.9)$$

DEM: Considerem les anteriors variables Y_i . Posem $E[Y_i] = 0$, $E[Y_i Y_j] = \delta_{ij}$ (la delta de Kronecker δ_{ij} val 1 si $i = j$ i val 0 en cas contrari). Veiem que impliquen aquestes relacions per a les variables originals X_i .

$$0 = E[Y] = E[B(X - m)] = B(E[X] - m) \Rightarrow E[X] - m = 0$$

ja que B és invertible. Tenim, per tant, $E[X_i] = m_i$ que ens dona el significat del vector m . Ara,

$$\delta_{ij} = E[Y_i Y_j] = E[(B(X - m))_i (B(X - m))_j] =$$

$$\sum_{k,l} B_{ik} B_{jl} E[(X - m)_k (X - m)_l] = \sum_{k,l} B_{ik} C[X_k, X_l] B_{lj}^T$$

que, en forma matricial és $I = BCB^T$ ($I = (\delta_{ij})$ és la matriu identitat) d'on $C = B^{-1}(B^T)^{-1} = (B^T B)^{-1} = A^{-1}$. ♣

Exemple 9.6 Recuperem els casos coneguts de variable gaussiana unidimensional i bidimensional.

Si $n = 1$ els vectors i les matrius es redueixen a un sol valor. $A = C[X, X]^{-1} = \sigma^{-2}$. Així, $|A|^{\frac{1}{2}} = \frac{1}{\sigma}$ i al substituir-ho a (9.8) s'obté (2.18).

Si $n = 2$ posem (X, Y) en lloc de (X_1, X_2) i ara

$$A = \begin{pmatrix} C[X, X] & C[X, Y] \\ C[X, Y] & C[Y, Y] \end{pmatrix}^{-1} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}^{-1} = \frac{1}{1 - \rho^2} \begin{pmatrix} \frac{1}{\sigma_1^2} & -\frac{\rho}{\sigma_1\sigma_2} \\ -\frac{\rho}{\sigma_1\sigma_2} & \frac{1}{\sigma_2^2} \end{pmatrix}.$$

$$|C| = \sigma_1^2 \sigma_2^2 (1 - \rho^2), \text{ d'on } |A|^{\frac{1}{2}} = \frac{1}{\sigma_1 \sigma_2 \sqrt{1 - \rho^2}}.$$

L'exponent a (9.8) és:

$$\begin{aligned} & -\frac{1}{2}(x - m_1, y - m_2) \frac{1}{1 - \rho^2} \begin{pmatrix} \frac{1}{\sigma_1^2} & -\frac{\rho}{\sigma_1\sigma_2} \\ -\frac{\rho}{\sigma_1\sigma_2} & \frac{1}{\sigma_2^2} \end{pmatrix} \begin{pmatrix} x - m_1 \\ y - m_2 \end{pmatrix} \\ & = -\frac{1}{2(1 - \rho^2)} \left\{ \frac{(x - m_1)^2}{\sigma_1^2} - 2\rho \frac{(x - m_1)(y - m_2)}{\sigma_1 \sigma_2} + \frac{(y - m_2)^2}{\sigma_2^2} \right\}. \end{aligned}$$

Així, obtenim la densitat (5.11). ♦

9.4.2 Processos estocàstics gaussians

El PE $X(t)$ s'anomena normal o gaussià si totes les seves mostres són gaussianes. És a dir, per a tot n , per a tot t_1, t_2, \dots, t_n les variables $X(t_1), X(t_2), \dots, X(t_n)$ són conjuntament normals. Llavors les densitats d'ordre n són:

$$f(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = \frac{|A|^{\frac{1}{2}}}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2}(x-m)^T A(x-m)} \quad (9.10)$$

on m i A ara depenen dels t_i . Concretament, $m_i = m(t_i)$ i A és la inversa de la matriu d'elements $C(t_i, t_j)$. Per tant, *un PE normal queda totalment determinat pel seu valor mitjà i la seva autocorrelació*.

9.5 Processos estacionaris

Es diu que $X(t)$ és un PE **estacionari en sentit estricte** si la seva distribució de probabilitat és invariant sota translacions temporals. Més concretament, les seves funcions de densitat verifiquen

$$f(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = f(x_1, x_2, \dots, x_n; t_1 + \tau, t_2 + \tau, \dots, t_n + \tau) \quad (9.11)$$

per a tot $n, t_1, t_2, \dots, t_n, \tau$. O, de manera equivalent, podem dir que els processos $X(t)$ i $X(t + \tau)$ tenen la mateixa distribució de probabilitat per a tot τ .

Això implica que les variables $X(t)$ estan idènticament distribuïdes per a tot t . Posant $n = 1$ i $\tau = -t$ en (9.11) trobem

$$f(x; t) = f(x; 0). \quad (9.12)$$

Amb $n = 2$ i $\tau = -t_1$ trobem

$$f(x_1, x_2; t_1, t_2) = f(x_1, x_2; 0, t_2 - t_1). \quad (9.13)$$

En general les funcions de densitat depenen només de les diferències entre els temps t_i .

Un PE es diu que és **estacionari en sentit ampli** si el seu valor mitjà és constant i la seva autocorrelació depèn només de la distància temporal considerada.

$$m(t) = m,$$

$$R(t_1, t_2) = R(t_1 - t_2). \quad (9.14)$$

Una conseqüència és que la potència és $R(0)$ constant.

Posant (9.12) i (9.13) en (9.4) i (9.5) veiem que tot PE estacionari en sentit estricte ho és en sentit ampli. L'invers no és cert en general però tenim el següent resultat:

Un PE *normal* estacionari en sentit ampli també ho és en sentit estricte.

DEM: Si el procés és estacionari en sentit ampli tenim que $C(t_i, t_j) = R(t_j - t_i) - m^2$ depèn només de les diferències de temps. Com tota l'estadística del procés normal queda determinada en (9.10) per m i $C(t_i, t_j)$ la distribució és invariant sota el canvi (9.11). ♣

Una característica de molts PE (estacionaris) és que

$$\lim_{|\tau| \rightarrow \infty} C(\tau) = 0. \quad (9.15)$$

De fet, hem vist a l'exemple 9.5 que (9.15) pot no complir-se. Però en l'exemple 9.1 veiem que la condició resulta plausible ja que, al ser el moviment continu, per petits increments de temps les posicions estan molt correlacionades. Per increments temporals grans la partícula "oblida" la posició inicial i tenim (9.15).

Existeix també una versió estadística de la periodicitat. Es diu que $X(t)$ és un PE **cicloestacionari en sentit estricte** si la seva distribució de probabilitat és periòdica sota translacions temporals. És a dir, existeix $T > 0$ fixat tal que per a tot $k \in \mathbb{Z}$

$$f(x_1, x_2, \dots, x_n; t_1 + kT, t_2 + kT, \dots, t_n + kT) = f(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) \quad (9.16)$$

Un PE es diu que és **cicloestacionari en sentit ampli** si per a tot $k \in \mathbb{Z}$

$$\begin{aligned} m(t + kT) &= m(t), \\ R(t_1 + kT, t_2 + kT) &= R(t_1, t_2). \end{aligned} \quad (9.17)$$

Notem que un PE on les variables $X(t_i)$ siguin sempre independents ens dóna una funció absolutament aleatòria que difícilment modelarà un fenomen real. Per una altra part, tampoc les funcions tan regulars com les de l'exemple 9.5 són les que ens interessin. Un PE sol ser massa complicat perquè sigui possible manipular directament les funcions $X(t)$. El que caracteritza realment el procés són les funcions com $m(t)$ i $C(t_1, t_2)$.

Capítol 10

Exemples de processos estocàstics

Donat un PE, si $X(t)$ és una v.a. discreta diem que el procés és d'**estat discret**. Anàlogament, si $X(t)$ és una v.a. contínua diem que el procés és d'**estat continu**. L'exemple més important de PE d'estat discret és el procés de Poisson que estudiem a continuació.

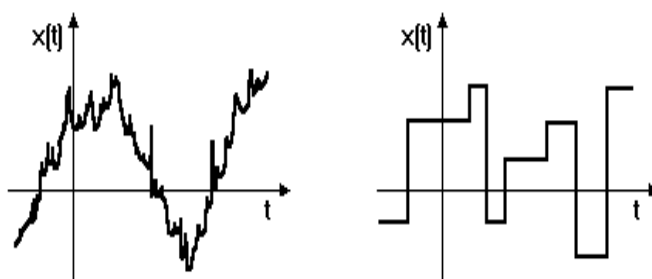


Figura 16: Processos d'estat continu i discret.

10.1 El procés de Poisson

Considerem que a mesura que passa el temps es van produint una sèrie d'esdeveniments idèntics de manera aleatòria. Per exemple, rebre trucades en una centralita telefònica, arribar vehicles a una gasolinera o desintegrar-se partícules radioactives. Suposem que els esdeveniments es produeixen de manera independent. Prendrem sempre temps positius començant a contar en $t = 0$. Indiquem per $N(t)$ el nombre d'esdeveniments que s'han produït en l'interval $[0, t)$. El nombre d'esdeveniments en un interval $[t_a, t_b)$ és $N(t_a, t_b) = N(t_b) - N(t_a)$.

Diem que el procés és de Poisson si

- El nombre d'esdeveniments en dos intervals temporals disjunts són variables aleatòries independents.
- Si $N = N(t, t + \tau)$, llavors $P(N = 1) = \lambda\tau + o(\tau)$ i $P(N > 1) = o(\tau)$.

Tal com vam justificar al parlar de la variable de Poisson (equació (2.9)) això implica que la v.a. $N = N(t_a, t_b)$ té la funció de probabilitat

$$P(N = n) = e^{-\alpha} \frac{\alpha^n}{n!}, \quad (10.1)$$

on $\alpha = \lambda(t_b - t_a)$ i $n = 0, 1, \dots$

El PE de Poisson consisteix en prendre com a funció el nombre total d'esdeveniments produïts en l'interval $[0, t]$:

$$X(t) = N(t).$$

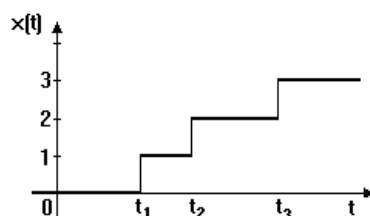


Figura 17: Procés de Poisson.

Fixat t , $X(t)$ és una v.a. de Poisson amb funció de probabilitat

$$P(X(t) = n) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}. \quad (10.2)$$

Es diu que el procés de Poisson és **no homogeni** quan λ depèn de t . El seu tractament és més complicat que el que hem vist ara. Només considerarem el cas **homogeni** on λ és constant.

Recordem que la variable (10.2) verifica $E[X(t)] = V[X(t)] = \lambda t$. Llavors, pel procés de Poisson

$$m(t) = \lambda t, \quad (10.3)$$

$$R(t_1, t_2) = \begin{cases} \lambda t_1 + \lambda^2 t_1 t_2 & \text{si } t_1 \leq t_2 \\ \lambda t_2 + \lambda^2 t_1 t_2 & \text{si } t_1 \geq t_2 \end{cases} \quad (10.4)$$

DEM: El valor mitjà és $m(t) = E[X(t)] = \lambda t$.

Ara considerem $t_1 \leq t_2$. Les variables aleatòries $X(t_1)$ i $X(t_2) - X(t_1)$ són variables de Poisson $N(0, t_1)$ i $N(t_1, t_2)$ independents ja que els intervals són disjunts. Llavors l'autocorrelació és

$$\begin{aligned} R(t_1, t_2) &= E[X(t_1)X(t_2)] = E[X(t_1)(X(t_2) - X(t_1))] + E[X(t_1)^2] \\ &= E[X(t_1)]E[(X(t_2) - X(t_1))] + E[X(t_1)^2] \\ &= E[X(t_1)]E[X(t_2)] + E[X(t_1)^2] - E[X(t_1)]^2 = E[X(t_1)]E[X(t_2)] + V[X(t_1)] \\ &= \lambda t_1 \lambda t_2 + \lambda t_1. \end{aligned}$$

El cas $t_1 \geq t_2$ s'obté intercanviant t_1 i t_2 ja que R és simètrica. ♣

Podem posar $R(t_1, t_2) = \lambda \min(t_1, t_2) + \lambda^2 t_1 t_2$. La funció d'autocovariància és $C(t_1, t_2) = \lambda \min(t_1, t_2)$.

Segons (10.3) el paràmetre λ és el nombre mig d'esdeveniments per unitat de temps.

10.2 Senyal telegràfic aleatori

Considerem la variable de Poisson $N(t)$. El senyal telegràfic **semialeatori** és el procés $X(t) = (-1)^{N(t)}$. És a dir:

$$X(t) = \begin{cases} 1 & \text{si } N(t) \text{ és parell} \\ -1 & \text{si } N(t) \text{ és senar} \end{cases} \quad (10.5)$$

Així, $X(0) = 1$ i canvia de signe cada vegada que es produeix un esdeveniment. La seva funció de probabilitat és

$$P(X(t)=1) = P(N(t)=0) + P(N(t)=2) + \dots = e^{-\lambda t} \left(1 + \frac{(\lambda t)^2}{2!} + \dots \right) = e^{-\lambda t} \cosh \lambda t.$$

$$P(X(t)=-1) = P(N(t)=1) + P(N(t)=3) + \dots = e^{-\lambda t} \left(\lambda t + \frac{(\lambda t)^3}{3!} + \dots \right) = e^{-\lambda t} \sinh \lambda t.$$

El valor mitjà és

$$m(t) = E[X(t)] = 1 \cdot e^{-\lambda t} \cosh \lambda t + (-1) \cdot e^{-\lambda t} \sinh \lambda t = e^{-2\lambda t}.$$

L'autocorrelació és

$$\begin{aligned} R(t_1, t_2) &= E[X(t_1)X(t_2)] = E[(-1)^{N(t_1)}(-1)^{N(t_2)}] = E[(-1)^{N(t_2)-N(t_1)}] = \\ &= E[(-1)^{N(t_1, t_2)}] = e^{-2\lambda|t_2-t_1|}. \end{aligned}$$

ja que $(-1)^N = (-1)^{-N}$ i $N(t_1, t_2)$ és una v.a. de Poisson de paràmetre $\lambda|t_2 - t_1|$.

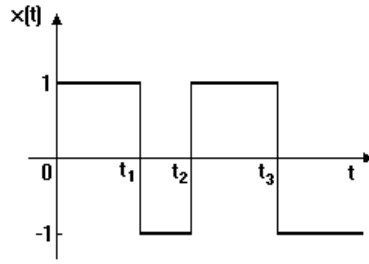


Figura 18: Senyal telegràfic semialeatori.

Aquest PE s'anomena semialeatori perquè hem fixat $X(0)$. El PE de **senyal telegràfic aleatori** és

$$Y(t) = SX(t) \quad (10.6)$$

on S és una v.a. de Bernoulli amb valors 1 i -1 equiprobables, independent de $X(t)$. Ara

$$m(t) = 0, \quad R(t_1, t_2) = e^{-2\lambda|t_2-t_1|}. \quad (10.7)$$

DEM: $E[Y(t)] = E[S]E[X(t)] = 0$, $E[Y(t_1)Y(t_2)] = E[S^2 X(t_1)X(t_2)] = E[X(t_1)X(t_2)] = e^{-2\lambda|t_2-t_1|}$. ♣

10.3 Tren d'impulsos de Poisson

Sigui $X(t)$ un PE de Poisson. El tren d'impulsos associat a ell és el procés

$$Z(t) = \frac{d}{dt}X(t). \quad (10.8)$$

Si els esdeveniments es produeixen en els instants t_1, t_2, \dots , podem representar X i Z com

$$X(t) = \sum_i u(t - t_i), \quad (10.9)$$

$$Z(t) = \sum_i \delta(t - t_i), \quad (10.10)$$

on u i δ són les funcions de Heaviside i Dirac.

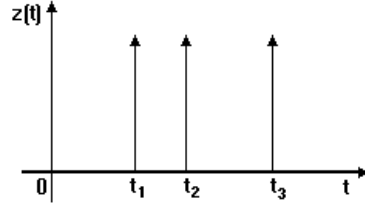


Figura 19: Impulsos de Poisson.

(10.10) representa un tipus de soroll de naturalesa impulsiva (**shot noise**). Els seus valor mitjà, autocorrelació i autocovariància són:

$$m(t) = \lambda, \quad (10.11)$$

$$R(t_1, t_2) = \lambda^2 + \lambda\delta(t_2 - t_1), \quad C(t_1, t_2) = \lambda\delta(t_2 - t_1). \quad (10.12)$$

DEM:

$$m(t) = E[Z(t)] = E\left[\frac{d}{dt}X(t)\right] = \frac{d}{dt}E[X(t)] = \frac{d}{dt}(\lambda t) = \lambda.$$

$$\begin{aligned} R(t_1, t_2) &= E[Z(t_1)Z(t_2)] = E\left[\frac{d}{dt_1}X(t_1)\frac{d}{dt_2}X(t_2)\right] = \frac{d^2}{dt_1dt_2}E[X(t_1)X(t_2)] \\ &= \frac{d^2}{dt_1dt_2}(\lambda^2 t_1 t_2 + \lambda \min(t_1, t_2)) = \lambda^2 + \lambda\delta(t_1 - t_2), \end{aligned}$$

ja que

$$\begin{aligned} \min(t_1, t_2) &= t_1 u(t_2 - t_1) + t_2 u(t_1 - t_2), \\ \frac{d}{dt_2} \min(t_1, t_2) &= t_1 \delta(t_2 - t_1) + u(t_1 - t_2) - t_2 \delta(t_1 - t_2) = u(t_1 - t_2) - (t_2 - t_1) \delta(t_2 - t_1) = u(t_1 - t_2), \\ \frac{d^2}{dt_1 dt_2} \min(t_1, t_2) &= \delta(t_1 - t_2). \clubsuit \end{aligned}$$

Notem que els processos de senyal telegràfic aleatori i el tren d'impulsos de Poisson són estacionaris en sentit ampli (veure (10.7), (10.11), (10.12)) mentre que el propi procés de Poisson òbviament no és estacionari. El fet és que l'arribada de fenòmens tipus Poisson és estrictament estacionària (l'origen de temps és arbitrari) i això és el que es manifesta en alguns processos derivats d'aquesta situació.

10.4 Distribució dels temps en el procés de Poisson

En un procés de Poisson $X(t)$ els esdeveniments es produeixen en els instants t_1, t_2, t_3, \dots . Considerem les v.a.'s T_1, T_2, T_3, \dots associades a aquests instants. La densitat de cadascuna d'elles és, per a $t \geq 0$:

$$f_{T_n}(t) = \lambda \frac{(\lambda t)^{n-1}}{(n-1)!} e^{-\lambda t}. \quad (10.13)$$

DEM:

$$F_{T_n}(t) = P(T_n \leq t) = P(X(t) \geq n) = 1 - P(X(t) < n) = 1 - e^{-\lambda t} \left(1 + \lambda t + \dots + \frac{(\lambda t)^{n-1}}{(n-1)!}\right),$$

$$f_{T_n}(t) = -\frac{d}{dt} \left(e^{-\lambda t} \left(1 + \lambda t + \dots + \frac{(\lambda t)^{n-1}}{(n-1)!}\right) \right)$$

(posem $\alpha = \lambda t$)

$$\begin{aligned}
&= -\lambda \frac{d}{d\alpha} \left(e^{-\alpha} \left(1 + \alpha + \dots + \frac{\alpha^{n-1}}{(n-1)!} \right) \right) \\
&= -\lambda \left[-e^{-\alpha} \left(1 + \alpha + \dots + \frac{\alpha^{n-1}}{(n-1)!} \right) + e^{-\alpha} \left(1 + \alpha + \dots + \frac{\alpha^{n-2}}{(n-2)!} \right) \right] \\
&= \lambda e^{-\alpha} \frac{\alpha^{n-1}}{(n-1)!} \clubsuit
\end{aligned}$$

En particular, notem que $f_{T_1}(t) = \lambda e^{-\lambda t}$. És a dir, T_1 és exponencial de paràmetre λ .

La distribució conjunta de (T_1, T_2, \dots, T_n) ve donada per la densitat:

$$f(t_1, t_2, \dots, t_n) = \lambda^n e^{-\lambda t_n}, \quad (10.14)$$

per a $0 \leq t_1 < t_2 < \dots < t_n < \infty$.

DEM:

Les funcions que considerarem seran 0 quan no es verifiqui $0 \leq t_1 < t_2 < \dots < t_n < \infty$. Calculem primer la distribució condicionada

$$\begin{aligned}
F(t_k \mid t_1, t_2, \dots, t_{k-1}) &= P(T_k \leq t_k \mid T_1 = t_1, T_2 = t_2, \dots, T_k = t_{k-1}) \\
&= P(n(t_{k-1}, t_k) \geq 1) = 1 - e^{-\lambda(t_k - t_{k-1})}.
\end{aligned}$$

Derivant respecte a t_k trobem

$$f(t_k \mid t_1, t_2, \dots, t_{k-1}) = \lambda e^{-\lambda(t_k - t_{k-1})},$$

per a $t_k \geq t_{k-1}$.

Ara factoritzem la densitat conjunta a partir de les marginals i les condicionades:

$$\begin{aligned}
f(t_1, t_2, \dots, t_n) &= f(t_n \mid t_1, t_2, \dots, t_{n-1}) f(t_1, t_2, \dots, t_{n-1}) = \dots \\
&= f(t_n \mid t_1, t_2, \dots, t_{n-1}) f(t_{n-1} \mid t_1, t_2, \dots, t_{n-2}) \dots f(t_3 \mid t_1, t_2) f(t_2 \mid t_1) f(t_1) \\
&= \lambda e^{-\lambda(t_n - t_{n-1})} \lambda e^{-\lambda(t_{n-1} - t_{n-2})} \dots \lambda e^{-\lambda(t_2 - t_1)} \lambda e^{-\lambda t_1} = \lambda^n e^{-\lambda t_n} \clubsuit
\end{aligned}$$

Considerem ara les variables que ens donen el temps entre esdeveniments consecutius:

$$\Delta_1 = T_1, \quad \Delta_2 = T_2 - T_1, \dots, \quad \Delta_n = T_n - T_{n-1}, \quad (10.15)$$

La densitat conjunta de la variable n -dimensional $(\Delta_1, \Delta_2, \dots, \Delta_n)$ és:

$$f(\delta_1, \delta_2, \dots, \delta_n) = \lambda^n e^{-\lambda(\delta_1 + \delta_2 + \dots + \delta_n)}, \quad (10.16)$$

per a $\delta_i \geq 0, i = 1, \dots, n$.

DEM: La regió $0 \leq t_1 < t_2 < \dots < t_n < \infty$ correspon a $\delta_1, \delta_2, \dots, \delta_n \geq 0$. La transformació (10.15) és bijectiva entre aquestes regions i el seu determinant jacobí val 1. Així la densitat s'obté substituint (10.15) en (10.14) que ens dona (10.16). ♣

En (10.16) veiem que els temps entre esdeveniments successius $\Delta_i = T_i - T_{i-1}$ són variables exponencials independents. Així $T_n = \Delta_1 + \Delta_2 + \dots + \Delta_n$ és la suma de n v.a.'s exponencials independents.

10.5 Senyals binaris

Un senyal binari és un PE d'estat discret que només pren valors 0 i 1.

Considerem un senyal binari a temps discret B_n . tal que, per a tot n , B_n pren valors 0 i 1 amb igual probabilitat. Si fixem T hi ha un procés a temps continu equivalent definit:

$$B(t) = B_n \quad \text{si } nT \leq t < (n+1)T \quad (10.17)$$

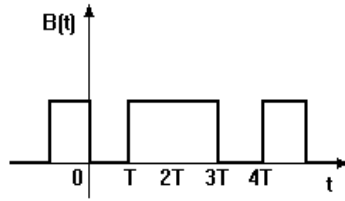


Figura 20: Procés binari.

Un procés d'aquest tipus pot representar la informació que conté un arxiu informàtic. Per transportar aquesta informació mitjançant un medi electromagnètic (per exemple, una línia telefònica) resulta més útil transformar el procés (10.17) en

$$X(t) = \cos(\omega t + \theta(t)), \quad (10.18)$$

on $\omega = 2\pi\nu_P$ és la freqüència portadora i

$$\theta(t) = \begin{cases} \frac{\pi}{2} & \text{si } B(t) = 1 \\ -\frac{\pi}{2} & \text{si } B(t) = 0 \end{cases} \quad (10.19)$$

La transformació (10.18) s'anomena modulació de freqüència o PSK (*phase shift keying*) i es pot aplicar a diferents tipus de senyal.

10.6 Oscil·lacions aleatòries

Es tracta de funcions oscil·latòries que contenen paràmetres aleatoris. L'exemple 9.3 de l'anterior capítol n'era un cas. Ara estudiem quan és estacionari el procés

$$X(t) = A \cos \omega t + B \sin \omega t \quad (10.20)$$

on ω està fixada i (A, B) és una v.a. bidimensional. Tenim que (10.20) és estacionari

- en sentit ampli si i només si $m_A = m_B = 0$, $\sigma_A = \sigma_B$ i $\rho_{AB} = 0$.
- en sentit estricte si i només si la distribució de (A, B) té simetria circular.

DEM:

Si (10.20) és estacionari en sentit ampli

$$m(t) = m_A \cos \omega t + m_B \sin \omega t = m.$$

Com les funcions \cos , \sin i 1 són linealment independents, ha de ser $m_A = m_B = m = 0$.

La funció $R(t, t + \tau)$ no ha de dependre de t . Posant $\tau = 0$ tenim que $R(t, t)$ és constant i, per tant, $R(0, 0) = E[X(0)^2] = E[A^2] = \sigma_A^2$ coincideix amb $R(\frac{\pi}{2\omega}, \frac{\pi}{2\omega}) = E[X(\frac{\pi}{2\omega})^2] = E[B^2] = \sigma_B^2$. Diem σ al valor comú de σ_A i σ_B . Tenim també $E[AB] = \rho\sigma^2$. Llavors

$$\begin{aligned} R(t, t + \tau) &= E[X(t)X(t + \tau)] = E[(A \cos \omega t + B \sin \omega t)(A \cos \omega(t + \tau) + B \sin \omega(t + \tau))] \\ &= \sigma^2 \cos \omega \tau + \rho \sigma^2 \sin \omega(2t + \tau). \end{aligned}$$

Perquè aquesta funció no depengui de t ha de ser $\rho = 0$.

Per veure quan és estacionari en sentit estricte utilitzem coordenades polars $A = R \cos \Theta$, $B = R \sin \Theta$. Llavors $X(t) = R \cos(\Theta - \omega t)$.

Si el procés és estacionari en sentit estricte la v.a. bidimensional $(X(t), X(t + \frac{\pi}{2\omega}))$ té la mateixa distribució per a tot t . En $t = 0$ i $t = \frac{\pi}{\omega}$ tenim les variables $(R \cos \Theta, R \sin \Theta)$ i $(R \cos(\Theta - \alpha), R \sin(\Theta - \alpha))$ respectivament. En

polars corresponen a (R, Θ) i $(R, \Theta - \alpha)$. Així, la funció de densitat no pot dependre de θ , és a dir, tenim simetria circular.

Inversament, si tenim simetria circular, el canvi $\Theta \rightarrow \Theta + \omega t_1$ deixa totes les distribucions invariants. Per tant la distribució conjunta de la v.a. $(X(t_1), X(t_2), \dots, X(t_n))$ és idèntica a la de $(X(0), X(t_2 - t_1), \dots, X(t_n - t_1))$ que és invariant sota translacions temporals. ♣

Capítol 11

Processos estacionaris

11.1 Realitzacions d'un PE

Considerem el PE $X(t)$. Si com a resultat de l'experiment obtenim la funció $X_1(t)$ diem que $X_1(t)$ és una **realització** del procés. En general, no sabem a priori quines propietats té $X_1(t)$. Si volem aplicar operacions al procés $X(t)$ com la derivació o la integració, pot ser massa restrictiu voler-ho fer sobre $X_1(t)$. Així, relaxem les definicions exigint només que la propietat es verifiqui en mitjana. Això requereix un nou concepte de límit.

11.2 Límit en mitjana quadràtica

Si X_s és una variable aleatòria dependent d'un paràmetre continu s diem que $\lim_{s \rightarrow a} X_s = l$ en *mitjana quadràtica* (MQ) si

$$\lim_{s \rightarrow a} E[(X_s - l)^2] = 0. \quad (11.1)$$

Si només ens interessa l'existència del límit i no el seu valor podem utilitzar el criteri de Cauchy que diu que X_s convergeix en MQ si

$$\lim_{s_1, s_2 \rightarrow a} E[(X_{s_1} - X_{s_2})^2] = 0. \quad (11.2)$$

Notem que el resultat del límit l és en general una v.a. Llavors es verifica que l'esperança del límit és el límit de l'esperança

$$E[l] = \lim_{s \rightarrow a} E[X_s]. \quad (11.3)$$

DEM: $(E[l] - E[X_s])^2 = E[l - X_s]^2 \leq E[(l - X_s)^2] \rightarrow 0$. (S'ha utilitzat que per qualsevol variable Y , $E[Y^2] \geq E[Y]^2$ degut a que la variància és no negativa.) ♣

11.3 Continuïtat estocàstica

Diem que un PE $X(t)$ és continu en MQ si

$$\lim_{\epsilon \rightarrow 0} E[(X(t + \epsilon) - X(t))^2] = 0. \quad (11.4)$$

Això vol dir que fixat t quasi totes les realitzacions del procés són contínues en t . Però donada una realització aquesta pot tenir diferents punts de discontinuïtat. La propietat (11.4) equival a

$$\lim_{\epsilon \rightarrow 0} \{R(t + \epsilon, t + \epsilon) - 2R(t + \epsilon, t) + R(t, t)\} = 0, \quad (11.5)$$

que es verifica, en particular, si la funció d'autocorrelació $R(t_1, t_2)$ és contínua. Això passa, per exemple, en el procés de Poisson on $R(t_1, t_2) = \lambda \min(t_1, t_2) + \lambda^2 t_1 t_2$. En canvi, tota realització de un PE de Poisson té un conjunt numerable de discontinuïtats.

Existeixen expressions equivalents a (11.4) i (11.5) per a la derivada d'un PE.

11.4 Integrals estocàstiques

Concentrem-nos ara en la integració. La integral (de Riemann) d'una funció sobre un interval $[a, b]$ consisteix en el límit de les seves sumes de Riemann quan les particions de l'interval són cada vegada més fines. Direm que la integral d'un PE sobre l'interval $[a, b]$ val I si

$$\lim_{\Delta t_i \rightarrow 0} E \left[\left(\sum_i X(t_i) \Delta t_i - I \right)^2 \right] = 0, \quad (11.6)$$

on els t_i formen una partició de $[a, b]$. Si (11.6) es verifica escriurem $I = \int_a^b X(t) dt$.

Si existeix la integral

$$\int_a^b \int_a^b R(t_1, t_2) dt_1 dt_2, \quad (11.7)$$

llavors el procés és integrable en $[a, b]$.

DEM:

Suposem que (11.7) existeix i utilitzem el criteri de Cauchy. Considerem dues particions, la primera indexada amb i o i' i la segona amb j o j'

$$\begin{aligned} & \lim_{\Delta t_i, \Delta t_j \rightarrow 0} E \left[\left(\sum_i X(t_i) \Delta t_i - \sum_j X(t_j) \Delta t_j \right)^2 \right] \\ &= \lim \left\{ \sum_{i, i'} R(t_i, t_{i'}) \Delta t_i \Delta t_{i'} - 2 \sum_{i, j} R(t_i, t_j) \Delta t_i \Delta t_j + \sum_{j, j'} R(t_j, t_{j'}) \Delta t_j \Delta t_{j'} \right\} = 0, \end{aligned}$$

ja que cada sumatori tendeix a (11.7). ♣

$\int_a^b X(t) dt$ és una variable aleatòria. De la linealitat de l'esperança i de (11.3) es pot concloure la propietat

$$E \left[\int_a^b X(t) dt \right] = \int_a^b E[X(t)] dt, \quad (11.8)$$

que serà àmpliament utilitzada en el que segueix.

Deixem de banda la consideració dels aspectes analítics fins de les anteriors definicions i ressaltem tan sols que les propietats d'un PE solen determinar-se a partir del comportament de les funcions de valor mitjà i d'autocorrelació.

Exemple 11.1 El següent cas presenta una aparent paradoxa. Sigui el procés $X(t) = te^{-Yt}$ on Y és una variable aleatòria uniforme en $[0, 1]$. Resulta que totes les seves realitzacions verifiquen

$$\lim_{t \rightarrow \infty} X_i(t) = 0$$

mentre que el seu valor mitjà $m(t) = E[X(t)] = \int_0^1 te^{-yt} \cdot 1 dy = 1 - e^{-t}$ tendeix a 1 en l'infinit.

La discrepància es deu a que el procés no tendeix a zero en MQ. En efecte,

$$\lim_{t \rightarrow \infty} E[(X(t) - 0)^2] = \lim_{t \rightarrow \infty} E[t^2 e^{-2Yt}] = \lim_{t \rightarrow \infty} \frac{t}{2} (1 - e^{-2t}) = \infty.$$

Així, si fem la mitjana de moltes realitzacions del procés, s'obtenen valors més alts per t 's més elevats. El missatge és que no s'han d'eleva les propietats de les realitzacions a propietats estadístiques sense verificar-ho. ♦

11.5 Ergodicitat

Sigui $X(t)$ un PE estacionari amb valor mitjà $m(t) = m$ i funció d'autocorrelació $R(t, t+\tau) = R(\tau)$. Per calcular m podríem fer moltes realitzacions del procés i avaluar

$$m = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n X_i(t). \quad (11.9)$$

En el que segueix estarem interessats en avaluar els paràmetres del procés no a través de mitjanes sobre moltes realitzacions sinó fent mitjanes temporals d'una sola realització. La **mitjana temporal** del procés és

$$m_T = \langle X(t) \rangle_T = \frac{1}{2T} \int_{-T}^T X(t) dt. \quad (11.10)$$

m_T és una v.a. ordinària. La seva esperança val m .

DEM:

$$E[m_T] = E\left[\frac{1}{2T} \int_{-T}^T X(t) dt\right] = \frac{1}{2T} \int_{-T}^T E[X(t)] dt = \frac{1}{2T} \int_{-T}^T m dt = m. \clubsuit$$

Un procés ergòdic és aquell en que les mitjanes temporals igualen les mitjanes sobre realitzacions. Més concretament definim:

- **Ergodicitat en valor mitjà.**

$X(t)$ és ergòdic en valor mitjà si m_T tendeix a m en MQ. És a dir,

$$\lim_{T \rightarrow \infty} E[(m_T - m)^2] = 0. \quad (11.11)$$

Aquesta condició és que $\sigma_T^2 \rightarrow 0$ quan $T \rightarrow \infty$, on σ_T és la desviació de la variable m_T .

- **Ergodicitat en autocorrelació.**

$X(t)$ és ergòdic en autocorrelació si el PE $Z_\lambda(t) = X(t + \lambda)X(t)$ és ergòdic en valor mitjà. En aquest cas les mitjanes temporals de $Z_\lambda(t)$ ens donen l'autocorrelació de $X(t)$.

Notem que

$$\sigma_T^2 = \frac{1}{4T^2} \int_{-T}^T \int_{-T}^T C(t_1, t_2) dt_1 dt_2. \quad (11.12)$$

DEM:

$$\begin{aligned} \sigma_T^2 &= E[(m_T - m)^2] = E\left[\left(\frac{1}{2T} \int_{-T}^T (X(t) - m) dt\right)^2\right] \\ &= \frac{1}{(2T)^2} \int_{-T}^T \int_{-T}^T E[(X(t_1) - m)(X(t_2) - m)] dt_1 dt_2, \end{aligned}$$

i la funció integrada és $C(t_1, t_2)$. \clubsuit

Exemple 11.2 Pel tren d'impulsos de Poisson $C(t_1, t_2) = \lambda \delta(t_1 - t_2)$.

$$\sigma_T^2 = \frac{1}{4T^2} \int_{-T}^T \int_{-T}^T \lambda \delta(t_1 - t_2) dt_1 dt_2 = \frac{1}{4T^2} \int_{-T}^T \lambda dt_1 = \frac{\lambda}{2T}.$$

Com $\frac{\lambda}{2T} \rightarrow 0$ quan $T \rightarrow \infty$, el procés és ergòdic en valor mitjà. \blacklozenge

Degut a que considerem processos estacionaris, l'expressió (11.12) es pot posar com

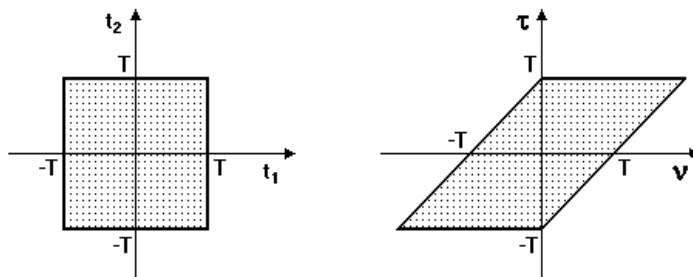
$$\sigma_T^2 = \frac{1}{2T} \int_{-2T}^{2T} C(\tau) \left(1 - \frac{|\tau|}{2T}\right) d\tau. \quad (11.13)$$

DEM:

Fem el canvi de variable

$$\begin{cases} \tau = t_1 - t_2 \\ \nu = t_1 \end{cases}$$

que té jacobiana 1 i transforma la regió d'integració segons el dibuix:



$$\begin{aligned} \int_{-T}^T \int_{-T}^T C(t_1, t_2) dt_1 dt_2 &= \int_{-2T}^0 \left(\int_{-T}^{T+\tau} C(\tau) d\nu \right) d\tau + \int_0^{2T} \left(\int_{-T+\tau}^T C(\tau) d\nu \right) d\tau \\ &= \int_{-2T}^0 C(\tau)(2T + \tau) d\tau + \int_0^{2T} C(\tau)(2T - \tau) d\tau = \int_{-2T}^{2T} C(\tau)(2T - |\tau|) d\tau. \clubsuit \end{aligned}$$

11.6 Condicions suficients d'ergodicitat

- Si $\int_{-\infty}^{\infty} C(\tau) d\tau < \infty$ el procés és ergòdic en valor mitjà.

DEM:

$$\frac{1}{2T} \int_{-2T}^{2T} C(\tau) \left(1 - \frac{|\tau|}{2T}\right) d\tau \leq \frac{1}{2T} \int_{-2T}^{2T} C(\tau) d\tau \rightarrow 0$$

quan $T \rightarrow \infty$. ♣

Recordem la propietat $|C[X, Y]| \leq \sigma_X \sigma_Y$. En el cas d'un procés estacionari, fent $X = X(t)$ i $Y = X(t + \tau)$ trobem que

$$|C(\tau)| \leq C(0).$$

- Si $C(0) < \infty$ i $C(\tau) \rightarrow 0$ quan $|\tau| \rightarrow \infty$, el procés és ergòdic en valor mitjà.

DEM:

Per a tot $\epsilon > 0$ existeix a tal que $|C(\tau)| < \epsilon$ si $|\tau| > a$. Llavors

$$\begin{aligned} 2T\sigma_T^2 &= \int_{-2T}^{2T} C(\tau) \left(1 - \frac{|\tau|}{2T}\right) d\tau \leq \int_{-2T}^{2T} |C(\tau)| d\tau \\ &\leq \int_{-a}^a |C(\tau)| d\tau + \int_{a < |\tau| < 2T} |C(\tau)| d\tau < 2aC(0) + 4T\epsilon. \end{aligned}$$

Així

$$\sigma_T^2 < \frac{a}{T} C(0) + 2\epsilon \rightarrow 2\epsilon$$

quan $T \rightarrow \infty$. Com això val per a tot ϵ , ha de ser $\sigma_T^2 \rightarrow 0$ quan $T \rightarrow \infty$. ♣

Exemple 11.3 Estudiar l'ergodicitat del senyal telegràfic aleatori i del shot noise amb els anteriors criteris.

Pel senyal telegràfic $C(\tau) = e^{-2\lambda|\tau|}$ que té integral finita ($= \frac{1}{\lambda}$). Segons el primer criteri, és ergòdic en valor mitjà. El mateix es conclou amb el segon criteri ja que $C(0) = 1$, finit, i $C(t)$ s'anul·la en l'infinit.

Pel shot noise $C(\tau) = \lambda\delta(\tau)$ que té integral finita ($= \lambda$). Segons el primer criteri, és ergòdic en valor mitjà. El segon criteri no es pot aplicar ja que $C(0) = \infty$. ♦

11.7 Espectre de potència d'un procés estacionari

La densitat espectral o espectre de potència d'un PE estacionari és la transformada de Fourier de la funció d'autocorrelació.

$$S(\omega) = \int_{-\infty}^{\infty} R(\tau) e^{-j\omega\tau} d\tau. \quad (11.14)$$

Invertint (11.14) trobem

$$R(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega) e^{j\omega\tau} d\omega. \quad (11.15)$$

Com, per processos reals, $R(-\tau) = R(\tau)$, si fem el canvi de variable $\tau' = -\tau$ en (11.14) trobem que $S(\omega)$ és real.

DEM: $S(\omega) = \int_{-\infty}^{\infty} R(\tau') e^{j\omega\tau'} d\tau' = S^*(\omega)$. ♣

Com $S(-\omega) = S^*(\omega) = S(\omega)$, $S(\omega)$ és una funció parell. Recordem que la potència mitjana era

$$R(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega) d\omega, \quad (11.16)$$

$S(\omega)$ és, en efecte, la densitat de potència distribuïda segons el contingut en freqüències.

Exemple 11.4 Shot noise.

Notem que al ser la transformada de Fourier de $\delta(\tau)$ igual a 1 la inversió ens dona la representació integral de la δ de Dirac:

$$\delta(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{j\omega\tau} d\omega. \quad (11.17)$$

En el cas dels impulsos de Poisson era $R(\tau) = \lambda^2 + \lambda\delta(\tau)$ d'on

$$S(\omega) = \int_{-\infty}^{\infty} (\lambda^2 + \lambda\delta(\tau)) e^{-j\omega\tau} d\tau = 2\pi\lambda^2\delta(\omega) + \lambda. \blacklozenge$$

Exemple 11.5 Senyal telegràfic. És $R(\tau) = \exp\{-2\lambda|\tau|\}$ i

$$S(\omega) = \int_{-\infty}^{\infty} e^{-2\lambda|\tau|} e^{-j\omega\tau} d\tau = 2 \int_0^{\infty} e^{-2\lambda\tau} \cos(\omega\tau) d\tau = \frac{4\lambda}{\omega^2 + 4\lambda^2}. \blacklozenge$$

11.8 Sistemes lineals

Considerem un PE $X(t)$. Igual que una variable aleatòria X dona lloc a noves variables fent transformacions del tipus $Y = g(X)$, podem transformar el procés $X(t)$ per obtenir el nou PE $Y(t)$. Diem L a aquesta transformació que anomenarem sistema i escrivim $Y(t) = L[X(t)]$. Gràficament ho pensem com un objecte que rep una entrada $X(t)$ i torna una sortida $Y(t)$.



Figura 21: Sistema lineal.

Un sistema és **determinista** si fixada la realització de $X(t)$ que entra, la sortida queda ja determinada. Si, pel contrari, una mateixa entrada pot donar lloc a sortides diferents diem que el sistema és **estocàstic**. En el que segueix només considerem sistemes deterministes.

Un sistema és **sense memòria** si el valor de $Y(t)$ per cada t depèn exclusivament del valor de $X(t)$ pel mateix t .

Un sistema és **lineal** si

$$L[\lambda_1 X_1(t) + \lambda_2 X_2(t)] = \lambda_1 L[X_1(t)] + \lambda_2 L[X_2(t)], \quad (11.18)$$

on λ_1, λ_2 poden ser variables aleatòries però no dependre de t .

Un tipus particular de transformació és la translació temporal definida

$$T_a[X(t)] = X(t - a). \quad (11.19)$$

Un sistema és **invariant en el temps** si $LT_a = T_aL$ per a tot a . Això es pot expressar de manera equivalent dient que si $X(t)$ dona la sortida $Y(t)$ llavors $X(t - a)$ dona la sortida $Y(t - a)$. Si pensem L com un dispositiu físic que transforma un senyal, la invariància en el temps vol dir que la constitució d'aquest dispositiu és idèntica en tot t . A partir d'ara considerem només sistemes lineals invariants en el temps.

Definim la **funció de resposta impulsiva** h d'un sistema lineal com

$$h(t) = L[\delta(t)]. \quad (11.20)$$

Coneixent h tenim que una resposta arbitrària es pot determinar fent una convolució:

$$L[X(t)] = (h * X)(t). \quad (11.21)$$

DEM:

$$\begin{aligned} L[X(t)] &= L \left[\int_{-\infty}^{\infty} \delta(t - \tau) X(\tau) d\tau \right] = \int_{-\infty}^{\infty} L[\delta(t - \tau)] X(\tau) d\tau \\ &= \int_{-\infty}^{\infty} h(t - \tau) X(\tau) d\tau = (h * X)(t). \clubsuit \end{aligned}$$

11.9 Valor mitjà, autocorrelació i espectre de potència d'un PE transformat linealment

Un procés estocàstic $X(t)$ es caracteritza principalment per les funcions valor mitjà $m(t)$ i autocorrelació $R(t_1, t_2)$. Si tenim altres processos definim en general la funció de correlació de $X(t)$ i $Y(t)$

$$R_{XY}(t_1, t_2) = E[X(t_1)Y(t_2)]. \quad (11.22)$$

La funció d'autocorrelació de $X(t)$ s'escriu ara $R_{XX}(t_1, t_2)$. També indicarem el seu valor mitjà $m_X(t)$. L'operació L es fa sempre sobre funcions d'una variable. Quan el seu argument té dues variables posem L_1 o L_2 per indicar que considerem la funció de la variable t_1 o t_2 deixant l'altra com a paràmetre.

El següent resultat indica com obtenir els paràmetres de $Y(t) = L[X(t)]$ a partir dels de $X(t)$.

$$m_Y(t) = L[m_X(t)], \quad (11.23)$$

$$R_{XY}(t_1, t_2) = L_2[R_{XX}(t_1, t_2)], \quad (11.24)$$

$$R_{YY}(t_1, t_2) = L_1[R_{XY}(t_1, t_2)]. \quad (11.25)$$

DEM:

$$\begin{aligned} m_Y(t) &= E[L[X(t)]] = E\left[\int_{-\infty}^{\infty} h(t-\tau)X(\tau)d\tau\right] = \int_{-\infty}^{\infty} h(t-\tau)E[X(\tau)]d\tau \\ &= \int_{-\infty}^{\infty} h(t-\tau)m_X(\tau)d\tau = (h * m_X)(t) = L[m_X(t)]. \end{aligned}$$

$$\begin{aligned} R_{XY}(t_1, t_2) &= E[X(t_1)Y(t_2)] = E\left[X(t_1)\int_{-\infty}^{\infty} h(t_2-\tau)X(\tau)d\tau\right] = \int_{-\infty}^{\infty} h(t_2-\tau)E[X(t_1)X(\tau)]d\tau \\ &= \int_{-\infty}^{\infty} h(t_2-\tau)R_{XX}(t_1, \tau)d\tau = L_2[R_{XX}(t_1, t_2)]. \end{aligned}$$

$$\begin{aligned} R_{YY}(t_1, t_2) &= E[Y(t_1)Y(t_2)] = E\left[\int_{-\infty}^{\infty} h(t_1-\tau)X(\tau)d\tau \int_{-\infty}^{\infty} h(t_2-\tau)X(\tau)d\tau\right] \\ &= \int_{-\infty}^{\infty} h(t_1-\tau)R_{XY}(\tau, t_2)d\tau = L_1[R_{XY}(t_1, t_2)]. \clubsuit \end{aligned}$$

Finalment trobem l'espectre de potència de $Y(t)$, $S_Y(\omega)$ a partir de $S_X(\omega)$. $X(t)$ és ara un PE estacionari amb funció d'autocorrelació $R_{XX}(t_1, t_2) = R_{XX}(t_2 - t_1)$. Definim la funció de transferència del sistema $H(\omega)$

$$H(\omega) = \int_{-\infty}^{\infty} h(\tau)e^{-i\omega\tau}d\tau. \quad (11.26)$$

Llavors tenim el següent resultat:

$$S_Y(\omega) = |H(\omega)|^2 S_X(\omega). \quad (11.27)$$

Abans de demostrar-ho recordem el teorema de convolució. Si les funcions $f(t)$ i $g(t)$ tenen les transformades de Fourier $F(\omega)$ i $G(\omega)$ llavors el seu producte de convolució $(f*g)(t)$ té transformada de Fourier $F(\omega)G(\omega)$. També tenim que $f(-t)$ es transforma en $F^*(\omega)$.

DEM:

L'espectre de potència de $Y(t)$ s'obté fent la transformada de Fourier de $R_{YY}(t)$.

$$\begin{aligned} R_{XY}(t_1, t_2) &= \int_{-\infty}^{\infty} h(t_2-\tau)R_{XX}(t_1, \tau)d\tau \\ &= \int_{-\infty}^{\infty} h(t_2-\tau)R_{XX}(\tau-t_1)d\tau = \int_{-\infty}^{\infty} h(t_2-t_1-\tau)R_{XX}(\tau)d\tau, \end{aligned}$$

on hem fet servir que R_{XX} depèn només de la diferència de temps i hem fet el canvi de variable $\tau \rightarrow \tau + t_1$. Així veiem que $R_{XY}(t_1, t_2)$ depèn només de $t_2 - t_1$. Ara obtenim

$$\begin{aligned} R_{YY}(t_1, t_2) &= \int_{-\infty}^{\infty} h(t_1-\tau)R_{XY}(\tau, t_2)d\tau \\ &= \int_{-\infty}^{\infty} h(t_1-\tau)R_{XY}(t_2-\tau)d\tau = \int_{-\infty}^{\infty} h(t_1-t_2-\tau)R_{XY}(-\tau)d\tau. \end{aligned}$$

Així veiem que R_{YY} s'obté de la convolució entre $h(t)$ i $R_{XY}(-t)$ i $R_{XY}(t)$ s'obté de la convolució entre $h(t)$ i $R_{XX}(t)$. Passant a transformades de Fourier

$$\begin{aligned} R_{XY}(-t) &\rightarrow [H(\omega)S_X(\omega)]^* = H^*(\omega)S_X(\omega) \\ R_{YY}(t) &\rightarrow S_Y(\omega) = H(\omega) \cdot H^*(\omega)S_X(\omega) = |H(\omega)|^2 S_X(\omega). \clubsuit \end{aligned}$$

Exemple 11.6 Circuit L-R.

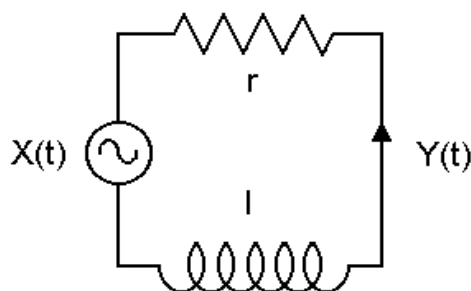


Figura 22: Circuit L-R.

Considerem el sistema format per una resistència r i una bobina d'inductància l . L'entrada al sistema és el voltatge aplicat al circuit donat pel procés $X(t)$. La sortida és la intensitat de corrent que hi circula $Y(t)$. Calculem la transformació $Y(t) = L[X(t)]$ i trobem com varia l'espectre de potència.

Donat $X(t)$ calculem $Y(t)$ resolent l'equació diferencial del circuit

$$l \frac{dY}{dt}(t) + rY(t) = X(t). \quad (11.28)$$

La solució de l'equació homogènia és $Y(t) = Ce^{-\frac{r}{l}t}$ on C és una constant. Ara trobem una solució particular de l'equació completa per variació de constants $Y_p(t) = C(t)e^{-\frac{r}{l}t}$ que implica al substituir en (11.28)

$$C(t) = \frac{1}{l} \int_{-\infty}^t e^{\frac{r}{l}\tau} X(\tau) d\tau.$$

Llavors la solució general de (11.28) és

$$Y(t) = Ce^{-\frac{r}{l}t} + \frac{1}{l} \int_{-\infty}^t e^{-\frac{r}{l}(t-\tau)} X(\tau) d\tau. \quad (11.29)$$

En (11.29) el primer terme és un transitori que desapareix passat un cert temps. Considerem com a sortida la corrent quan s'ha arribat a la situació estacionària i, per tant, fem $C = 0$ en (11.29). Comparant amb l'expressió general

$$Y(t) = \int_{-\infty}^{\infty} h(t-\tau) X(\tau) d\tau \quad (11.30)$$

trobem

$$h(t) = \frac{1}{l} e^{-\frac{r}{l}t} u(t).$$

Ara tenim

$$\begin{aligned} H(\omega) &= \int_{-\infty}^{\infty} \frac{1}{l} e^{-\frac{r}{l}t} u(t) e^{-i\omega t} dt = \frac{1}{l} \int_0^{\infty} e^{-(\frac{r}{l} + i\omega)t} dt \\ &= \frac{1}{r + il\omega}. \end{aligned}$$

Llavors (11.27) ens dona

$$S_Y(\omega) = \frac{S_X(\omega)}{r^2 + l^2\omega^2} \blacklozenge$$

Part IV

Taules

Funció d'error

La primera columna son valors de x . La segona columna son valors de

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt.$$

La funció de distribució d'una variable normal (gaussiana) amb paràmetres m i σ és

$$F(x) = \frac{1}{2} \left(1 + \operatorname{erf} \left(\frac{x - m}{\sqrt{2}\sigma} \right) \right).$$

La taula conté valors $0 \leq x < 3$. La funció erf és imparella així que, per exemple, $\operatorname{erf}(-1.5) = -\operatorname{erf}(1.5)$. Per $x > 3$ es pot utilitzar l'aproximació

$$\operatorname{erf}(x) \simeq 1 - \frac{\sqrt{\pi}}{(\pi - 1)x + \sqrt{x^2 + \pi}} e^{-x^2}.$$

La funció d'error esta predefinida en MAPLE i és accessible fent directament, per exemple,
> erf(0.34);

x	$\text{erf}(x)$	x	$\text{erf}(x)$	x	$\text{erf}(x)$	x	$\text{erf}(x)$	x	$\text{erf}(x)$	x	$\text{erf}(x)$
.00	.000000	.50	.520499	1.00	.842700	1.50	.966105	2.00	.995322	2.50	.999593
.01	.011283	.51	.529243	1.01	.846810	1.51	.967276	2.01	.995524	2.51	.999614
.02	.022564	.52	.537898	1.02	.850838	1.52	.968413	2.02	.995719	2.52	.999634
.03	.033841	.53	.546464	1.03	.854784	1.53	.969516	2.03	.995906	2.53	.999653
.04	.045111	.54	.554939	1.04	.858649	1.54	.970585	2.04	.996085	2.54	.999671
.05	.056371	.55	.563323	1.05	.862436	1.55	.971622	2.05	.996258	2.55	.999689
.06	.067621	.56	.571615	1.06	.866143	1.56	.972628	2.06	.996423	2.56	.999705
.07	.078857	.57	.579815	1.07	.869773	1.57	.973602	2.07	.996582	2.57	.999721
.08	.090078	.58	.587922	1.08	.873326	1.58	.974547	2.08	.996734	2.58	.999736
.09	.101280	.59	.595936	1.09	.876803	1.59	.975462	2.09	.996880	2.59	.999750
.10	.112462	.60	.603856	1.10	.880205	1.60	.976348	2.10	.997020	2.60	.999763
.11	.123622	.61	.611681	1.11	.883533	1.61	.977206	2.11	.997154	2.61	.999776
.12	.134758	.62	.619411	1.12	.886787	1.62	.978038	2.12	.997283	2.62	.999788
.13	.145867	.63	.627046	1.13	.889970	1.63	.978842	2.13	.997407	2.63	.999800
.14	.156947	.64	.634585	1.14	.893082	1.64	.979621	2.14	.997525	2.64	.999811
.15	.167995	.65	.642029	1.15	.896123	1.65	.980375	2.15	.997638	2.65	.999821
.16	.179011	.66	.649376	1.16	.899096	1.66	.981104	2.16	.997747	2.66	.999831
.17	.189992	.67	.656627	1.17	.902000	1.67	.981810	2.17	.997851	2.67	.999840
.18	.200935	.68	.663782	1.18	.904837	1.68	.982492	2.18	.997950	2.68	.999849
.19	.211839	.69	.670840	1.19	.907608	1.69	.983152	2.19	.998045	2.69	.999857
.20	.222702	.70	.677801	1.20	.910313	1.70	.983790	2.20	.998137	2.70	.999865
.21	.233521	.71	.684665	1.21	.912955	1.71	.984407	2.21	.998224	2.71	.999873
.22	.244295	.72	.691433	1.22	.915533	1.72	.985002	2.22	.998307	2.72	.999880
.23	.255022	.73	.698103	1.23	.918050	1.73	.985578	2.23	.998387	2.73	.999886
.24	.265700	.74	.704678	1.24	.920505	1.74	.986134	2.24	.998464	2.74	.999893
.25	.276326	.75	.711155	1.25	.922900	1.75	.986671	2.25	.998537	2.75	.999899
.26	.286899	.76	.717536	1.26	.925235	1.76	.987190	2.26	.998607	2.76	.999905
.27	.297418	.77	.723821	1.27	.927513	1.77	.987690	2.27	.998673	2.77	.999910
.28	.307880	.78	.730010	1.28	.929734	1.78	.988174	2.28	.998737	2.78	.999915
.29	.318283	.79	.736103	1.29	.931898	1.79	.988640	2.29	.998798	2.79	.999920
.30	.328626	.80	.742100	1.30	.934007	1.80	.989090	2.30	.998856	2.80	.999924
.31	.338908	.81	.748003	1.31	.936063	1.81	.989524	2.31	.998912	2.81	.999929
.32	.349125	.82	.753810	1.32	.938065	1.82	.989943	2.32	.998965	2.82	.999933
.33	.359278	.83	.759523	1.33	.940015	1.83	.990346	2.33	.999016	2.83	.999937
.34	.369364	.84	.765142	1.34	.941913	1.84	.990735	2.34	.999064	2.84	.999940
.35	.379382	.85	.770668	1.35	.943762	1.85	.991111	2.35	.999110	2.85	.999944
.36	.389329	.86	.776100	1.36	.945561	1.86	.991472	2.36	.999154	2.86	.999947
.37	.399205	.87	.781439	1.37	.947312	1.87	.991820	2.37	.999196	2.87	.999950
.38	.409009	.88	.786687	1.38	.949016	1.88	.992156	2.38	.999236	2.88	.999953
.39	.418738	.89	.791843	1.39	.950673	1.89	.992479	2.39	.999275	2.89	.999956
.40	.428392	.90	.796908	1.40	.952285	1.90	.992790	2.40	.999311	2.90	.999958
.41	.437969	.91	.801882	1.41	.953852	1.91	.993089	2.41	.999346	2.91	.999961
.42	.447467	.92	.806767	1.42	.955376	1.92	.993378	2.42	.999379	2.92	.999963
.43	.456886	.93	.811563	1.43	.956857	1.93	.993655	2.43	.999410	2.93	.999965
.44	.466225	.94	.816271	1.44	.958296	1.94	.993922	2.44	.999440	2.94	.999967
.45	.475481	.95	.820890	1.45	.959695	1.95	.994179	2.45	.999469	2.95	.999969
.46	.484655	.96	.825423	1.46	.961053	1.96	.994426	2.46	.999496	2.96	.999971
.47	.493745	.97	.829870	1.47	.962372	1.97	.994663	2.47	.999522	2.97	.999973
.48	.502749	.98	.834231	1.48	.963654	1.98	.994892	2.48	.999547	2.98	.999974
.49	.511668	.99	.838508	1.49	.964897	1.99	.995111	2.49	.999570	2.99	.999976

Taula de variables aleatòries

X	Funció de probabilitat	E[X]	V[X]
Bernouilli (p)	$P(X=1) = p, \quad P(X=0) = q$	p	pq
Binomial (n, p)	$P(X=k) = \binom{n}{k} p^k q^{n-k}, \quad k = 0, 1, 2, \dots, n$	np	npq
Geomètrica (p)	$P(X=k) = q^{k-1} p, \quad k = 1, 2, 3, \dots$	$\frac{1}{p}$	$\frac{q}{p^2}$
Poisson (α)	$P(X=k) = \frac{\alpha^k}{k!} e^{-\alpha}, \quad k = 0, 1, 2, \dots$	α	α

X	Funció de densitat	Funció de distribució	E[X]	V[X]
Uniforme (a, b)	$f(x) = \frac{1}{b-a}$ per $x \in (a, b)$	$F(x) = \frac{x-a}{b-a}$ si $x \in [a, b]$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Exponencial (λ)	$f(x) = \lambda e^{-\lambda x}$ per $x \geq 0$	$F(x) = 1 - e^{-\lambda x}$ si $x \geq 0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Normal (m, σ)	$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}} \quad \forall x \in \mathbb{R}$	$F(x) = \frac{1}{2}(1 + \operatorname{erf}(\frac{x-m}{\sqrt{2}\sigma}))$	m	σ^2

En les anteriors taules: $0 < p < 1$ (els casos $p = 0$ i $p = 1$ poden considerar-se com situacions límit), $q = 1 - p$. Les funcions de densitat valen 0 fora dels intervals indicats. La desviació σ es calcula en cada cas com $\sigma = \sqrt{V[X]}$.

Part V

Apèndixs

Apèndix A

Lògica i conjunts

Els objectes elementals en la lògica són els predicats o **proposicions lògiques**. Aquestes proposicions poden tenir dos valors, 0 o 1, assignats. Si la proposició p és certa li assignem el valor 1 i si és falsa el valor 0. A partir de proposicions elementals en construïm d'altres a través d'operacions lògiques. Llavors la qüestió és conèixer el valor d'una proposició complexa a partir del valor de les proposicions elementals que la formen.

Si p i q són proposicions lògiques definim la conjunció $p \wedge q$ (p i q), la disjunció $p \vee q$ (p o q) i la negació $\neg p$ (no p) amb les **taules de veritat**:

p	q	$p \wedge q$
0	0	0
1	0	0
0	1	0
1	1	1

p	q	$p \vee q$
0	0	0
1	0	1
0	1	1
1	1	1

p	$\neg p$
0	1
1	0

Amb aquestes taules es poden demostrar algunes equivalències veient que els valors dels dos costats coincideixen per a totes les eleccions de valors de les proposicions elementals. Algunes d'aquestes propietats són:

$$p \wedge q \leftrightarrow q \wedge p, \quad p \vee q \leftrightarrow q \vee p, \quad (\text{A.1})$$

$$p \wedge (q \wedge r) \leftrightarrow (p \wedge q) \wedge r, \quad p \vee (q \vee r) \leftrightarrow (p \vee q) \vee r, \quad (\text{A.2})$$

$$p \wedge (q \vee r) \leftrightarrow (p \wedge q) \vee (p \wedge r), \quad p \vee (q \wedge r) \leftrightarrow (p \vee q) \wedge (p \vee r), \quad (\text{A.3})$$

$$\neg(p \wedge q) \leftrightarrow (\neg p) \vee (\neg q), \quad \neg(p \vee q) \leftrightarrow (\neg p) \wedge (\neg q). \quad (\text{A.4})$$

Algunes proposicions fan referència a un argument i tenen la forma $p(x)$. El seu valor de veritat depèn llavors del valor de l'argument. Aquestes proposicions ens connecten la lògica amb la teoria de conjunts de la següent manera. Donat el conjunt A li associem la proposició $p(x)$ que és certa si i només si $x \in A$. Inversament, donada una proposició $p(x)$ definim el conjunt $A = \{x \mid p(x)\}$, és a dir, el conjunt dels x tals que $p(x)$ és certa.

Les operacions de conjunts \cap (intersecció), \cup (unió) i complementari tradueixen llavors les operacions lògiques a aquest nou llenguatge. Si A i B són conjunts:

$$A = B \leftrightarrow (x \in A \leftrightarrow x \in B), \quad (\text{A.5})$$

$$A \cap B = \{x \mid x \in A \text{ i } x \in B\}, \quad (\text{A.6})$$

$$A \cup B = \{x \mid x \in A \text{ o } x \in B\}, \quad (\text{A.7})$$

$$\overline{A} = \{x \mid x \notin A\}. \quad (\text{A.8})$$

Les propietats de les operacions lògiques passen ara trivialment a les operacions entre conjunts:

$$A \cap B = B \cap A, \quad A \cup B = B \cup A, \quad (\text{A.9})$$

$$A \cap (B \cap C) = (A \cap B) \cap C, \quad A \cup (B \cup C) = (A \cup B) \cup C, \quad (\text{A.10})$$

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C), \quad A \cup (B \cap C) = (A \cup B) \cap (A \cup C), \quad (\text{A.11})$$

$$\overline{A \cap B} = \overline{A} \cup \overline{B}, \quad \overline{A \cup B} = \overline{A} \cap \overline{B}. \quad (\text{A.12})$$

Les propietats s'anomenen (A.1) i (A.9) commutatives, (A.2) i (A.10) associatives i (A.3) i (A.11) distributives.

Quan la proposició $p(x)$ es falsa per a tot valor de l'argument x li associem el conjunt buit \emptyset . Aquest conjunt no conté cap element i verifica per a tot A

$$A \cap \emptyset = \emptyset, \quad A \cup \emptyset = A. \quad (\text{A.13})$$

Quan la proposició $p(x)$ es certa per a tot valor de l'argument x li associem el conjunt total Ω . Aquest conjunt conté qualsevol element al que es pugui fer referència i és necessari en teoria de conjunts per evitar algunes paradoxes. Ω verifica per a tot A

$$A \cap \Omega = A, \quad A \cup \Omega = \Omega. \quad (\text{A.14})$$

El conjunt $\mathcal{P}(\Omega)$ de tots els subconjunts de Ω amb les operacions descrites anteriorment adquireix una estructura que s'anomena **àlgebra de Boole**. Una característica important és la propietat de **dualitat**: si en una propietat certa es fa el canvi $\cup \leftrightarrow \cap$ i $\emptyset \leftrightarrow \Omega$ s'obté una altra propietat certa.

Apèndix B

Combinatòria

Per a tot enter no negatiu n definim el seu factorial $n!$ de la següent manera:

$$0! = 1, \quad (\text{B.1})$$

$$n! = 1 \cdot 2 \cdot 3 \cdots (n-1) \cdot n. \quad (\text{B.2})$$

La funció factorial és de creixement molt ràpid. Una forma asimptòtica la dóna la fórmula de Stirling

$$n! \simeq \sqrt{2\pi n} \left(\frac{n}{e}\right)^n \quad (\text{B.3})$$

quan $n \rightarrow \infty$. De fet per $n = 10$ l'error és ja només de l'1% . La fórmula de Stirling és útil també per trobar formes asimptòtiques de resultats en combinatòria.

Considerem ara que tenim n objectes distingibles. Els podem pensar com n boles numerades de 1 a n . D'aquests objectes en seleccionem k . Això ho podem fer de dues maneres:

- Ordenats. Quan considerem que, per exemple, triant dos objectes, 1,2 i 2,1 són dues eleccions diferents.
- No ordenats. Quan considerem que, per exemple, triant dos objectes, 1,2 i 2,1 són la mateixa elecció.

Per exemple, suposem que tenim deu ampolles amb diferents licors. Si volem seleccionar tres ampolles per fer un còctel les triarem no ordenades ja que l'ordre en que les posem és irrelevant. Si volem triar tres ampolles per donar el primer, segon i tercer premis d'un concurs les triarem ordenades ja que de les tres ampolles triades hem de decidir quina és la primera, quina la segona i quina la tercera, és a dir, ordenar-les.

Així per $0 \leq k \leq n$ definim els nombres:

- **Variacions** V_n^k : Nombre de maneres de triar k objectes diferents ordenats d'un total de n .
- **Combinacions** C_n^k : Nombre de maneres de triar k objectes diferents no ordenats d'un total de n .
- **Permutacions** P_n : Nombre de maneres d'ordenar n objectes diferents.

Llavors resulten evidents les relacions

$$P_n = V_n^n, \quad (\text{B.4})$$

$$V_n^k = C_n^k P_k. \quad (\text{B.5})$$

Per tant és suficient trobar les variacions per tenir-ho tot. Per calcular V_n^k veiem que podem triar el primer objecte de n maneres, el segon de $n - 1$ maneres, etc. Llavors

$$V_n^k = m \cdot (n - 1) \cdots (n - (k - 1)) = \frac{n!}{(n - k)!}. \quad (\text{B.6})$$

Ara trobem, al fer $k = n$, $P_n = n!$. Igualment veiem

$$C_n^k = \frac{V_n^k}{P_k} = \frac{n!}{k!(n - k)!}. \quad (\text{B.7})$$

Generalment s'escriu $\binom{n}{k} = \frac{n!}{k!(n - k)!}$. La següent propietat és immediata

$$\binom{n}{k} = \binom{n}{n - k}. \quad (\text{B.8})$$

Una altra propietat fàcil de demostrar per càlcul directe és

$$\binom{n}{k} = \binom{n - 1}{k} + \binom{n - 1}{k - 1}. \quad (\text{B.9})$$

A vegades és útil considerar $\binom{n}{k} = 0$ per $k > n$. Amb aquest criteri no cal recordar els límits de sumació en moltes fórmules. Això és consistent amb l'extensió del factorial a través de la funció gamma (D.10) $x! = \Gamma(x + 1)$ que dona infinit pel factorial dels enters negatius.

Exemple B.1 Binomi de Newton

$$(1 + x)^n = \sum_{k=0}^n \binom{n}{k} x^k. \quad (\text{B.10})$$

En efecte,

$$(1 + x)^n = \overbrace{(1 + x) \cdot (1 + x) \cdots (1 + x)}^n = \sum_{k=0}^n a_k x^k.$$

El producte dels n $1 + x$ dona 2^n termes. a_k és el nombre d'aquests termes iguals a x^k . Aquest es genera quan dels n termes, en k agafem x i en $n - k$ agafem 1. La multiplicitat apareix perquè poden triar les k x de diferents maneres, exactament C_n^k .

Exemple B.2 Sigui A un conjunt finit de n elements. Diem llavors que el seu cardinal és n i posem $|A| = n$. El conjunt dels subconjunts de A el designem $\mathcal{P}(A)$ (parts de A). Que val $|\mathcal{P}(A)|$?

És a dir, quants subconjunts té un conjunt de n elements?

Contem primer quants subconjunts de k elements té A i després sumem per $k = 0, \dots, n$. La llibertat que tenim per triar subconjunts de k elements és quins són aquests elements. Com un conjunt no implica cap ordenació dels seus elements hi han C_n^k subconjunts de k elements. Ara, en total tenim

$$\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{n} = (1 + 1)^n = 2^n.$$

subconjunts. Així

$$|\mathcal{P}(A)| = 2^{|A|}. \quad (\text{B.11})$$

Una manera alternativa de trobar aquest resultat és la següent. Podem representar cada subconjunt de A amb una paraula binària de n lletres. Cada posició en la paraula representa un

element de A . Per un subconjunt donat cada lletra val 1 o 0 segons l'element hi pertanyi o no. Per exemple, al subconjunt $\{2, 3, 5\} \subset \{1, 2, 3, 4, 5, 6\}$ li correspon la paraula 011010. És evident que hi ha una correspondència biunívoca entre aquestes paraules i els subconjunts. També es immediat el fet que podem formar 2^n paraules.

Exemple B.3 Funcions generadores

Hi han nombroses propietats dels nombres combinatoris i aquestes poden resultar difícils de demostrar directament. Una manera d'obtenir moltes relacions és utilitzar funcions generadores. Per exemple, tots els combinatoris de n elements apareixen en (B.10). Així les propietats d'aquests nombres estan codificades en les propietats de la funció $(1+x)^n$. Es pot, llavors, anar traduint d'un costat a l'altre per diferents propietats i funcions generadores.

Per exemple, sabem trivialment que $(1+x)^{2n} = (1+x)^n(1+x)^n$. Llavors

$$\sum_{k=0}^{2n} \binom{2n}{k} x^k = \sum_{r=0}^n \binom{n}{r} x^r \sum_{s=0}^n \binom{n}{s} x^s = \sum_{r,s=0}^n \binom{n}{r} \binom{n}{s} x^{r+s} = \sum_{k=0}^{2n} \left\{ \sum_{r+s=k} \binom{n}{r} \binom{n}{s} \right\} x^k.$$

Comparant els polinomis dels dos extrems deduïm

$$\binom{2n}{k} = \sum_{r=0}^k \binom{n}{r} \binom{n}{k-r}. \quad (\text{B.12})$$

Apèndix C

Sèries

Les variables aleatòries discretes amb espai mostral infinit numerable porten a la suma de sèries per resoldre alguns problemes. Donat el context en que apareixen i el fet que aquestes sèries solen ser de termes positius, la qüestió de la convergència és generalment bastant simple. Les sèries més corrents són les geomètriques, associades a les variables del mateix nom. Si definim

$$S_n = \sum_{k=0}^n r^k, \quad (\text{C.1})$$

es veu immediatament que $rS_n = S_n + r^{n+1} - 1$ d'on traiem

$$\sum_{k=0}^n r^k = \frac{1 - r^{n+1}}{1 - r}. \quad (\text{C.2})$$

Aquest resultat val per a tot $r \neq 1$. Si $|r| < 1$ podem fer el límit $n \rightarrow \infty$ i concloure

$$\sum_{k=0}^{\infty} r^k = \frac{1}{1 - r}. \quad (\text{C.3})$$

Notem que la sèrie següent és un cas de la sèrie geomètrica. Per $\alpha > 0$

$$\sum_{k=0}^{\infty} e^{-\alpha k} = \frac{1}{1 - e^{-\alpha}}. \quad (\text{C.4})$$

A vegades els sumatoris anteriors comencen en $k = 1$. Si restem el terme de la sèrie per $k = 0$ trobem

$$\sum_{k=1}^n r^k = \frac{r - r^{n+1}}{1 - r}, \quad (\text{C.5})$$

$$\sum_{k=1}^{\infty} r^k = \frac{r}{1 - r}. \quad (\text{C.6})$$

Moltes de les sèries que se solen utilitzar s'obtenen com a sèrie de Taylor d'alguna funció

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k. \quad (\text{C.7})$$

En (C.7) $f^{(k)}$ indica la derivada k -èsima. La sèrie es considera al voltant de 0 per ser el cas més utilitzat. La igualtat en (C.7) només és certa per x pertanyent a una regió anomenada

interval de convergència. En molts casos l'interval de convergència és tot \mathbb{R} . En particular les sèries geomètriques (C.3) i (C.6) es poden considerar sèries de Taylor de les funcions de la dreta convergint en l'interval $(-1, 1)$. Altres exemples importants són

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots = \sum_{k=0}^{\infty} \frac{x^k}{k!}. \quad (\text{C.8})$$

$$\cosh(x) = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots = \sum_{k=0}^{\infty} \frac{x^{2k}}{(2k)!}. \quad (\text{C.9})$$

$$\sinh(x) = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots = \sum_{k=0}^{\infty} \frac{x^{2k+1}}{(2k+1)!}. \quad (\text{C.10})$$

Les sèries (C.8), (C.9) i (C.10) convergeixen per a tot x .

Apèndix D

Integrals

Algunes integrals impròpies apareixen sovint en l'estudi de variables aleatòries contínues, especialment en les exponencials i les gaussianes.

Un resultat fàcil de demostrar per inducció és

$$\int_0^\infty x^n e^{-x} dx = n!. \quad (\text{D.1})$$

Amb el canvi de variable $\lambda x = t$ trobem que

$$\int_0^\infty x^n e^{-\lambda x} dx = \frac{n!}{\lambda^{n+1}}. \quad (\text{D.2})$$

Les integrals gaussianes són les del tipus polinomi per exponencial d'un polinomi de segon grau. El resultat bàsic és

$$\int_{-\infty}^\infty e^{-x^2} dx = \sqrt{\pi}. \quad (\text{D.3})$$

DEM: Si I és l'anterior integral

$$\begin{aligned} I^2 &= \int_{-\infty}^\infty e^{-x^2} dx \int_{-\infty}^\infty e^{-y^2} dy = \int_{-\infty}^\infty \int_{-\infty}^\infty e^{-(x^2+y^2)} dx dy \\ &= \int_0^\infty \int_0^{2\pi} e^{-r^2} r dr d\theta = 2\pi \int_0^\infty r e^{-r^2} dr = \pi \int_0^\infty e^{-t} dt = \pi \end{aligned}$$

on s'ha fet la integral doble passant a polars i després s'ha fet el canvi $r^2 = t$. ♣

Degut a la simetria de e^{-x^2}

$$\int_0^\infty e^{-x^2} dx = \frac{\sqrt{\pi}}{2}. \quad (\text{D.4})$$

Un resultat més general és

$$\int_{-\infty}^\infty e^{-ax^2+bx+c} dx = e^{\frac{b^2}{4a}+c} \sqrt{\frac{\pi}{a}}. \quad (\text{D.5})$$

La integral és convergent per $a > 0$.

DEM: Posem l'exponent $-ax^2 + bx + c = -a(x - \frac{b}{2a})^2 + \frac{b^2}{4a} + c$ i fem el canvi $\sqrt{a}(x - \frac{b}{2a}) = t$. Llavors la integral és

$$\int_{-\infty}^\infty e^{-t^2 + \frac{b^2}{4a} + c} \frac{dt}{\sqrt{a}} = e^{\frac{b^2}{4a} + c} \sqrt{\frac{\pi}{a}}. \clubsuit$$

El següent resultat és directe pel fet que la funció integrada és imparell ($f(-x) = -f(x)$):

$$\int_{-\infty}^\infty x e^{-x^2} dx = 0. \quad (\text{D.6})$$

El mateix és cert si posen x^n amb n senar en lloc de x davant l'exponencial. També és trivial veure, fent el canvi $x^2 = t$

$$\int_0^\infty x e^{-x^2} dx = \frac{1}{2}. \quad (\text{D.7})$$

El mateix canvi i l'ús de (D.1) ens dona que, per n senar,

$$\int_0^\infty x^n e^{-x^2} dx = \frac{1}{2} \left(\frac{n-1}{2} \right)!. \quad (\text{D.8})$$

El cas de potències parells es pot deduir derivant amb respecte a el resultat

$$\int_{-\infty}^\infty e^{-ax^2} dx = \sqrt{\frac{\pi}{a}}, \quad (\text{D.9})$$

o també integrant per parts. Per exemple

$$\int_0^\infty x^2 e^{-x^2} dx = -\frac{x}{2} e^{-x^2} \Big|_0^\infty + \frac{1}{2} \int_0^\infty e^{-x^2} dx = \frac{\sqrt{\pi}}{4}.$$

Molts dels resultats anteriors es sistematitzen amb l'ús de la **funció gamma d'Euler**. Motivada per (D.1) definim per $z > 0$

$$\Gamma(z) = \int_0^\infty x^{z-1} e^{-x} dx. \quad (\text{D.10})$$

Per integració per parts es demostra

$$\Gamma(z+1) = z\Gamma(z). \quad (\text{D.11})$$

Definim el factorial d'un nombre real com $z! = \Gamma(z+1)$. Per a $z = 0, 1, 2, \dots$ coincideix amb el factorial ordinari. Per valors generals de z s'ha de recórrer a taules o calculadores però per z semienter es pot calcular. Si fem el canvi $x = t^2$:

$$\Gamma(z) = 2 \int_0^\infty t^{2z-1} e^{-t^2} dt. \quad (\text{D.12})$$

Ara és immediat veure que

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}. \quad (\text{D.13})$$

i per $n = 1, 2, 3, \dots$, utilitzant (D.11),

$$\Gamma\left(n + \frac{1}{2}\right) = \left(n - \frac{1}{2}\right) \left(n - \frac{3}{2}\right) \cdots \frac{1}{2} \Gamma\left(\frac{1}{2}\right) = \frac{(2n-1)(2n-3) \cdots 3 \cdot 1}{2^n} \sqrt{\pi} = \frac{(2n)!}{2^{2n} n!} \sqrt{\pi}. \quad (\text{D.14})$$

Apèndix E

Distribucions

Considerem un espai \mathcal{F} els elements del qual són funcions reals d'una variable real $f(t)$. Suposem que aquestes funcions són contínues i prou regulars perquè les operacions que considerarem tinguin sentit. Una **distribució** és una aplicació lineal que assigna un nombre real a cada element d'aquest espai. Si anomenem T a una d'aquestes distribucions llavors indiquem amb $\langle T, f \rangle$ la imatge de $f(t)$ sota T . Donada una funció $\alpha(t)$ podem associar-li una distribució T_α definida

$$\langle T_\alpha, f \rangle = \int_{-\infty}^{\infty} \alpha(t) f(t) dt. \quad (\text{E.1})$$

(Suposem que l'espai \mathcal{F} conté només funcions que assegurin la convergència de l'anterior integral.) També està clar que existeixen distribucions que no poden obtenir-se a partir de cap funció pel procediment anterior. La distribució **delta de Dirac** es defineix

$$\langle \delta_a, f \rangle = f(a). \quad (\text{E.2})$$

Així el concepte de distribució es més ampli que el de funció i per això a vegades s'anomena funcions generalitzades a les distribucions.

Tot i que la delta de Dirac no correspon a cap funció, sempre es representa de manera simbòlica com si vingués donada per una funció $\delta(t)$. Llavors representem (E.2) quan $a = 0$ com un cas de (E.1) i posem

$$\int_{-\infty}^{\infty} \delta(t) f(t) dt = f(0). \quad (\text{E.3})$$

En general (E.2) es posa

$$\int_{-\infty}^{\infty} \delta(t-a) f(t) dt = f(a), \quad (\text{E.4})$$

que es pot obtenir fent el canvi de variable $t \rightarrow t+a$ en la integral i aplicant (E.3).

Encara que cap funció verifiqui la propietat (E.4) si que podem obtenir-la com a límit d'una successió de distribucions donades per funcions. Així posant per exemple

$$\phi_\epsilon(t) = \frac{1}{\sqrt{2\pi\epsilon}} e^{-\frac{t^2}{2\epsilon^2}}$$

tenim que

$$f(a) = \langle \delta_a, f \rangle = \lim_{\epsilon \rightarrow 0} \int_{-\infty}^{\infty} \phi_\epsilon(t-a) f(t) dt.$$

Si bé no té sentit rigurós podem visualitzar la “funció” delta com $\delta(t-a) = \lim_{\epsilon \rightarrow 0} \phi_\epsilon(t-a)$ que ens dóna

$$\delta(t-a) = \begin{cases} 0 & \text{si } t \neq a \\ \infty & \text{si } t = a \end{cases} \quad (\text{E.5})$$

Això ens dóna una imatge intuitiva de la delta però el que la defineix realment és (E.4).

Les distribucions es poden sotmetre a operacions definides en principi només per funcions. Per exemple, veiem quina és la distribució associada a la derivada d'una funció.

$$\langle T_{\alpha'}, f \rangle = \int_{-\infty}^{\infty} \alpha'(t) f(t) dt = - \int_{-\infty}^{\infty} \alpha(t) f'(t) dt = - \langle T_{\alpha}, f' \rangle$$

on s'ha fet una integració per parts i considerat que les funcions tendeixen prou ràpidament cap a 0 en els extrems. Ara podem definir la derivada de qualsevol distribució T com una nova distribució T' que actua

$$\langle T', f \rangle = - \langle T, f' \rangle \quad (\text{E.6})$$

Això implica que, per exemple,

$$\int_{-\infty}^{\infty} \delta'(t-a) f(t) dt = -f'(a). \quad (\text{E.7})$$

Igualment considerem la **funció de Heaviside** $u_a(t) = u(t-a)$ definida

$$u(t-a) = \begin{cases} 0 & \text{si } t < a \\ 1 & \text{si } t \geq a \end{cases} \quad (\text{E.8})$$

La seva distribució associada la designem amb el mateix nom i posem

$$\langle u_a, f \rangle = \int_a^{\infty} f(t) dt. \quad (\text{E.9})$$

Ara és fàcil demostrar que $u'(t-a) = \delta(t-a)$.

$$\langle u'_a, f \rangle = - \int_a^{\infty} f'(t) dt = - f(t)|_a^{\infty} = f(a) = \langle \delta_a, f \rangle.$$

Una altra propietat de la delta és que per a tota funció g

$$g(t)\delta(t-a) = g(a)\delta(t-a), \quad (\text{E.10})$$

que es demostra multiplicant els dos costats per una funció arbitrària i integrant. Una conseqüència de (E.10) és que $t\delta(t) = 0$.

Apèndix F

Simulació numèrica de variables i processos aleatoris

Els llenguatges i entorns de programació solen oferir algun generador de nombres aleatoris. Aquest generador es presenta sota la forma d'una funció (mètode, procediment, subrutina) com `random()` que dona com a sortida un valor real (`float` o `double`) uniforme en $[0, 1]$. El mètode d'obtenció d'aquests nombres pseudo-aleatoris consisteix en realitzar operacions modulars amb enters grans i utilitzar part dels seus dígitos per construir-los. El generador sol incorporar un procediment d'inicialització que utilitza, per exemple, la data (en mil·lisegons) proveïda pel sistema operatiu.

Aquest generador bàsic ve donat pel sistema i no podem manipular-lo directament però hi han tècniques per, a partir dels seus valors obtenir-ne d'altres amb millors propietats. Els requisits dependran de la quantitat de nombres que necessitem. En simulacions molt extenses, el fet que el generador bàsic és necessàriament periòdic podria falsejar els resultats. Aquestes qüestions es poden estudiar en els llibres tipus *numerical recipes*. En el que segueix, suposem que tenim un generador amb bones propietats.

Denotem G la variable uniforme en $[0, 1]$ donada pel generador. Si volem obtenir una variable contínua X amb funció de distribució $F_X(x)$ ho podem fer com $X = F_X^{-1}(G)$ ja que la variable $F_X(X)$ pren valors en $[0, 1]$ i té densitat

$$f_X(x) \cdot \frac{1}{F'_X(x)} = 1.$$

Per exemple si volem obtenir una variable exponencial de paràmetre λ , $F_X(x) = 1 - e^{-\lambda x}$ i serà $X = -\lambda^{-1} \ln(1 - G)$.

Aquest mètode es pot aplicar si F_X és estrictament creixent sobre Ω_X . També cal poder invertir explícitament F_X . Això no es pot fer en el cas de les variables gaussianes. Per generar una variable normal amb $m = 0, \sigma = 1$, considerem una normal bidimensional amb paràmetres $m_1 = m_2 = \rho = 0, \sigma_1 = \sigma_2 = 1$, és a dir, dues normals independents amb paràmetres normalitzats. La seva densitat és

$$f(x_1, x_2) = \frac{1}{2\pi} e^{-\frac{x_1^2 + x_2^2}{2}}$$

i, en polars ($x_1 = r \cos \theta, x_2 = r \sin \theta$):

$$f(r, \theta) = \frac{1}{2\pi} r e^{-\frac{r^2}{2}}$$

on veiem que θ és uniforme en $[0, 2\pi]$, r té densitat $f(r) = r e^{-\frac{r^2}{2}}$ i les dues són independents. A partir de dos valors G_1, G_2 del generador obtenim $\theta = 2\pi G_2, r = \sqrt{-2 \ln(1 - G_1)}$. Així s'obtenen variables normals per parelles:

$$X_1 = \sqrt{-2 \ln(1 - G_1)} \cos(2\pi G_2),$$

$$X_2 = \sqrt{-2 \ln(1 - G_1)} \sin(2\pi G_2).$$

(En les anteriors expressions es pot canviar $1 - G_1$ per G_1 ja que són el mateix tipus de variable.)

Per simular processos, es generen seqüències de valors corresponents a instants temporals equiespaiats. Per un procés normal estacionari incorrelat amb $m(t) = 0$ simplement cal una llista de valors normals independents. Per exemple, la rutina que segueix torna el procés en forma d'array de reals

```
float PI=3.1425927;
float x[1000]; //proces amb 1000 mostres

x=proces_normal_incorrelat(1000);

float x[] proces_normal_incorrelat(int n){
    for(int k=0;k<(n/2);k++){
        r=sqrt(-2*ln(random()));
        teta=2*PI*random()
        x[2*k]=r*cos(teta);
        x[2*k+1]=r*sin(teta);
    }
}
```

Veiem ara com es pot simular un procés normal estacionari $X(t)$ amb $m_X(t) = 0$ i espectre de potència $S_X(\omega)$. Sigui $\hat{Y}(\omega)$ un procés normal incorrelat, és a dir

$$E[\hat{Y}(\omega)] = 0, \quad E[\hat{Y}(\omega_1)\hat{Y}^*(\omega_2)] = \delta(\omega_1 - \omega_2).$$

Pensem en aquest procés com la transformada de Fourier d'un cert procés temporal. Per una funció fixada $A(\omega)$ considerem el procés $\hat{X}(\omega) = A(\omega)\hat{Y}(\omega)$ i sigui $X(t)$ el procés donat per la seva transformada inversa

$$X(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{X}(\omega) e^{j\omega t} d\omega.$$

És immediat veure que $m_X(t) = 0$. La seva funció d'autocorrelació val

$$\begin{aligned} R_X(t_1, t_2) &= E[X(t_1)X^*(t_2)] = E\left[\frac{1}{2\pi} \int_{-\infty}^{\infty} A(\omega)\hat{Y}(\omega)e^{j\omega t_1} d\omega \frac{1}{2\pi} \int_{-\infty}^{\infty} A^*(\omega')\hat{Y}^*(\omega')e^{-j\omega' t_2} d\omega'\right] \\ &= \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A(\omega)A^*(\omega')e^{j\omega t_1}e^{-j\omega' t_2} E[\hat{Y}(\omega)\hat{Y}^*(\omega')] d\omega d\omega' \\ &= \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A(\omega)A^*(\omega')e^{j\omega t_1}e^{-j\omega' t_2} \delta(\omega - \omega') d\omega d\omega' = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} |A(\omega)|^2 e^{j\omega(t_1 - t_2)} d\omega \end{aligned}$$

d'on veiem que $S_X(\omega) = \frac{|A(\omega)|^2}{2\pi}$. Així només cal triar $A(\omega) = \sqrt{2\pi S_X(\omega)}$. El procés es genera fent $\hat{Y}(\omega)$ procés normal incorrelat, escalant les seves components amb $A(\omega)$ i fent una FFT inversa.

Un altre procediment és partir d'un procés temporal de valor mitjà 0, incorrelat, $Y(t)$ i convolucionar-lo amb una funció fixada $K(t)$ de suport compacte (caiguda ràpida). El procés $X(t) = (K * Y)(t)$ ja té una certa correlació que depèn de K :

$$R_X(t_1, t_2) = E[X(t_1)X(t_2)] = E\left[\int_{-\infty}^{\infty} K(t_1 - \tau_1)Y(\tau_1) d\tau_1 \int_{-\infty}^{\infty} K(t_2 - \tau_2)Y(\tau_2) d\tau_2\right]$$

$$\begin{aligned}
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K(t_1 - \tau_1) K(t_2 - \tau_2) E[Y(\tau_1) Y(\tau_2)] d\tau_1 d\tau_2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K(t_1 - \tau_1) K(t_2 - \tau_2) \delta(\tau_1 - \tau_2) d\tau_1 d\tau_2 \\
&= \int_{-\infty}^{\infty} K(t_1 - \tau_1) K(t_2 - \tau_1) d\tau_1 = \int_{-\infty}^{\infty} K(-\tau_1) K(t_2 - t_1 - \tau_1) d\tau_1
\end{aligned}$$

que, si K és parell, resulta ser $K * K$. Si K té forma gaussiana, R_X també la té. Si, per exemple, K és un pols quadrat, R_X és un pols triangular, etc.

Per obtenir processos de Poisson o relacionats amb ell es suficient tenir la llista d'instants en que es produeixen els esdeveniments, i això s'obté sumant exponencials independents:

```
float t[1000]; //proces amb 1000 esdeveniments
```

```
t=instants_poisson(1000, 2.5); // lambda=2,5
```

```
float t[] instants_poisson(int n, float lambda){
    float lambda_inv = 1/lambda;
    float temp = 0;
    for(int k=0;k<n;k++){
        temp += -2*lambda_inv*ln(random());
        t[k]=temp;
    }
}
```


Part VI

Problemes

PROBLEMES

1. Es genera un nombre de n xifres a l'atzar. Calculeu la probabilitat que surti cap-i-cua.
2. Es tira un dau sis vegades. Quina és la probabilitat que surtin els sis resultats diferents?
3. Què és més probable: Treure alguna vegada un as tirant un dau quatre vegades o treure alguna vegada dos asos tirant dos daus 24 vegades?
4. En una loteria de tipus **6/49** tenim un requadre amb 49 nombres dels quals hem de marcar-ne 6. El resultat del sorteig són els 6 nombres premiats més un altre anomenat complementari. El premi depèn de quants encerts tenim. En el cas de 5 encerts, el premi és diferent si el restant nombre que tenim és el complementari o no. Calculeu les probabilitats dels diferents nombres d'encerts ($P(0), P(1), P(2), P(3), P(4), P(5 - c), P(5 + c), P(6)$).
Pels 4 resultats menys probables calculeu la probabilitat d'obtenir-los alguna vegada si juguem cada dia durant 30 anys.
5. Una baralla francesa sense comodins té 52 cartes (4 pals, 13 nombres). Calculeu la probabilitat de les següents jugades en una ma de pòquer (5 cartes): parella, doble parella, trio, escala, full, color, pòquer, escala de color, escala real.
6. Es permuten a l'atzar les 8 lletres AAAACRTT. Quina és la probabilitat que surti CATARATA?
7. Quant d'un total de n objectes distingibles em volem triar k de manera ordenada i amb la possibilitat de repetició, es parla de **variacions amb repetició**. El nombre de configuracions és evident que val

$$VR_n^k = n^k.$$

El cas en que tenim repetició però no fem la tria ordenada s'anomena **combinacions amb repetició** i és menys immediat de calcular. (Per exemple, si $n = 3$ i $k = 2$ hi han 6 eleccions: 11, 22, 33, 12, 23, 13.) Demostreu que

$$CR_n^k = \binom{n+k-1}{k}.$$

(Indicació: Pot ajudar pensar el problema així. Tenim n caselles i hem de posar k boles idèntiques en les caselles de manera que cada casella pot tenir cap, una o varies boles.)

8. De quantes maneres es poden posar k boles idèntiques en n caselles ($k \geq n$) si cap casella pot quedar buida?
9. De quantes maneres podem posar dues torres blanques en un tauler d'escacs sense que s'amenacin?
De quantes maneres podem posar n torres blanques en un tauler de mida $N \times N$ sense que s'amenacin?
10. Extraïem $2k$ boles d'una urna que conté n boles blanques i n boles negres. Calculeu la probabilitat d'extraure el mateix nombre de boles de cada color si fem l'extracció
 - (a) sense reemplaçament.
 - (b) amb reemplaçament.

Calculeu el límit d'aquestes probabilitats quan $n \rightarrow \infty$. (Indicació: Per a fer el límit, demostreu primer que per k fixada i n gran $\binom{n}{k} \simeq \frac{n^k}{k!}$.)

11. Una urna conté 3 boles blanques, 5 boles vermelles i 8 boles negres. S'hi extreuen dues boles (sense reemplaçament) a l'atzar.
 - (a) Quina és la probabilitat que siguin de color diferent?
 - (b) Si les boles extretes són del mateix color, quina és la probabilitat que siguin negres?
 - (c) Són els esdeveniments $A = \text{"Surt alguna bola blanca"}$ i $B = \text{"Surt alguna bola negra"}$ independents?
12. Algunes qüestions sobre independència:
 - (a) Demostreu que A i B independents implica A i \overline{B} independents. Demostreu que també implica \overline{A} i \overline{B} independents.
 - (b) Demostreu que, per a tot esdeveniment A , Ω i A són independents i \emptyset i A són independents.
 - (c) Poden dos esdeveniments no trivials ser incompatibles i independents a la vegada?
13. En l'experiment de tirar un dau, trobeu dos esdeveniments no trivials que siguin independents.
14. La urna A conté 2 boles blanques i 3 boles negres mentre que la urna B conté 4 boles blanques i 3 boles negres. Es passen dues boles a l'atzar de la urna A a la B i després s'extreuen dues boles de B .
 - (a) Si aquestes boles són blanques, quina és la probabilitat que haguem passat alguna bola negra de A a B ?
 - (b) El fet que les boles finals siguin blanques afavoreix que les boles que hagin passat de A a B siguin blanques. Però la probabilitat anterior és bastant gran. Calculeu la probabilitat anterior, ara sense condicionar, per veure que el resultat és consistent.
15. Dos jugadors tiren alternativament un dau. Guanya el joc el primer que treu un 6. Quina és la probabilitat que guanyi el que comença a jugar?
16. Tres jugadors tiren alternativament una moneda amb probabilitat de cara p . Guanya el primer que treu cara. Calculeu la probabilitat que té cada jugador de guanyar.
17. Dos jugadors A i B tiren simultàniament un dau cadascun. El joc es repeteix fins que algú treu un 6 i guanya, llevat que treguin un 6 els dos jugadors declarant-se un empat. Calculeu les probabilitats $P(\text{empat})$, $P(\text{guanya } A)$, $P(\text{guanya } B)$.
18. Una xarxa de comunicacions consisteix de nodes units per línies. La xarxa transmet paquets de manera que si un paquet es troba en un node intern x (node intern és el que està connectat a més d'un node) tria aleatòriament el node de sortida. La probabilitat de sortir pel node y connectat a x val p_{xy} de manera que $\sum_y p_{xy} = 1$. Quan un paquet arriba a un node extern X ja s'hi queda. Denotem p_X a la probabilitat d'anar-hi quan estem al node connectat a ell. Ens plantegem calcular la probabilitat $P(xX)$ que el paquet, trobant-se al node intern x , acabi al node extern X .
 - (a) Resoleu el problema per una xarxa amb un node extern A connectat a un node intern a connectat a un node intern b connectat a un node extern B . (Noteu que els únics paràmetres lliures són p_A i p_B .)
 - (b) Resoleu el problema per una xarxa triangular amb nodes a, b, c connectats respectivament a nodes externs A, B, C .

(Indicació: L'arbre d'esdeveniments en el segon cas és complicat. Resoleu el problema plantejant equacions lineals que han de verificar les diferents probabilitats $P(xY)$.)

19. Sis jugadors s'assignen els nombres 1, 2, 3, 4, 5, 6 i fan el següent joc. Un d'ells tira un dau. Si li surt el seu nombre, guanya. Si no, passa el dau al jugador corresponent al nombre que ha sortit. Aquest jugador repeteix el joc, etc. Calculeu la probabilitat que té de guanyar cadascun del sis jugadors. (Indicació: Formuleu el problema en els termes del problema anterior.)
20. A partir del resultat conegut $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ dedueix una fórmula anàloga per $P(A \cup B \cup C)$.
Generalitzeu el resultat per $P(A_1 \cup A_2 \cup \dots \cup A_n)$. Això es coneix com fórmula o **principi d'inclusió-exclusió**. Els tres problemes següents es resolen aplicant aquesta fórmula.
21. Una secretària ha d'enviar n cartes. Com té pressa per acabar, col·loca les cartes en els sobres de manera aleatòria. Quina és la probabilitat que ninguna carta arribi al seu destinatari?
Calculeu el límit d'aquesta probabilitat per $n \rightarrow \infty$. Compareu-la per $n = 10$ i $n = 100.000$.
22. Tirant un dau 10 vegades, quina és la probabilitat que apareguin els sis resultats possibles?
Generalitzeu a n tirades. Comproveu que $n = 6$ dóna el mateix que el segon problema.
23. El detergent HOMO dóna una lletra del conjunt $\{H, O, M\}$ en cada paquet i fa un regal als que completen la paraula HOMO. Si les tres lletres apareixen amb la mateixa probabilitat, quina es la probabilitat de formar la paraula comprant 10 paquets?
24. Us esteu enfrontant a un pistoler en un descampat i tant a vosaltres com a ell us queden dues bales. Com us trobeu a certa distància, la probabilitat de liquidar a l'adversari amb un tret val $p = 0,3$.
 - (a) Teniu que decidir si convé fer els dos trets abans que el vostre adversari, o si és preferible esperar a que aquest dispari primer (si un esgota les seves bales sense fer "diana" l'altre es pot acostar i disparar amb certesa). Discutiu també com depèn l'estratègia del valor de p .
 - (b) Si no podem evitar que l'altre dispari primer, que és millor a priori (sense saber l'efecte d'aquest tret): esperar a que faci el seu segon tret, disparar dues vegades nosaltres o disparar una vegada i esperar?
25. Donat un grup de 30 persones, quina és la probabilitat que totes celebrin l'aniversari en dies diferents?
26. Un canal transmet missatges consistents en seqüències de 0's i 1's. La probabilitat que un símbol arribi canviat val ϵ . Comproveu si els següents procediments redueixen la probabilitat que un símbol es detecti erròniament:
 - (a) Cada símbol s'envia tres vegades i es tria el resultat més abundant.
 - (b) Cada símbol es va enviant fins que algun valor ha aparegut dues vegades. Es tria aquest valor.
27. Es repeteix un joc amb probabilitat de guanyar p fins que hem guanyat n vegades. Determineu la funció de probabilitat del nombre necessari de jugades X . (Noteu que $n = 1$ és la variable geomètrica.)
28. Una variable aleatòria X discreta pren valors $0, 1, 2, \dots$ amb probabilitat $P(n) = Aa^n$. Digueu quins valors pot prendre a i calculeu A . Calculeu les probabilitats que
 - (a) $X > 4$.
 - (b) X sigui parell.

(c) X sigui divisible per 3 sabent que X és parell.

29. Si X és una variable aleatòria de Cauchy, demostreu que $Y = \frac{1}{X}$ també ho és. Com es relacionen els seus paràmetres?

30. X és una variable uniforme en $[0, 2\pi]$.

(a) Calculeu les funcions de densitat i de distribució de la variable $Y = \cos X$.

(b) Calculeu la funció de distribució de la variable

$$Z = \begin{cases} \sin X & 0 \leq X \leq \pi \\ 0 & \text{altrament} \end{cases}$$

(c) Utilitzeu el teorema de l'esperança per calcular les esperances de Y i Z . Calculeu també $E[Y]$ amb $f_Y(y)$.

31. Calculeu l'esperança de les variables X dels problemes 27 i 28. (En el primer cas sabeu que les probabilitats elementals sumen 1 d'on deduiu

$$\sum_{k=n}^{\infty} \binom{k-1}{n-1} q^{k-n} = \frac{1}{(1-q)^n}$$

i després es pot utilitzar derivació respecte a q .)

32. En un examen la nota mitja val 6 i la desviació 1,8. Quin és el tant per cent d'aprovat?

33. Una variable aleatòria té densitat $f(x) = Kx^2e^{-\lambda x}$, $x > 0$. Calculeu K , $E[X]$, $V[X]$, m_n .

34. Donada una variable X exponencial d'esperança $\frac{1}{3}$,

(a) Calculeu $P(\sin X > \frac{1}{2})$.

(b) Determineu la funció de probabilitat i l'esperança de la variable $Y = [X]$ (part entera de X).

(c) Calculeu $E[X \mid X > \frac{1}{3}]$.

35. La variable aleatòria de *Laplace* és la que té densitat:

$$f(x) = Ke^{-\lambda|x-m|}$$

per $-\infty < x < \infty$. Calculeu K , $E[X]$, $V[X]$. Pel cas $m = 0$, que val la probabilitat que $-\sigma < X < 2\sigma$, on σ és la desviació típica de X ?

36. La temperatura màxima que assoleix un dispositiu electrònic al llarg d'un dia és una variable aleatòria X amb funció de densitat:

$$f_X(x) = \begin{cases} \frac{10-|x-40|}{100} & \text{si } |x-40| < 10 \\ 0 & \text{altrament} \end{cases}$$

El valor de X en qualsevol dia és independent dels altres. Anomenem T el nombre de dies que han de passar fins que $X > 45$. Anomenem N el nombre de dies amb $X < 38$ en un període de 30 dies.

(a) Dibuixeu $f_X(x)$. Calculeu la funció de distribució de X i la probabilitat que $38 < X < 45$.

(b) Calculeu l'esperança i la desviació de X .

(c) Quins tipus de variable són T i N ? Què valen els seus valors mitjans?

(d) Calculeu $P(X \geq 10)$ i $P(N = 7)$.

37. Donada una variable X normal amb paràmetres m, σ :

(a) Calculeu $P(|X - m| < k\sigma)$. Avalueu-ho en tant per cent per a $k = 1$ i $k = 2$.

(b) Calculeu $E[|X - m|]$.

(c) La variable aleatòria *log-normal* és la que té la forma $Y = e^X$. Calculeu els seus moments i, a partir d'ells, la seva esperança i la seva variància.

38. El temps de vida d'una partícula radioactiva és una variable exponencial de paràmetre $\lambda = \frac{1}{\tau_m}$ on τ_m és la vida mitjana de la partícula. En l'instant $t = 0$ tenim una mostra de N_0 partícules. Quin és el nombre mig de partícules en l'instant t ? És correcte formular una llei física que digui que el nombre de partícules en l'instant t val aquesta quantitat?

39. De les taules de mortalitat es dedueix que moren $\alpha(t)$ persones d'edat t per unitat de temps. És a dir, si T és la variable aleatòria que dona el temps de vida d'una persona,

$$f_T(t \mid T \geq t) = \alpha(t).$$

Una hipòtesi realista és la fórmula de Makeham: $\alpha(t) = \alpha + \beta e^{\gamma t}$.

Deduiu amb aquestes dades la funció de distribució $F_T(t)$. Compareu amb el que passaria si T fos exponencial.

40. Una variable aleatòria bidimensional té densitat conjunta $f_{XY}(x, y) = K e^{-2x+y}$ sobre la regió \mathcal{R} . Calculeu les densitats marginals i determineu si són independents en els casos

(a) $\mathcal{R} = \{(x, y) \in \mathbb{R}^2 \mid 0 < x < \infty, 0 < y < 1\}$.

(b) $\mathcal{R} = \{(x, y) \in \mathbb{R}^2 \mid 0 < x < \infty, 0 < y < x\}$.

41. Una moneda de diàmetre 2cm cau a un terra de rajoles quadrades de costat 25cm. Quina és la probabilitat que la moneda no trepitgi més d'una rajola?

Calculeu el mateix amb rajoles hexagonals.

42. Un material presenta impureses aleatòries de forma el·líptica. Els semieixos d'aquestes el·lipses són variables independents X, Y uniformes en $[0, L_1]$, $[0, L_2]$ respectivament. Calculeu la funció de densitat i l'esperança de l'àrea d'aquestes impureses.

43. Tirem un dau 10 vegades.

(a) Calculeu la probabilitat de treure exactament dos uns, dos dosos i dos tresos.

(b) Calculeu la probabilitat de treure algun u i algun dos.

44. $Y = X_1 + X_2$ on X_1 i X_2 són variables aleatòries independents uniformes en $[0, 1]$. Calculeu la densitat de Y .

45. Si A és un esdeveniment aleatori, podem considerar una variable aleatòria Δ_A que val 1 si A es verifica i 0 en cas contrari. Que val $E[\Delta_A]$?

Si tenim dos esdeveniments A, B . Expresseu $\Delta_{A \cap B}$, $\Delta_{A \cup B}$ i $\Delta_{\overline{A}}$ en funció de Δ_A i Δ_B . Comproveu que al prendre esperances en aquestes relacions s'obté la fórmula bàsica de la probabilitat. (Aquestes variables s'anomenen indicadors.)

46. Un centre de comunicacions disposa de 20 ordinadors connectats a un node. Quan arriba una petició de servei en aquest node s'assigna a l'atzar un dels 20 ordinadors (cada ordinador pot atendre més d'un usuari). Si tenim 15 usuaris:

- (a) Per $0 \leq k \leq 15$, calculeu el nombre mig d'ordinadors que atenen exactament a k usuaris.
- (b) En particular, quin és el nombre mig d'ordinadors desocupats?
- (c) Que val la suma per $k = 0, \dots, 15$ dels anteriors resultats? Raoneu el resultat.
- (Indicació: utilitzeu indicadors, noteu la dificultat en tractar la variable que conta els ordinadors amb k usuaris.)
47. Calculeu la covariància en el cas (b) del problema 40. Què podríeu dir de la independència sabent només això?
48. n corredors fan una carrera. El temps que triga el corredor i és una variable exponencial X_i de paràmetre λ_i ($i = 1, \dots, n$). Sigui T la variable aleatòria que dona el temps d'arribada del guanyador. Calculeu la seva densitat i la seva esperança.
49. Dues persones queden per trobar-se en un lloc entre les 11h i les 12h. El primer que arribi només s'esperarà 15 minuts.
- (a) Quina és la probabilitat que es trobin?
- (b) Considereu ara que el primer s'espera fins que arriba el segon. Calculeu la densitat i l'esperança del temps d'espera.
- (c) Calculeu la densitat i l'esperança de l'instant en que arriba la primera persona.
50. L'empresa de pizzes a domicili *PizzaPlus* cobra 6 Euros per una pizza. El temps T d'entrega és una variable aleatòria exponencial d'esperança m . L'empresa considera que una pizza està entregada a temps si $T < 2m$. Si el temps d'entrega és superior a $2m$ es dona la pizza gratuïtament. Si a més T és superior a $3m$ *PizzaPlus* regala al client un postre de 2 Euros.
- (a) Quin és l'ingrés mitjà per pizza repartida?
- (b) De 70 pizzes repartides quin és el nombre mitjà de pizzes entregades a temps? Quina és la probabilitat que se n'entreguin més d'un 20% fora de temps?
- (c) Dos repartidors de pizza independents han entregat les seves pizzes a temps. Quina és la probabilitat que la suma dels seus temps sigui inferior a $3m$?
51. El temps que passa fins que un sistema que funciona té una avaria és una variable exponencial de paràmetre λ . Supposeu que les avaries es produeixen amb independència una de l'altra i que es reparen de manera immediata. Si $T_i, i = 1, \dots, n$ són els instants en que es produeixen les n primeres avaries calculeu la densitat conjunta $f(t_1, t_2, \dots, t_n)$. (Indicació: trobeu primer la densitat dels intervals entre avaries i feu el canvi de variable pertinent.)
52. Al fer una mesura física apareix un error experimental. Degut a això podem considerar que el resultat obtingut a la mesura és una variable aleatòria amb esperança m , el valor exacte de la magnitud que mesurem, i desviació σ , error característic de la mesura. Si s'obté el resultat X_1 llavors expresseu la nostra mesura de m com a $X_1 \pm \sigma$.
- Una manera de reduir l'error és fer varies mesures X_1, X_2, \dots, X_n i donar com a resultat la seva mitjana $Y = \frac{X_1 + X_2 + \dots + X_n}{n}$. Calculeu la desviació de Y (suposant que les mesures són incorrelades) i comproveu que és inferior a σ .
53. Si X i Y són variables exponencials de paràmetre λ independents, determineu el tipus de variable que és $Z = \frac{X}{X+Y}$.
54. Donada la variable bidimensional gaussiana (X, Y) tal que $m_1 = m_2 = 0$, calculeu la densitat conjunta de les variables $Z = \frac{Y}{X}, T = XY$. Demostreu que si X i Y són independents, Z és de Cauchy amb $\alpha = \frac{\sigma_2}{\sigma_1}$.
55. Una urna conté n boles numerades de 1 a n . N'extraiem k sense reemplaçament.

- (a) Calculeu l'esperança de la suma dels seus valors. Que passa si fem l'extracció amb reemplaçament?
- (b) Calculeu la probabilitat que el màxim valor extret sigui m .
56. Donada la variable aleatòria bidimensional (X, Y) uniforme en la regió $x > 0, y > 0, x^2 + y^2 < 1$, calculeu
- (a) Les millors estimacions de X donada Y i de Y donada X .
- (b) Els seus paràmetres utilitzant coordenades polars.
- (c) Les millors estimacions lineals de X donada Y i de Y donada X .
57. Un canal de comunicació li suma al senyal transmès X un soroll Z incorrelat amb X . Volem corregir l'efecte multiplicant la sortida per una constant λ . Trobeu el millor valor de λ si $m_X = m_Z = 0, \sigma_X = 10$ i $\sigma_Z = 5$.
- Calculeu-ho ara en el cas que el soroll sigui independent de X i multiplicatiu en lloc d'additiu. Penseu alguna manera de millorar la predicció.

58. Calculeu el valor mitjà del procés estocàstic

$$X(t) = \begin{cases} 1 - \frac{t}{Y} & 0 \leq t \leq Y \\ \frac{t - Y}{1 - Y} & Y \leq t \leq 1 \end{cases}$$

on Y és una variable aleatòria uniforme en $[0, 1]$.

59. Donat el procés estocàstic

$$X(t) = \begin{cases} t^2 & 0 \leq t \leq A \\ \frac{A^3}{t} & t > A \end{cases}$$

on A és una variable aleatòria exponencial de paràmetre $\lambda = 1$,

- (a) Representeu gràficament el procés: primer per A fixat com a funció de t (realització); després per t fixat com a funció de A .
- (b) Calculeu la seva funció de valor mitjà $m(t)$.
- (c) Calculeu la seva funció de distribució de primer ordre $F(x; t)$. Representeu-la gràficament.
- (d) Calculeu la seva funció de densitat de primer ordre $f(x; t)$ i utilitzeu-la per tornar a calcular $m(t)$.
- (e) Calculeu l'esperança de la variable aleatòria $J = \int_0^\infty X(t)^2 dt$.
60. Un procés estocàstic estacionari té funció d'autocorrelació $R(\tau) = Ke^{-a|\tau|}$. Calculeu la millor estimació lineal en mitjana quadràtica de $X(t)$ a partir de $X(t - T)$ i $X(t - 2T)$. Calculeu l'error quadràtic mig.
61. Donat un procés $X(t)$ amb valor mitjà $m(t)$ i autocovariància $C(t_1, t_2)$, i una constant $T > 0$:
- (a) Calculeu la millor estimació de $X(t)$ a partir de l'estimador $C = \alpha X(t - T) + \beta$, on α i β són constants.
- (b) Particularitzeu el resultat anterior al cas que $X(t)$ sigui estacionari amb valor mitjà m i autocovariància $C(\tau)$.
- (c) Particularitzeu el resultat del primer apartat al cas que $X(t)$ sigui un procés de Poisson de paràmetre λ .

62. En un procés de Poisson de paràmetre λ , T_i , $i = 1, 2, 3$, són els instants en que es produeixen els tres primers esdeveniments. Quin tipus de variable és T_2 si sabem que $T_1 = t_1$ i $T_3 = t_3$? (Indicació: calculeu $f_{T_2}(t_2|T_1 = t_1, T_3 = t_3)$ utilitzant 10.14.)

63. (X, Y) és una variable aleatòria bidimensional uniforme en el triangle delimitat per les rectes $x = 1, y = 0, x = y$.

(a) Calculeu els moments conjunts $m_{r,s}$ de (X, Y) .

(b) Quina és la millor estimació de X donada Y ?

(c) Calculeu la funció de valor mitjà i l'autocorrelació del procés estocàstic:

$$Z(t) = Xt(t - Y).$$

(d) Quina és la millor estimació lineal homogènia de $\int_0^Y Z(t)dt$ donat $Z(\frac{Y}{2})$?

64. Donat el procés estocàstic $X(t) = e^{-At}$ on A és una variable aleatòria uniforme en $[0, 1]$

(a) Calculeu la seva funció de valor mitjà i la seva autocorrelació.

(b) Si $X_1(t)$ és una realització d'aquest procés: que val $\lim_{t \rightarrow \infty} X_1(t)$?

(c) Que podem dir de $\lim_{t \rightarrow \infty} X(t)$?

65. Donat el procés $X(t) = V \cos(\omega t + \phi)$ on ω està fixat, V és una v.a. exponencial de paràmetre $\frac{1}{V_0}$ i ϕ és una v.a. uniforme en $[0, 2\pi]$, V i ϕ independents

(a) Que se'n dedueix amb els criteris d'ergodicitat?

(b) Determineu si és un procés ergòdic.

EL COSTAT FOSC

1. El 28 de gener de 1986 el transbordador espacial *Challenger* amb set tripulants a dins va explotar poc després d'enlairar-se. Aquest desastre, vist per TV per tots els americans va suposar un fort cop a la carrera espacial i va deixar una marca inesborrable al subconscient ianqui. A la comissió investigadora s'hi va voler posar un element independent de prestigi i es va triar al premi Nobel de Física Richard P. Feynman que en el seu llibre de memòries "*What do you care what other people think?*" explica profusament l'experiència.

Un dels elements que més ressalten, i que més són criticats per Feynman, és l'ús de la probabilitat en qüestions d'enginyeria espacial. Un dels errors comesos per la NASA és utilitzar l'experiència històrica per determinar la fiabilitat d'una missió. Per justificar una probabilitat d'accident molt poc realista ($1/100.000$) els oficials de la NASA diuen (en documentació oficial prèvia a l'accident):

since the shuttle is a manned vehicle, the probability of mission success is necessarily very close to 1.0.

Historically, this extremely high degree of mission success has given rise to a difference in philosophy between manned space flight programs and unmanned programs; i.e., numerical probability usage versus engineering judgement.

Una mostra del poc sentit que poden tenir les estimacions probabilístiques la tenim en la disparitat de valors que donaven com a probabilitat P d'accident fatal diferents grups:

Enginyers de Rocketdyne (els fabricants): $P = 1/10.000$.

Enginyers del Marshall Space Center (els dissenyadors): $P = 1/300$.

El *management* de la NASA: $P = 1/100.000$.

Un enginyer independent consultat per la NASA: P entre $1/100$ i $2/100$.

Finalment Feynman, en les seves conclusions recomana

For a successful technology, reality must take precedence over public relations, for Nature cannot be fooled.

2. En el concurs televisiu del 123, al final el presentador ens ofereix triar entre tres sobres un dels quals conté un apartament a la costa. Després que n'hem seleccionat un, el presentador agafa un dels sobres de la taula i, somrient, mostra que no conté res. Ara ens ofereix la possibilitat de canviar el sobre que tenim pel que queda a la taula. Que hem de fer?
3. Apostant a la ruleta a vermell i negre, una estratègia és començar apostant una unitat. Si guanyem tornem a començar, si perdem apostem el doble que la passada vegada. D'aquesta manera sempre que guanyem recuperem tot el que hem perdut abans més una unitat. Discutiu la validesa d'aquesta estratègia. (En els casinos està prohibit utilitzar-la.)
4. Tenim el següent desenvolupament de Taylor ($|x| < 1$):

$$\frac{1}{(1-x)^n} = \sum_{k=0}^{\infty} C R_n^k x^k.$$

Demostreu-lo des del punt de vista combinatori. Després calculeu els coeficients de la sèrie i haureu demostrat novament quin valor pren $C R_n^k$.

5. L'esdeveniment impossible \emptyset té probabilitat 0. Si un esdeveniment A té probabilitat 0, es impossible que es produeixi?

6. Donada una variable aleatòria X contínua sempre es verifica $P(X=a) = 0$ (la probabilitat d'un punt es zero). Vol dir això que la variable no pren cap valor? No és una paradoxa?
7. Donada una variable aleatòria contínua X considerem les noves variables $Y = |X|$ i $S = \text{sg}(X)$. Com ha de ser la funció de densitat de X per que Y i S siguin independents?
8. El conjunt de Cantor \mathcal{K} és l'exemple més simple de fractal. Es parteix de l'interval $[0, 1] \subset \mathbb{R}$ i se li resta el terç del mig. Després es resta el terç del mig dels dos intervals que queden. Després es resta el terç del mig dels 4 intervals que queden... El conjunt $\mathcal{K} \subset \mathbb{R}$ és el que queda després d'aplicar aquest procediment indefinidament.
 - (a) Demostreu que \mathcal{K} és un borelià.
 - (b) Si X és una variable aleatòria uniforme en $[0, 1]$, que val $P(X \in \mathcal{K})$?
 - (c) Considereu ara una variable Y uniforme en \mathcal{K} . Demostreu que Y no és ni discreta, ni contínua, ni mixta! La seva funció de distribució es diu *escala del diable*. Quin aspecte té?
9. La “qualitat” d'un joc d'apostes es pot mesurar amb l'esperança del guany (per Euro apostat, per exemple). Si apostem a la tirada d'una moneda i, en cas d'encertar ens paguen el doble del que hem apostat, el guany esperat (apostant 1 euro) val $E[G] = 0 \cdot \frac{1}{2} + 2 \cdot \frac{1}{2} = 1$. En aquest cas diem que el joc és just. Apostant a vermell o negre a la ruleta (18 nombres vermells, 18 nombres negres i el zero), $E[G] = 0 \cdot \frac{19}{37} + 2 \cdot \frac{18}{37} = 0,973$. El joc és quasi just, amb un petit avantatge per a la banca.

Calculeu el guany esperat pel sorteig de la Loteria de Nadal a partir de les següents dades (Nadal 2008).

El sorteo consta de 195 series de 85.000 billetes cada una. Al precio de 200 euros el billete, dividido en décimos de 20 euros. El total de la emisión asciende a 3.315.000.000 euros. El total destinado a premios asciende a poco más de 2.000 millones de euros.

PREMIOS POR SERIE	TOTAL PREMIOS	PROPORCIÓN AL PREMIO
EL GORDO (Primer Premio)	1	15.000 Euros por Euro jugado
Segundo Premio	1	5.000 Euros por Euro jugado
Tercer Premio	1	2.500 Euros por Euro jugado
Cuartos Premios	2	1.000 Euros por Euro jugado
Quintos Premios	8	250 Euros por Euro jugado
Pedreas de 5 Cifras	1.774	5 Euros por Euro jugado
Aproximación Primer premio	2	100 Euros por Euro jugado
Aproximación Segundo premio	2	62,5 Euros por Euro jugado
Aproximación Tercer premio	2	48 Euros por Euro jugado
Centenas del Primer premio	99	5 Euros por Euro jugado
Centenas del Segundo premio	99	5 Euros por Euro jugado
Centenas del Tercer premio	99	5 Euros por Euro jugado
Centenas de dos Cuartos Prem	198	5 Euros por Euro jugado
Dos últimas del Primer premio	849	6 Euros por Euro jugado
Dos últimas del Segundo prem.	849	5 Euros por Euro jugado
Dos últimas del Tercer premio	849	5 Euros por Euro jugado
Reintegros (última cifra Gordo)	8.499	1 Euro por Euro jugado

10. De les consideracions mètriques fetes hem arribat a definir la distància entre dues variables aleatòries X, Y com

$$d(X, Y) = E[(X - Y)^2].$$

Sembla també que si aquesta distancia mesura la diferència entre les dues variables, la podríem definir

$$d(X, Y) = \int_{-\infty}^{\infty} (f_X(t) - f_Y(t))^2 dt$$

on f_X, f_Y són les respectives funcions de densitat.

Quina diferència hi ha entre les dues definicions?

SOLUCIONS DELS PROBLEMES

1. $\frac{1}{10^{n/2}}$ si n és parell i $\frac{1}{10^{(n-1)/2}}$ si n és senar.

2. 0,0154.

3. Algun as en 4 tirades: 0,5177. Algun doble as en 24 tirades: 0,4914.

4.

k	0	1	2	3	4	5-c	5+c	6
$P(k)$	0,4359	0,4130	0,1323	0,01765	0,0009686	$1,802 \cdot 10^{-5}$	$4,290 \cdot 10^{-7}$	$7,15 \cdot 10^{-8}$

En 10.950 apostes:

k	4	5-c	5+c	6
	0,999975	0,1790	0,00468	0,000782

5.

Parella	Doble parella	Trio
0,422	0,04753	0,0211
Escala	Color	Full
0,00392	0,0019	0,00144
Poker	Escala de color	Escala real
0,00024	0,0000138	$1,5 \cdot 10^{-6}$

6. 0,00119.

7.

8. $\binom{k-1}{k-n}$.

9. 1568. $\frac{N!^2}{(N-n)!^2 n!}$.

10. (a) $\frac{\binom{n}{k}^2}{\binom{2n}{2k}}$. (b) $\binom{2k}{k} \frac{1}{2^{2k}}$. El primer tendeix al resultat (b) que és independent de n .

11. (a) 0,6583. (b) 0,683. (c) No són independents.

12. (c) Algun ha de tenir probabilitat nul·la.

13. Només poden tenir cardinal 2 i 3 amb un element comú o 3 i 4 amb dos elements comuns. Per exemple, $\{1, 2\}$ i $\{1, 3, 4\}$ o $\{1, 2, 3\}$ i $\{1, 2, 4, 5\}$.

14. (a) $\frac{26}{31} = 0,838$. (b) 0,9. Efectivament, la condició ha disminuït la probabilitat de passar boles negres.

15. $\frac{6}{11}$.

16. Primer: $1/(1+q+q^2)$. Segon: $q/(1+q+q^2)$. Tercer: $q^2/(1+q+q^2)$. ($q = 1-p$)

17. Empat: $\frac{1}{11}$. A: $\frac{5}{11}$. B: $\frac{5}{11}$.

18. (a) $p_{aA} = \frac{p_A}{p_A+p_B-p_A p_B}$. (b) $p_{aA} = \frac{p_A(1-p_{bc}p_{cb})}{1-p_{ab}p_{bc}p_{ca}-p_{ac}p_{cb}p_{ba}-p_{ab}p_{ba}-p_{ac}p_{ca}-p_{cb}p_{bc}}$.

19. El primer $\frac{2}{7}$, els altres $\frac{1}{7}$.

20. $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) - P(A \cap C) + P(A \cap B \cap C)$.

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = \sum_i P(A_i) - \sum_{i < j} P(A_i \cap A_j) + \sum_{i < j < k} P(A_i \cap A_j \cap A_k) - \dots + (-1)^{n+1} P(A_1 \cap A_2 \cap \dots \cap A_n)$$

21. $P = 1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \dots + \frac{(-1)^n}{n!}$. El límit val $e^{-1} = 0,3678\dots$. Per $n = 10$ val 0,36787946, per $n = 100.000$ val 0,36787944.
22. 0,2718. En general $P = \sum_{k=0}^5 (-1)^k \binom{6}{k} (1 - \frac{k}{6})^n$.
23. 0,861657.
- 24.
25. 0,2937.
26. Els dos casos són equivalents: $P(\text{error}) = 3\epsilon^2 - 2\epsilon^3$. Si ϵ és petit, $P(\text{error}) < \epsilon$.
27. $\Omega_X = \{n, n+1, \dots\}$ i $P(k) = \binom{k-1}{n-1} p^n q^{k-n}$, $k = n, n+1, \dots$
28. Ha de ser $0 < a < 1$. $A = 1 - a$. (a) a^5 . (b) $\frac{1}{1+a}$. (c) $\frac{1}{1+a^2+a^4}$.
29. Si X té paràmetre α , $\frac{1}{X}$ té paràmetre α^{-1} .
30. (a) $\Omega_Y = [-1, 1]$. $f_Y(y) = \frac{1}{\pi\sqrt{1-y^2}}$, $-1 < y < 1$. $F_Y(y) = \frac{1}{2} + \frac{1}{\pi} \arcsin y$, $-1 \leq y \leq 1$.
(b) $\Omega_Y = [0, 1]$.
- $$F_Z(z) = \begin{cases} 0 & z < 0 \\ 1/2 & z = 0 \\ \frac{1}{2} + \frac{1}{\pi} \arcsin z & 0 \leq z \leq 1 \\ 1 & z \geq 1 \end{cases}$$
- (c) $E[Y] = 0$, $E[Z] = \frac{1}{\pi}$.
31. Primer cas: $E[X] = \frac{n}{p}$. Segon cas: $E[X] = \frac{1}{1-a}$.
32. 71% (fent la suposició raonable que la variable és gaussiana).
33. $K = \frac{\lambda^3}{2}$, $m_n = \frac{(n+2)!}{2\lambda^n}$, $E[X] = \frac{3}{\lambda}$, $V[X] = \frac{3}{\lambda^2}$.
34. (a) $\frac{e^{-\pi/2} - e^{-5\pi/2}}{1 - e^{-6\pi}}$. (b) $\Omega_Y = \{0, 1, 2, \dots\}$, $P_Y(n) = (1 - e^{-3})e^{-3n}$, $E[Y] = \frac{1}{e^3 - 1}$. (c) $\frac{2}{3}$.
35. $K = \frac{\lambda}{2}$, $E[X] = m$, $V[X] = \frac{2}{\lambda^2}$, $P = 0,8488$.
36. (a) $F_X(x) = \begin{cases} 0 & \text{si } x < 30 \\ \frac{(x-30)^2}{200} & \text{si } 30 \leq x < 40 \\ 1 - \frac{(50-x)^2}{200} & \text{si } 40 \leq x < 50 \\ 1 & \text{si } x \geq 50 \end{cases}$ 0,555 (b) $m_X = 40$, $\sigma_X = 4,082$. (c) T és geomètrica amb $p = \frac{1}{8}$. N és binomial amb $n = 30$, $p = \frac{8}{25}$. $E[T] = 8$, $E[N] = 9,6$. (d) 0,3006, 0,09828.
37. (a) $\text{erf}(k/\sqrt{2})$, 68,3%, 95,4%. (b) $\sqrt{\frac{2}{\pi}}\sigma$. (c) $m_n = e^{n^2 \frac{\sigma^2}{2} + nm}$, $E[Y] = e^{\frac{\sigma^2}{2} + m}$, $V[Y] = e^{2m}(e^{2\sigma^2} - e^{\sigma^2})$.
38. $\overline{N(t)} = N_0 e^{-\frac{t}{\tau_m}}$. Per mostres macroscòpiques: $\sigma/\overline{N} < \overline{N}^{-1/2}$ que és negligible (per exemple, en un gram d'urani hi ha 10^{21} partícules).
39. $F_T(t) = 1 - \exp(-\alpha t - \frac{\beta}{\gamma}(e^{\gamma t} - 1))$. El cas exponencial és $\beta = 0$.

40. (a) $K = 2/(e - 1)$, X és exponencial amb $\lambda = 2$ i $f_Y(y) = (e - 1)^{-1}e^y, 0 < y < 1$. Són independents. (b) $K = 2$, $f_X(x) = 2(e^{-x} - e^{-2x}), x > 0$ i Y és exponencial amb $\lambda = 1$. No són independents.
41. Quadrat: $P = 0,8464$. Hexàgon: $P = 0,9097$.
42. Si Z és l'àrea: $E[Z] = \frac{\pi}{4}L_1L_2$, $f_Z(z) = -(\pi L_1L_2)^{-1} \ln(\frac{z}{\pi L_1L_2}), 0 < z < \pi L_1L_2$.
43. (a) 0,02531. (b) 0,6943.
44. $\Omega_Y = [0, 2]$ i $f_Y(y) = \begin{cases} y & 0 \leq y \leq 1 \\ 2 - y & 1 \leq y \leq 2 \end{cases}$
45. $E[\Delta_A] = P(A)$, $\Delta_{A \cap B} = \Delta_A \Delta_B$, $\Delta_{A \cup B} = \Delta_A + \Delta_B - \Delta_A \Delta_B$, $\Delta_{\bar{A}} = 1 - \Delta_A$.
46. (a) Si N_k és la variable que dona el nombre d'ordinadors amb k usuaris, tenim que $E[N_k] = 20 \binom{15}{k} \frac{19^{15-k}}{20^{15}}$. (b) En particular $E[N_0] = 9,2$. (c) La suma d'aquestes esperances dona el nombre total d'ordinadors, 20.
47. $C[X, Y] = 1$. Això implica que no són independents.
48. T és exponencial amb $\lambda = \sum_i \lambda_i$. $E[T] = \frac{1}{\sum_i \lambda_i}$.
49. (a) 0,4375. (b) $f(z) = (60 - z)/1800, 0 < z < 60$. $E[Z] = 20$ min. (c) $f(u) = (60 - u)/1800, 0 < u < 60$. $E[U] = 20$ min.
50. (a) 5,1 Euros. (b) 60,5, 0,057. (c) 0,9755.
51. $f(t_1, t_2, \dots, t_n) = \lambda^n e^{-\lambda t_n}, 0 < t_1 < t_2 < \dots < t_n < \infty$.
52. $\frac{\sigma}{\sqrt{n}}$.
53. Uniforme en $[0, 1]$.
- 54.
55. (a) En els dos casos $\frac{k(n+1)}{2}$. (b) $\frac{m^k - (m-1)^k}{n^k}$.
56. (a) $y = 4\sqrt{1 - x^2}/\pi$, $x = 4\sqrt{1 - y^2}/\pi$. (b) $m_X = m_Y = \frac{4}{3\pi}$, $\sigma_X^2 = \sigma_Y^2 = \frac{1}{4} - \frac{16}{9\pi^2}$ ($\sigma_X = \sigma_Y = 0,264$), $C[X, Y] = \frac{1}{2\pi} - \frac{16}{9\pi^2}$ ($\rho = -0,3$). (c) $y = \rho x + m_X(1 - \rho)$, $x = \rho y + m_X(1 - \rho)$.
57. Additiu: $\lambda = \frac{4}{5}$. Multiplicatiu: $\lambda = 0$. Aquest segon cas es podria millorar fent una estimació de $|X|$ a partir de $|ZX|$. Només ens caldria conèixer $E[|Z|]$.
58. $m(t) = 1 + t \ln t + (1 - t) \ln(1 - t)$.
59. (b) $m(t) = 6 \frac{1 - e^{-t}}{t} - (3t + 6)e^{-t}$. (e) 144.
60. $C = \alpha X(t - T) + \beta X(t - 2T)$, $\alpha = e^{-aT}$, $\beta = 0$. $\bar{\epsilon} = K(1 - e^{-2aT})$.
61. (a) $\alpha = \frac{C(t-T, t)}{C(t-T, t-T)}$, $\beta = m(t) - m(t-T)\alpha$. (b) $\alpha = \frac{C(T)}{C(0)}$, $\beta = m(1 - \frac{C(T)}{C(0)})\alpha$. (c) $\alpha = 1$, $\beta = \lambda T$.
62. Uniforme en $[t_1, t_2]$.
63. (a) $m_{r,s} = \frac{2}{(r+s+2)(s+1)}$. (b) $x = \frac{1+y}{2}$. (c) $m(t) = \frac{2t^2}{3} - \frac{t}{4}$. $R(t_1, t_2) = t_1 t_2 (\frac{t_1 t_2}{2} - \frac{(t_1 + t_2)}{5} + \frac{1}{9})$. (d) $\lambda = \frac{40}{81}$.
64. (a) $m(t) = \frac{1 - e^{-t}}{t}$, $R(t_1, t_2) = \frac{1 - e^{-(t_1 + t_2)}}{t_1 + t_2}$. (b) 0. (c) 0 en MQ.
65. (a) Els criteris no decideixen res. (b) És ergòdic ja que $\sigma_T^2 = V_0^2 \frac{1 - \cos 2\omega T}{2\omega^2 T^2}$ tendeix a zero per $T \rightarrow \infty$.