SECCIÓ DOCENT I PROBLEMES

La «Secció docent i problemes» té l'objectiu de publicar articles de caire docent, difícilment publicables en revistes de recerca. A cada número de *Qüestiió* s'inclouen d'un a tres problemes i les solucions es donen en el número següent.

Els lectors poden proposar problemes amb les solucions pertinents i enviar-los a *Qüestiió*, que farà una selecció i en publicarà els més adequats, fent la corresponent referència a l'autor.

També seran ben rebudes solucions alternatives a les propostes fetes per l'autor dels problemes. L'editorial es reservarà, però, el dret a publicar-les.

QÜESTIIÓ, vol. 23, 1, p. 171-192, 1999

MODELS GRÀFICS D'INDEPENDÈNCIA

J.M. DURAN RÚBIES
Universitat Autònoma de Barcelona*

Els models gràfics d'independència són una eina de l'anàlisi multivariant que utilitza gràfics per representar models. En particular, els grafs d'independència resumeixen i clarifiquen les interaccions entre variables, interaccions no sempre fàcils d'interpretar, especialment quan hi intervenen tres o més variables.

En aquest treball es dóna, en clau pedagògica, una introducció a la teoria de grafs d'independència, començant per les nocions d'independència necessàries i arribant a les propietats de Markov, claus per a la interpretació d'aquests grafs.

S'estudia també un exemple d'aplicació amb variables numèriques, que mostra la viabilitat d'aquesta teoria a l'hora de simplificar i aclarir les relacions entre variables.

Independence graphs models.

Paraules clau: Independència, independència condicional, paradoxa de Simpson, vectors aleatoris, teoria de grafs, grafs d'independència condicional, propietats de Markov.

Classificació AMS: 62 H 20

* J.M. Duran Rúbies. Professor associat U.A.B. Departament de Matemàtiques. 08193 Bellaterra (Espanya). E-mail: duran@mat.uab.es

– Rebut el novembre de 1996.
– Acceptat el desembre de 1998.

1. INTRODUCCIÓ

En anàlisi multivariant, quan hi intervenen tres o més variables, les relacions d'independència condicional no són fàcils de veure ni d'interpretar. Els grafs d'independència condicional resumeixen i clarifiquen aquestes interaccions.

La teoria de grafs d'independència és relativament recent. Fou Speed [8] un dels primers en relacionar els mètodes de la teoria de grafs amb les interaccions entre variables. El treball clau, però, és el de Darroch, Lauritzen i Speed [1], seguit després per altres com Wermuth i Lauritzen [9], Edwards i Kreiner [3] i Whittaker [10], el qual ha estat el motivador principal del treball que segueix, recull de tot aquest material, presentat en clau pedagògica. Més recentment Edwards [2] ha publicat un llibre amb aplicacions dels models gràfics d'independència. Inclou una guia de referència del programa MIM per a PC dissenyat per estudiar aquests models.

L'objectiu d'aquest treball és l'enunciat, demostració i aplicació de la propietat global de Markov, que permet interpretar els grafs d'independència i està dirigit a tots aquells que, sense haver d'anar a llibres com Whittaker [10] i Edwards [2], desitgin tenir una primera idea dels models gràfics.

L'exposició està estructurada de la manera següent: en les seccions 2 a 4 es dóna una revisió ràpida de les nocions d'independència i d'independència condicional, fent especial èmfasi en les propietats que s'utilitzaran en la secció 7 per demostrar el teorema de separació. En la secció 5 definim els conceptes de la teoria de grafs necessaris per definir en la secció 6 els grafs d'independència condicional. En la secció 7 es presenta el teorema de separació i en la secció 8 les propietats de Markov. Finalment s'exposa, en la secció 9, un exemple d'aplicació amb variables numèriques.

La majoria de les demostracions estan agrupades fora del text principal, en l'apèndix, per poder seguir així més fàcilment els resultats que són el fil conductor d'aquest treball.

2. CONCEPTES D'INDEPENDÈNCIA

Dos esdeveniments A i B, d'un espai de probabilitat (Ω,$\mathcal{A}$,P), són ***independents*** i ho escriurem $A \perp\!\!\!\perp B$ si:

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

Tenint en compte que, suposant $P(B) > 0$, la probabilitat de l'esdeveniment A condicionat per l'esdeveniment B és:

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

una altra manera de definir la independència és:

$$A \perp\!\!\!\perp B \iff P(A) = P(A/B).$$

Es pot provar que:

- $A \perp\!\!\!\perp B \iff A \perp\!\!\!\perp B^c$. (vegeu pàg. 186).
- $A \perp\!\!\!\perp B \iff A^c \perp\!\!\!\perp B$, i
- $A \perp\!\!\!\perp B \iff A^c \perp\!\!\!\perp B^c$.

També es pot provar que la relació $\perp\!\!\!\perp$ definida en un mateix espai de probabilitat és simètrica, però no és reflexiva ni transitiva. (vegeu pàg. 186).

En el cas de tenir tres esdeveniments o més, podem definir la ***independència marginal*** com la independència dos a dos. El concepte d'independència però té dues formes de generalització més naturals.

Direm que els esdeveniments A, B i C són ***independents*** [1] si i només si compleixen:

- Són marginalment independents
- $P(A \cap B \cap C) = P(A)P(B)P(C)$.

Cal fer notar la necessitat de les dues condicions ja que ni la primera implica la segona, ni viceversa.

Exemple. Per veure que $P(A \cap B \cap C) = P(A)P(B)P(C)$ no implica que A, B i C siguin independents dos a dos, observem el següent exemple:

C	A	A^c
B	$1/16$	$1/16$
B^c	$1/16$	$1/16$

C^c	A	A^c
B	$1/16$	$5/16$
B^c	$5/16$	$1/16$

Per una banda tenim que $P(A \cap B \cap C) = 1/16$ i en canvi A i B no són independents, ja que $P(A) = P(B) = 1/2$ i $P(A \cap B) = 1/8$.

[1] Alguns autors prefereixen parlar de esdeveniments ***mútuament independents***.

Per veure que A, B i C independents dos a dos no implica $P(A \cap B \cap C) = P(A)P(B)$ $P(C)$, l'exemple següent és clarificador:

C	A	A^c
B	0	$1/4$
B^c	$1/4$	0

C^c	A	A^c
B	$1/4$	0
B^c	0	$1/4$

El punt clau d'aquesta última afirmació és de fet que:

$$A \perp\!\!\!\perp B, A \perp\!\!\!\perp C \not\Leftrightarrow A \perp\!\!\!\perp (B \cap C).$$

En els dos exemples anteriors es pot comprovar. En el primer exemple no es verifica la suficiència i en el segon no es verifica la necessitat.

Tenim, per tant, que la independència implica la independència marginal però no és certa l'afirmació contrària.

La segona possibilitat de generalització és la següent. Un esdeveniment A és ***independent respecte l'àlgebra generada pels*** esdeveniments B i C, i ho escriurem $A \perp\!\!\!\perp [B, C]$, si i només si:

$$A \perp\!\!\!\perp (B \cap C), A \perp\!\!\!\perp (B \cap C^c), A \perp\!\!\!\perp (B^c \cap C) \text{ i } A \perp\!\!\!\perp (B^c \cap C^c).$$

Amb aquesta definició tenim que:

- $A \perp\!\!\!\perp [B, C] \iff A \perp\!\!\!\perp B, A \perp\!\!\!\perp C$ i $A \perp\!\!\!\perp (B \cap C)$. (vegeu pàg. 187).
- $A \perp\!\!\!\perp [B, C] \iff A \perp\!\!\!\perp E,$ per tot esdeveniment, E, generat per unions, interseccions i pas al complementari dels esdeveniments B i C. (vegeu pàg. 187).

3. INDEPENDÈNCIA CONDICIONAL

Donarem dues definicions d'esdeveniments condicionalment independents.

Direm que dos esdeveniments A i B són ***dèbilment independents condicionats per*** C si i només si:

$$P(A \cap B/C) = P(A/C)P(B/C), \text{ on suposem que } P(C) > 0.$$

Escriurem aquest fet de la manera següent:

$$A \perp\!\!\!\perp B/C.$$

Aquesta relació entre esdeveniments és simètrica, $A \perp\!\!\!\perp B/C \iff B \perp\!\!\!\perp A/C$, però no és reflexiva ni transitiva.

Compleix també:

- $A \perp\!\!\!\perp B/C \iff A \perp\!\!\!\perp B^c/C$.
- $A \perp\!\!\!\perp B/C \iff A^c \perp\!\!\!\perp B/C$.
- $A \perp\!\!\!\perp B/C \iff A^c \perp\!\!\!\perp B^c/C$.

Atenció, però, al fet que $A \perp\!\!\!\perp B/C$ no implica $A \perp\!\!\!\perp B/C^c$, com es pot veure en l'exemple de la pàg. 173, ni tampoc, per simetria, $A \perp\!\!\!\perp B/C^c$ implica $A \perp\!\!\!\perp B/C$.

També hem de tenir en compte els següents resultats aparentment paradoxals:

- $A \perp\!\!\!\perp B/C$ i $A \perp\!\!\!\perp B/C^c \not\Rightarrow A \perp\!\!\!\perp B$.
- $A \perp\!\!\!\perp B \not\Rightarrow A \perp\!\!\!\perp B/C$ ó $A \perp\!\!\!\perp B/C^c$.

Aquests resultats queden il·lustrats per la paradoxa de Simpson.

Exemple. Dels molts exemples que s'han publicat escollim el del mateix E.H.Simpson [7]. Uns laboratoris volen provar un cert tractament. S'estudien 52 casos donant els resultats de la taula següent:

C	A	A^c
B	$4/52$	$8/52$
B^c	$3/52$	$5/52$

C^c	A	A^c
B	$2/52$	$12/52$
B^c	$3/52$	$15/52$

on C i C^c representen home i dona respectivament. A i A^c ens diu si ha rebut tractament o no i B i B^c si s'ha curat o ha mort.

A la vista d'aquestes dades veiem que el nou tractament és beneficiós tant per la població masculina com per la població femenina, mentre que si observéssim els resultats globals, sense tenir en compte el sexe, veuríem que és independent, pel desenvolupament de la malaltia, el fet d'aplicar o no aquest nou tractament.

La segona definició és una versió forta de l'anterior i està relacionada amb la definició d'independència d'A respecte l'àlgebra generada per B i C.

Dos esdeveniments A i B són ***fortament independents condicionats per l'esdeveniment*** C, i ho escriurem $A \perp\!\!\!\perp B/[C]$, si i només si:

$$A \perp\!\!\!\perp B/C \text{ i } A \perp\!\!\!\perp B/C^c.$$

La definició $A \perp\!\!\!\perp B/[C]$ és més forta que $A \perp\!\!\!\perp B/C$ en el sentit de que la primera implica la segona però no al contrari.

Més en general, dos esdeveniments A i B són ***fortament independents condicionats pels esdeveniments*** C i D, i ho escriurem $A \perp\!\!\!\perp B/[C, D]$, si i només si:

$$A \perp\!\!\!\perp B/(C \cap D),\ A \perp\!\!\!\perp B/(C \cap D^c),\ A \perp\!\!\!\perp B/(C^c \cap D) \text{ i } A \perp\!\!\!\perp B/(C^c \cap D^c).$$

Per poder arribar a demostrar el teorema de separació, secció 7, necessitem el següent resultat anomenat ***independència de blocs***, que ens relaciona la definició forta d'independència condicionada i la definició d'independència entre un esdeveniment i una àlgebra:

Sigui l'espai de probabilitat (Ω,$\mathcal{A}$,P). Suposant que P és estrictament positiva en la partició generada pels esdeveniments A, B i C, les propietats següents són equivalents:

1. $A \perp\!\!\!\perp [B, C]$.
2. $A \perp\!\!\!\perp [B]/[C]$ i $A \perp\!\!\!\perp [C]/[B]$. (vegeu pàg. 188).

4. EXTENSIÓ A VECTORS ALEATORIS

Direm que els vectors aleatoris X i Y són ***independents***, $X \perp\!\!\!\perp Y$, si i només si la funció de densitat conjunta, f_{XY}, compleix:

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

O també:

$$X \perp\!\!\!\perp Y \iff f_{X/Y}(x;y) = f_X(x),\ \forall x.$$

Recordem que $f_{X/Y}(x;y) = \dfrac{f_{XY}(x,y)}{f_Y(y)}$ i que per tant hem de suposar que $f_Y(y) \neq 0$ quasi-per-tot.

Dos resultats importants per poder demostrar més endavant el teorema de separació són els següents:

1. Criteri de factorització.

   Els vectors aleatoris X i Y són independents si i només si existeixen dues funcions g i h tals que:

   $$f_{XY}(x,y) = g(x)h(y), \forall x \text{ i } \forall y. \quad \text{(vegeu pàg. 188).}$$

2. Lema de reducció.

   Si (X,Y,Z) és un vector aleatori, aleshores:

   $$X \perp\!\!\!\perp (Y,Z) \Rightarrow X \perp\!\!\!\perp Y \text{ i } X \perp\!\!\!\perp Z,$$

   és a dir, la independència implica la independència marginal. (vegeu pàg. 188).

La implicació contrària no es verifica.

Vegem-ho en el següent exemple [2]:

Exemple. Sigui el cub $\{(x,y,z); 0 < x < 1, 0 < y < 1, 0 < z < 1\}$ i la funció de densitat conjunta:

$$f_{XYZ}(x,y,z) = 2(\chi_1(x,y,z) + \chi_2(x,y,z) + \chi_3(x,y,z) + \chi_4(x,y,z))$$

on

$$\chi_1(x,y,z) = \begin{cases} 1 & \text{si } x < 0.5, y < 0.5 \text{ i } z < 0.5 \\ 0 & \text{en cas contrari} \end{cases} \quad ;$$

[2] Aquest exemple així com d'altres de semblants tenen el seu origen en l'exemple de S. Bernstein, vegeu per exemple [4] pg. 126. Es pot trobar un exemple en el cas continu quelcom diferent en [5].

$$\chi_2(x,y,z) = \begin{cases} 1 & \text{si } x < 0.5, y > 0.5 \text{ i } z > 0.5 \\ 0 & \text{en cas contrari} \end{cases} \quad ;$$

$$\chi_3(x,y,z) = \begin{cases} 1 & \text{si } x > 0.5, y < 0.5 \text{ i } z > 0.5 \\ 0 & \text{en cas contrari} \end{cases} \quad \text{i}$$

$$\chi_4(x,y,z) = \begin{cases} 1 & \text{si } x > 0.5, y > 0.5 \text{ i } z < 0.5 \\ 0 & \text{en cas contrari} \end{cases}$$

són les funcions indicadores de cubs d'aresta 1/2 situats dins del cub gran d'aresta unitat. En aquest cas tenim per $0 < x < 1$ i $0 < y < 1$ que $f_{XY}(x,y) = 1$ que és igual al producte $f_X(x)f_Y(y)$. Per tant $X \perp\!\!\!\perp Y$.

De la mateixa manera $X \perp\!\!\!\perp Z$. Però $f_X(x)f_{YZ}(y,z) = 1$ que no és igual a $f_{XYZ}(x,y,z)$.

El criteri de factorització dóna una caracterització de la independència i el lema de reducció permet escriure-la amb marginals.

Definim ara la independència condicional.

Els vectors aleatoris Y i Z són ***independents condicionats pel*** vector aleatori X, $Y \perp\!\!\!\perp Z/X$, si i només si la densitat conjunta dels vectors Y i Z condicionada a $X = x$ és igual al producte de les seves marginals:

$$f_{YZ/X}(y,z;x) = f_{Y/X}(y;x)f_{Z/X}(z;x), \ \forall z, y \text{ i } \forall x \text{ tal que } f_X(x) > 0.$$

O equivalentment:

$$\begin{aligned} f_{Y/XZ}(y;x,z) &= f_{Y/X}(y;x), \text{ o} \\ f_{XYZ}(x,y,z) &= f_{XY}(x,y)\frac{f_{XZ}(x,z)}{f_X(x)}. \end{aligned}$$

De fet és una extensió de la definició forta d'independència condicional d'esdeveniments, $A \perp\!\!\!\perp B/[C]$, ja que és per tot x.

Generalitzem també el criteri de factorització, i els lemes de reducció i independència de blocs.

- Criteri de factorització.

  $Y \perp\!\!\!\perp Z/X \iff \exists g, h$ funcions tals que:
  $f_{XYZ}(x,y,z) = g(x,y)h(x,z), \forall y, z$ i $\forall x$ amb $f_X(x) > 0$. (vegeu pàg. 189).

- Lema de reducció.

  Siguin X, Y, Z_1, Z_2 vectors aleatoris, aleshores:

$$Y \perp\!\!\!\perp (Z_1, Z_2)/X \Rightarrow Y \perp\!\!\!\perp Z_1/X \text{ i } Y \perp\!\!\!\perp Z_2/X.$$ (vegeu pàg. 189).

- Independència de blocs.

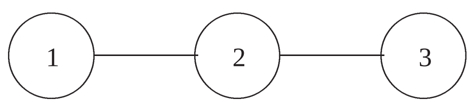
  Siguin X, Y, Z_1, Z_2 vectors aleatoris i $f > 0$, aleshores les afirmacions següents són equivalents:

  1. $Y \perp\!\!\!\perp (Z_1, Z_2)/X$.
  2. $Y \perp\!\!\!\perp Z_1/(X, Z_2)$, i $Y \perp\!\!\!\perp Z_2/(X, Z_1)$.
  3. $Y \perp\!\!\!\perp Z_2/(X, Z_1)$, i $Y \perp\!\!\!\perp Z_1/X$. (vegeu pàg. 189).

5. TEORIA DE GRAFS

Un ***graf*** és un parell (V, F) on V és un conjunt no buit de ***vèrtexs*** i F és una col·lecció de parells no ordenats de punts diferents de V, ***arestes***.

Per exemple, el graf amb $V = \{1, 2, 3\}$ i $F = \{(1, 2), (2, 3)\}$ es pot representar com:



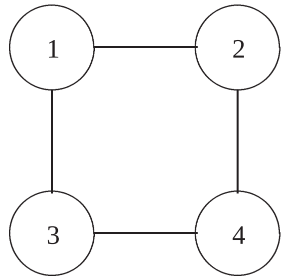
Un parell de vèrtexs, i, j, són ***contigus*** si $(i, j) \in F$.

Un ***camí*** és una seqüència de vèrtexs contigus. Un camí és un ***cicle*** si el primer i l'últim vèrtex coincideixen.

Dos vèrtexs i i j estan ***connectats*** si existeix un camí amb extrems i i j.

Un conjunt de vèrtexs S ***separa*** dos vèrtexs i, j si tot camí amb origen en i i final en j conté al menys un vèrtex de S.

Per exemple, en el graf:



el conjunt $S = \{1,4\}$ separa els vèrtexs 3 i 2.

Un conjunt $S \subset V$ de vèrtexs ***separa*** dos conjunts V_1 i V_2 de vèrtexs si V separa tot parell de vèrtexs, un de V_1 i l'altre de V_2.

Frontera d'un conjunt de vèrtexs V és el conjunt de tots els vèrtexs contigus a un vèrtex de V i que no són de V. Per exemple, en el graf anterior el conjunt $\{3,2\}$ és la frontera del conjunt $V = \{1\}$.

6. GRAFS D'INDEPENDÈNCIA CONDICIONAL

Sigui $X = (X_1, \ldots, X_k)$ un vector aleatori i sigui $V = \{1, \ldots, k\}$ el corresponent conjunt de vèrtex.

Un graf (V, F) direm que és un ***graf d'independència condicional*** si

$$X_i \perp\!\!\!\perp X_j / X_{V-\{i,j\}} \iff (i,j) \notin F.$$

Exemple. Sigui $X = (X_1, X_2, X_3, X_4)$ amb les següents independències:

$$X_1 \perp\!\!\!\perp X_3/\{X_2, X_4\}, X_1 \perp\!\!\!\perp X_4/\{X_2, X_3\} \text{ i } X_2 \perp\!\!\!\perp X_4/\{X_1, X_3\}.$$

El graf d'independència condicional vindrà donat per $F = \{(1,2), (2,3), (3,4)\}$.

Gràficament:

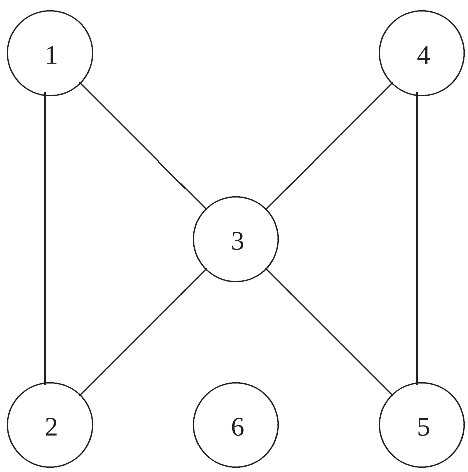
7. TEOREMA DE SEPARACIÓ

Enunciem en primer lloc el teorema.

Sigui $X = (X_1, \dots, X_k)$ un vector aleatori i sigui $V = \{1, \dots, k\}$ el corresponent conjunt de vèrtexs. Siguin a, b i c subconjunts disjunts de V. En aquestes condicions, si a separa b i c en el graf d'independència condicional (V, F), aleshores $X_b \perp\!\!\!\perp X_c / X_a$.

La demostració d'aquest teorema es basa en l'aplicació dels lemes de reducció i independència de blocs. Per donar una idea del seu mecanisme vegem com aniria en un cas senzill.

Sigui el següent graf d'independència condicional:



on $a = \{3\}$, $b = \{1, 2\}$ i $c = \{4, 5\}$. És clar que a separa b i c.

Per definició de graf d'independència tenim:

$$X_1 \perp\!\!\!\perp X_4/(X_2X_3X_5X_6) \text{ i } X_1 \perp\!\!\!\perp X_6/(X_2X_3X_5X_4).$$

Ara pel lema d'independència de blocs amb $Y = X_1$, $Z_1 = X_4$, $Z_2 = X_6$ i $X = (X_2X_3X_5)$ tenim

$$X_1 \perp\!\!\!\perp (X_4X_6)/(X_2X_3X_5)$$

i pel teorema de reducció arribem a

$$X_1 \perp\!\!\!\perp X_4/(X_2X_3X_5).$$

De la mateixa manera podríem haver arribat a $X_1 \perp\!\!\!\perp X_5/(X_2X_3X_4)$.

D'aquestes dues últimes afirmacions i aplicant el lema d'independència de blocs trobem que $X_1 \perp\!\!\!\perp (X_4X_5)/(X_2X_3)$.

També hauríem pogut trobar de la mateixa manera que $X_2 \perp\!\!\!\perp (X_4X_5)/(X_1X_3)$. Tornant aplicar el lema d'independència de blocs tenim que:

$$(X_1X_2) \perp\!\!\!\perp (X_4X_5)/X_3$$

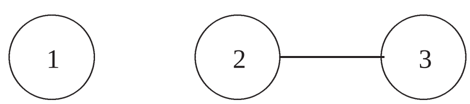
que acabaria la demostració.

Una demostració complerta d'aquest teorema es pot trobar a [10].

Una conseqüència del teorema de separació és que alguna de les variables condicionants en una relació d'independència pot ser redundant.

Direm que una independència condicional entre un parell de variables és ***minimal*** si no es pot aplicar més el teorema de separació per eliminar variables condicionants. En definitiva, quan hem eliminat totes les variables redundants en una relació d'independència.

Exemple. Sigui el graf d'independència condicional:



Totes les afirmacions d'aquest graf són: $X_1 \perp\!\!\!\perp X_2/X_3$ i $X_1 \perp\!\!\!\perp X_3/X_2$. En canvi per aplicació del teorema de separació tenim: $X_1 \perp\!\!\!\perp X_2$ i $X_1 \perp\!\!\!\perp X_3$ que són a la vegada independències condicionals minimals.

8. PROPIETATS DE MARKOV

Presentem a continuació tres propietats que són equivalents, com demostrarem tot seguit, i que, segons la definició donada anteriorment, un graf és d'independència condicional si té la primera d'aquestes propietats. Són les propietats de Markov:

1. Propietat ***dos a dos***: $\forall i, j$ no contigus, $X_i \perp\!\!\!\perp X_j/X_a$ on $a = V \perp \{i, j\}$.
2. Propietat ***global***: $\forall a, b, c$ subconjunts de V disjunts i b i c separats per a implica que $X_b \perp\!\!\!\perp X_c/X_a$.
3. Propietat ***local***: $\forall i$ i $a = frontera(i)$ i b els restants vèrtexs, tenim que $X_i \perp\!\!\!\perp X_b/X_a$.

- Les tres propietats de Markov són equivalents.

La demostració és com segueix:

- La propietat dos a dos implica la propietat global a causa del teorema de separació.
- La propietat global implica la propietat local pel fet que la frontera sempre és un conjunt separador.
- Falta demostrar que la propietat local implica la propietat dos a dos. Suposem un graf amb vèrtex $V = \{1 \ldots k\}$ que satisfà la propietat local. En aquestes condicions, per cada vèrtex i tenim $X_i \perp\!\!\!\perp X_{V-(\{i\}\cup a)}/X_a$ on a és la frontera del vèrtex i.

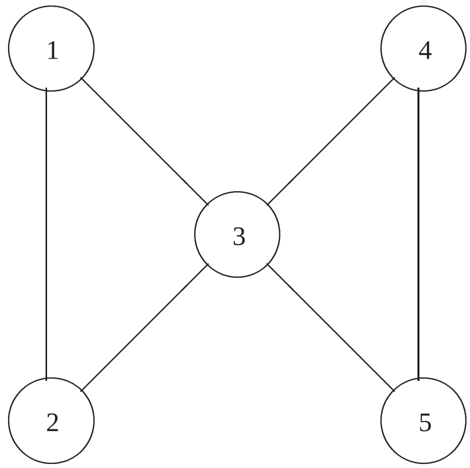
  Siguin ara un vèrtex $j \in V \perp (\{i\} \cup a)$ i el conjunt $c = V \perp (\{i, j\} \cup a)$.

  Podem escriure $X_i \perp\!\!\!\perp (X_j, X_c)/X_a$ i pel lema d'independència de blocs això equival a $X_i \perp\!\!\!\perp X_j/(X_a, X_c)$ i $X_i \perp\!\!\!\perp X_c/(X_a, X_j)$.

  La primera afirmació ja ens dóna la propietat dos a dos ja que $a \cup c = V \perp \{i, j\}$

Vegem en un exemple aquesta equivalència.

Exemple. Sigui el següent graf d'independència condicional:



- Per la propietat dos a dos tenim:

  $\perp$ $X_1 \perp\!\!\!\perp X_4 / (X_2, X_3, X_5)$
  $\perp$ $X_1 \perp\!\!\!\perp X_5 / (X_2, X_3, X_4)$
  $\perp$ $X_2 \perp\!\!\!\perp X_4 / (X_1, X_3, X_5)$
  $\perp$ $X_2 \perp\!\!\!\perp X_5 / (X_1, X_3, X_5)$.

- Per la propietat global tenim:

  $\perp$ $X_1 \perp\!\!\!\perp X_4 / X_3$
  $\perp$ $X_2 \perp\!\!\!\perp X_4 / X_3$
  $\perp$ $X_1 \perp\!\!\!\perp X_5 / X_3$
  $\perp$ $X_2 \perp\!\!\!\perp X_5 / X_3$ és a dir
  $\perp$ $(X_1, X_2) \perp\!\!\!\perp (X_4, X_5) / X_3$.

- Per la propietat local tenim:

  $\perp$ $X_1 \perp\!\!\!\perp (X_4 X_5) / (X_2 X_3)$
  $\perp$ $X_2 \perp\!\!\!\perp (X_4 X_5) / (X_1 X_3)$
  $\perp$ $X_5 \perp\!\!\!\perp (X_1 X_2) / (X_3 X_4)$
  $\perp$ $X_4 \perp\!\!\!\perp (X_1 X_2) / (X_3 X_5)$.

9. EXEMPLE D'APLICACIÓ

Per acabar vegem un exemple d'aplicació de la teoria de grafs d'independència condicional.

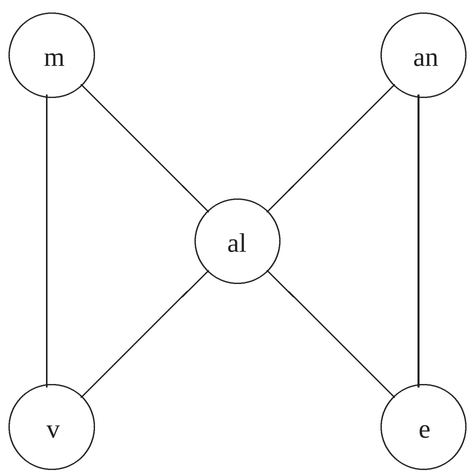
L'exemple està desenvolupat a [10] amb dades extretes de [6]. Les dades corresponen a les notes obtingudes per 88 alumnes en 5 assignatures: mecànica, vectors, àlgebra, anàlisi i estadística, respectivament.

La matriu de correlacions dóna:

$$\begin{pmatrix} 1 & & & & \\ 0.3371 & 1 & & & \\ 0.2294 & 0.2784 & 1 & & \\ \ldots & \ldots & 0.4336 & 1 & \\ \ldots & \ldots & 0.3554 & 0.2530 & 1 \end{pmatrix}$$

on els valors marcats amb punts suspensius són les correlacions simples no significatives, és a dir, aproximadament zero.

Una primera interpretació és simplement que cada parell de variables amb correlació parcial zero són independents condicionalment a les altres variables. Això ens porta a construir el següent graf d'independència condicional:



El gràfic és idèntic al de l'exemple de l'apartat 9 on ara, x_1 és mecànica, x_2 és vectors, x_3 és àlgebra, x_4 és anàlisi i x_5 és estadística.

Repassant l'exemple citat trobem que la propietat global de Markov ens permet escriure

$$\text{[mecànica, vectors]} \perp\!\!\!\perp \text{[anàlisi, estadística]/àlgebra}$$

la qual cosa ens porta a concloure que hi ha dos grups de variables, mecànica, vectors i àlgebra, per una banda, i anàlisi, estadística i àlgebra, per l'altra.

Pel que fa a la predicció de variables, per la propietat local de Markov, tenim:

- àlgebra i vectors són suficients per predir mecànica.
- mecànica i àlgebra són suficients per predir vectors.
- àlgebra i anàlisi són suficients per predir estadística.
- àlgebra i estadística són suficients per predir anàlisi.
- totes les variables són necessàries per predir àlgebra.

10. APÈNDIX

- $A \perp\!\!\!\perp B \iff A \perp\!\!\!\perp B^c$.
  secció 2, pàgina 173.

  $\perp$ Cal notar tan sols que:
  $P(A \cap B^c) = P[A \perp (A \cap B)] = P(A) \perp P(A)P(B) =$
  $P(A)[1 \perp P(B)] = P(A)P(B^c)$.

- *La relació* $\perp\!\!\!\perp$ *no és transitiva.*
  secció 2, pàgina 173.

  $\perp$ Tirem un dau. Siguin els esdeveniments següents:
  A: «parell», B: «múltiple de tres» i C: «Imparell».
  $A \perp\!\!\!\perp B$ i $B \perp\!\!\!\perp C$ però no és cert que $A \perp\!\!\!\perp C$.

- $A \perp\!\!\!\perp [B, C] \iff A \perp\!\!\!\perp B, A \perp\!\!\!\perp C$ *i* $A \perp\!\!\!\perp (B \cap C)$.

  secció 2, pàgina 174.

  Abans de provar-ho demostrem el següent resultat:

  $\perp$ $A \perp\!\!\!\perp B, A \perp\!\!\!\perp C$ i B i C disjunts $\Rightarrow A \perp\!\!\!\perp (B \cup C)$.

  Amb les hipòtesis establertes tenim: $P(A \cap (B \cup C)) = P((A \cap B) \cup (A \cap C)) = P(A \cap B) + P(A \cap C) \perp P(A \cap B \cap C) = P(A \cap B) + P(A \cap C) = P(A)P(B) + P(A)P(C) = P(A)(P(B) + P(C)) = P(A)P(B \cup C)$.

  Ara estem en condicions de demostrar el resultat.

  $\perp$ Per $\Rightarrow$ cal recordar que $B = (B \cap C^c) \cup (B \cap C)$, amb $B \cap C$i$B \cap C^c$ disjunts. Igual per C.

  Per veure $\Leftarrow$ hem de demostrar que

  $A \perp\!\!\!\perp (B \cap C), A \perp\!\!\!\perp (B \cap C^c), A \perp\!\!\!\perp (B^c \cap C)$ i $A \perp\!\!\!\perp (B^c \cap C^c)$

  Per exemple, $P(A)P(B^c \cap C) = P(A)(P(C) \perp P(B \cap C)) = P(A)P(C) \perp P(A)P(B \cap C) = P(A \cap C) \perp P(A \cap B \cap C) = P(A \cap B^c \cap C)$.

  O també, $P(A)P(B^c \cap C^c) = P(A)(P(B^c) \perp P(B^c \cap C)) = P(A)P(B^c) \perp P(A)P(B^c \cap C) = P(A \cap B^c) \perp P(A \cap B^c \cap C) = P(A \cap B^c \cap C^c)$.

  I així també $A \perp\!\!\!\perp (B \cap C^c)$.

- $A \perp\!\!\!\perp [B, C] \Rightarrow A \perp\!\!\!\perp E$, *per tot esdeveniment* E *generat per unions, interseccions i pas al complementari dels esdeveniments B i C.*

  secció 2, pàgina 174.

  De la definició i de la propietat anterior ens queda tan sols demostrar-ho en el cas $E = B \cup C, E = B \cup C^c, E = B^c \cap C$i $E = B^c \cap C^c$.

  $\perp$ $A \perp\!\!\!\perp (B \cup C)$.

  De $A \perp\!\!\!\perp B, A \perp\!\!\!\perp CiA \perp\!\!\!\perp (B \cap C)$ podem escriure:

  $P(A \cap (B \cup C)) = P((A \cap B) \cup (A \cap C)) = P(A \cap B) + PA \cap C) \perp P(A \cap B \cap C) = P(A)P(B) + P(A)P(C) \perp P(A)P(B \cap C) = P(A)P(B \cup C)$.

  $\perp$ $A \perp\!\!\!\perp (B^c \cup C)$.

  La mateixa idea però tenint en compte $A \perp\!\!\!\perp B^c, A \perp\!\!\!\perp CiA \perp\!\!\!\perp (B^c \cap C)$.

  $\perp$ Igual per les restants independències.

- *Lema d'independència de blocs.*

  secció 3, pàgina 176.

  $\perp$ (1) $\Rightarrow$ (2). Tenim: $P(A \cap B/C) = \dfrac{P(A \cap B \cap C)}{P(C)} = \dfrac{P(A)P(B \cap C)}{P(C)} = P(A)P(B/C) = P(A/C)P(B/C)$, és a dir $A \perp\!\!\!\perp B/C$.

  De la mateixa manera es demostraria $A \perp\!\!\!\perp B^c/C, A \perp\!\!\!\perp B/C^c$ i $A \perp\!\!\!\perp B^c/C^c$, amb el que tindríem $A \perp\!\!\!\perp [B]/[C]$. De manera semblant arribaríem a veure $A \perp\!\!\!\perp [C]/[B]$.

  $\perp$ (2) $\Rightarrow$ (1). De (2) tenim que d'una banda, $P(A \cap B \cap C) = P(A \cap B/C)P(C) = P(A/C)P(B/C)P(C) = P(A/C)P(B \cap C)$, i d'altra banda $P(A \cap B \cap C) = P(A \cap C/B)P(B) = P(A/B)P(C/B)P(B) = P(A/B)P(B \cap C)$.

  Així doncs $P(A/C) = P(A/B)$.

  De la mateixa manera i partint de $P(A \cap B^c \cap C)$ arribem a $P(A/C) = P(A/B^c)$, per tant $P(A/B^c) = P(A/B)$.

  Això ens permet veure que $P(A) = P(A/B)P(B) + P(A/B^c)P(B^c) = P(A/B)P(B) + +P(A/B)P(B^c) = P(A/B)[P(B) + P(B^c] = P(A/B)$ per tant $A \perp\!\!\!\perp B$.

  De la mateixa manera veuríem $A \perp\!\!\!\perp C$.

  Finalment, de $P(A \cap B/C) = P(A/C)P(B/C)$, en ser $A \perp\!\!\!\perp B/C$, tenim $P(A \cap B \cap C) = P(A/C)P(B \cap C) = P(A)P(B \cap C)$, el que demostra que $A \perp\!\!\!\perp (B \cap C)$.

- *Criteri de factorització.*

  secció 4, pàgina 177.

  $\perp$ La necessitat passa per fer $f_X = g$ i $f_Y = h$. Provem la suficiència.

  De $f_{XY}(x, y) = g(x)h(y)$, integrant els dos costats sobre y, tenim $f_X(x) = g(x)k$. De la mateixa manera, integrant sobre x: $f_Y(y) = h(y)k'$, d'on $f_{XY}(x, y) = k'' f_X(x) f_Y(y)$; però $k'' = 1$ en integrar sobre x i y.

- *Lema de reducció.*

  secció 4, pàgina 177.

  $\perp$ $X \perp\!\!\!\perp (Y, Z) \Rightarrow f_{XYZ}(x, y, z) = f_X(x) f_{YZ}(y, z)$. Integrant els dos costats sobre z tenim:

  $f_{XY}(x, y) = f_X(x) f_Y(y)$, per tant $X \perp\!\!\!\perp Y$.

  De la mateixa manera però integrant sobre y es demostraria $X \perp\!\!\!\perp Z$.

- *Criteri de factorització.*

  secció 4, pàgina 178.

  - $\perp$ Per veure la necessitat cal fer $g(x,y) = f_{XY}(x,y)$ i $h(x,z) = f_{XZ}(x,z)/f_X(x)$.
  - $\perp$ Per veure la suficiència integrem els dos costats de la igualtat $f_{XYZ}(x,y,z) = g(x,y)h(x,z)$ primer sobre y i després sobre z. Així tenim: $f_{XZ}(x,z) = h(x,z)k(x)$ i $f_{XY}(x,y)g(x,y)k'(x)$, per tant:

    $f_{XYZ}(x,y,z) = f_{XZ}(x,z)f_{XY}(x,y)k''(x)$, però integrant aquesta igualtat sobre y i z alhora tenim que $k''(x) = 1/f_X(x)$.

- *Lema de reducció.*

  secció 4, pàgina 179.

  - $\perp$ De $Y \perp\!\!\!\perp (Z_1, Z_2)/X$ tenim que $f_{XYZ_1Z_2}(x,y,z_1,z_2) = f_{XY}(x,y)f_{XZ_1Z_2}(x,z_1,z_2)/f_X(x)$. Ara integrant sobre z_1 tenim: $f_{XYZ_2}(x,y,z_2) = f_{XY}(x,y)f_{XZ_2}(x,z_1)/f_X(x)$, que implica $Y \perp\!\!\!\perp Z_2/X$. De la mateixa manera sortiria $Y \perp\!\!\!\perp Z_1/X$.

- *Lema d'independència de blocs.*

  secció 4, pàgina 179.

  - $\perp$ (1) $\Rightarrow$ (2). Pel criteri de factorització sabem que si $Y \perp\!\!\!\perp (Z_1, Z_2)/X$ aleshores tenim $f(x,y,z_1,z_2) = g(x,y)h(x,z_1,z_2)$.

    Cal observar tan sols ara que a la dreta de l'expressió no hi ha cap funció que tingui a la y i a les (z_1, z_2) com arguments. En particular no hi ha cap funció que tingui com arguments y i z_1 a la vegada. Aplicant un altre cop el criteri de factorització deduïm que $Y \perp\!\!\!\perp Z_1/(X, Z_2)$. De la mateixa manera demostraríem que $Y \perp\!\!\!\perp Z_2/(X, Z_1)$.
  - $\perp$ (2) $\Rightarrow$ (1). De $Y \perp\!\!\!\perp Z_1/(X, Z_2)$ tenim que

    $f_{XZ_2YZ_1}(x,z_2,y,z_1) = g'(x,z_2,y)h'(x,z_2,z_1)$ però de $Y \perp\!\!\!\perp Z_2/(X, Z_1)$, $Y \perp\!\!\!\perp Z_2/X$ podem escriure $g'(x,z_2,y) = g''(x,y)g'''(z_2,x)$ per tant $f_{XZ_2YZ_1}(x,z_2,y,z_1) = g''(x,y)g'''(z_2,x)h'(x,z_2z_1) = g''(x,y)h''(x,z_1,z_2)$ i això implica (1).
  - $\perp$ (2) $\Rightarrow$ (3). Resulta del lema de reducció.
  - $\perp$ (3) $\Rightarrow$ (1). De $Y \perp\!\!\!\perp Z_2/(X, Z_2)$ tenim que

    $f_{XYZ_1Z_2} = f_{XYZ_1}\dfrac{f_{XZ_1Z_2}}{f_{XZ_1}}$ però com $Y \perp\!\!\!\perp Z_1/X$ tenim $f_{XYZ_1Z_2} = f_{XY}f_{XZ_1}f_{XZ_1Z_2}/f_Xf_{XZ_1} = f_{XY}f_{XZ_1Z_2}/f_X$ que vol dir $Y \perp\!\!\!\perp (Z_1, Z_2)/X$.

11. BIBLIOGRAFIA

[1] **Darroch, J.N., Lauritzen, S.L.** i **Speed, T.P.** (1980). «Markov fiels and log linear interaction models for contingency tables». *Ann. Stat.*, **8**, 522-539.

[2] **Edwards, D.** (1995). *Introduction to Graphical Modelling*. Springer-Verlag.

[3] **Edwards, D.E.** i **Kreiner, S.** (1983). «The analysis of contingency tables by graphical models». *Biometrika*, **70, 3**, 553-565.

[4] **Feller, W.** (1968). *An Introduction to Probability Theory and its Applications*, Volume 1. John Willey & Sons.

[5] **Geisser, S.** i **Mantel, N.** (1962). «Pairwise independence of jointly dependent variables». *Ann. Math. Statist.*, **33**, 290.

[6] **Mardia, K.V., J. Kent** and **J. Bibby** (1979). *Multivariate Analysis*. Academic Press Limited.

[7] **Simpson, E.H.** (1951). «The Interpretation of Interaction in Contingency Tables». *Journal of the Royal Statistical Society*, **B-13, 2**, 238-241.

[8] **Speed T.P.** (1978). «Graph-teoretic methods in the analysis of interaction». *Lecture Notes*, Institute of Mathematical Statistics, University of Copenhagen.

[9] **Wermuth, N.** i **Lauritzen, S.L.** (1983). «Graphical and recursive models for contingency tables». *Biometrika*, **70, 3**, 537-552.

[10] **Whittaker, J.** (1990). *Graphical Models in Applied Multivariate Statistics*. John Wiley & Sons.

ENGLISH SUMMARY

INDEPENDENCE GRAPHS MODELS

J.M. DURAN RÚBIES
Universitat Autònoma de Barcelona*

Independence Graphical Models are a tool of the Multivariate Analysis which uses graphs to represent models. Independence graphs specially sum up and clarify the interactions among variables; interactions which are not always easy to interpret, specially when three or more take part.

This work is of pedagogical interest and gives an introduction to the Independence Graphs theory, starting with notions of independence and getting to the Markov's properties, keys to the interpretations of these graphs.

It also studies an application example using numerical variables that shows the feasibility of this theory when it comes to simplifying and clarifying the relations among variables.

Keywords: Independence, conditional independence, Simpson's paradox, random vectors, graph theory, conditional independence graphs, Markov properties.

AMS Classification (MSC 2000): 62 H 20

* J.M. Duran Rúbies. Professor associat U.A.B. Departament de Matemàtiques. 08193 Bellaterra (Espanya). E-mail: duran@mat.uab.es

– Received November 1996.
– Accepted December 1998.

The independence graph theory is quite new. One of the first to relate the graph theory methods with the interactions among variables was Speed [8]. Conditional independence graphs sum up and clarify the relations between three of more variables, relations which are not easy to see nor to interpret.

This work is of pedagogical interest and gives an introduction to the independence graphs theory and it is an approach for all those who, without having to look at books, such as Whittaker [10] and Edwards [2], would like to have a graphical models early idea.

Firstly we study Independence notions (the marginal independence definitions, the mutual independence) and conditional independence (the weak conditional independence, the strong conditional independence, the block independence) which are needed to develop the theory. An example is introduced to show the Simpson's paradox which is useful to clarify some results. To be able to show the separation theorem later, previous concepts are extended to random vectors and important results are given, like factorisation criterion and reduction lemma.

Next we study the basic notions of the graph theory in order to define what it is understood as conditional independence graph. The separation theorem, that allows us to detect and remove redundant variables among conditioning variables in an independent relation, it is shown through an example.

Further on, Markov properties are given and their equivalence is demonstrated allowing us to interpret Independence graphs. An application example with numerical variables is shown. In this example and from the correlation matrix a conditional independence graph is built, allowing us, thanks to the global Markov property, to sum up and clarify the conditional independence among variables. In the same way, and thanks the local Markov property, the prediction of variables also remains clarify.

Demonstrations of the some of the results stablished in the text are assembled in the appendix.

SOLUCIONS ALS PROBLEMES PROPOSATS AL VOLUM 22 N. 3

PROBLEMA N. 72

La función característica de la distribución uniforme en el intervalo $(-1,\,+1)$ es

$$\int_{-1}^{+1} \frac{1}{2}\, e^{itx}\, d\,x = \frac{1}{2}\left[\frac{e^{itx}}{it}\right]_{-1}^{+1} = \frac{e^{it}-e^{-it}}{2it} = \frac{\operatorname{sen}(t)}{t}$$

Si $\varphi(t)$ es la función característica de X, entonces $\varphi(-t)$ es la función característica de $-Y$ y la función característica de $X-Y$ es

$$\varphi_{X\,-\,Y}(t) = \varphi(t)\,\varphi(-t)$$

por ser $X,\,Y$ independientes.

Si existe X tal que $X-Y$ es uniforme en $(-1,\,+1)$, donde Y sigue la misma distribución que X y es independiente de X, entonces necesariamente

$$\alpha\ <\ X\ <\ \alpha+1$$

para alguna constante α. Pero entonces $X-\alpha$ tiene la misma propiedad, así que podemos suponer que el soporte de X es

$$0\ <\ X\ <\ 1$$

Se verifica pues

$$\varphi(t)\,\varphi(-t) = \frac{\operatorname{sen}(t)}{t}$$

Pero

$$\begin{aligned}
\varphi(t)\,\varphi(-t) &= \big(E\,\cos(Xt)+i\operatorname{sen}(Xt)\big)\,\big(E\,\cos(Xt)-i\operatorname{sen}(Xt)\big)\\
&= \big(E\cos(Xt)\big)^2+\big(E\operatorname{sen}(Xt)\big)^2\\
&= \frac{\operatorname{sen}(t)}{t}
\end{aligned}$$

y para $t=\pi$ es $\operatorname{sen}(\pi)/\pi=0$

$$\big(E\,\cos(Xt)\big)^2+\big(E\,\operatorname{sen}(X\pi)\big)^2=0$$

Luego

$$E\big(\,\mathrm{sen}(X\pi)\big) = 0$$

Sin embargo, el soporte de $\mathrm{sen}(X\pi)$ es

$$0 \;<\; \mathrm{sen}(X\pi) \;<\; 1$$

y su esperanza no puede ser $0.$ Luego la variable aleatoria X no puede existir.

C.M. Cuadras
Universitat de Barcelona

PROBLEMA N. 73

Si $H(x, y)$ es la función de distribución conjunta de (X, Y), la distribución de $\big(X,\ G^{-1}(1 - G(Y)\big)$ es

$$\begin{aligned} H^*(x, y) &= P\big(X \leqslant x,\ G^{-1}(1 - G(Y)\big) \leqslant y\big) \\ &= P\big(X \leqslant x,\ Y > G^{-1}(1 - G(y)\big) \\ &= F(x) - H\big(x,\ G^{-1}(1 - G(y)\big) \end{aligned}$$

pues

$$\big[X \leqslant x\big] = \big[X \leqslant x,\ Y \leqslant G^{-1}\big(1 - G(y)\big)\big] \cup \big[X \leqslant x,\ Y > G^{-}(1 - G(y)\big)\big]$$

Si $H = H^- = \max\{F + G - 1\}$ entonces

$$\begin{aligned} H^*(x, y) &= F(x) - H^-\Big(x,\ G^{-1}\big(1 - G(y)\big)\Big) \\ &= F(x) - \max\big\{F(x) + GG^{-1}\big(1 - G(y)\big) - 1,\ 0\big\} \\ &= F(x) - \max\{F(x) - G(y),\ 0\} \\ &= \left\{\begin{array}{lllll} F(x) - \big(F(x) - G(y)\big) & = & G(y) & \text{si} & F(x) \geqslant G(y) \\ F(x) - 0 & = & F(x) & \text{si} & F(x) < G(y) \end{array}\right\} \\ &= \min\{F(x),\ G(y)\} = H^+(x, y) \end{aligned}$$

La demostración de la segunda relación

$$H^+(x, y) = G(y) - H^-\big(F^{-1}(1 - F(x)),\ y\big)$$

es similar.

C.M. Cuadras
Universitat de Barcelona

PROBLEMES PROPOSATS

PROBLEMA N. 74

Demostrar por inducción en n la siguiente identidad

$$\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2=\frac{1}{n^2}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}(x_i-x_j)^2\,,$$

donde

$$\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i.$$

M. Ruíz Espejo
U.N.E.D.

PROBLEMA N. 75

Demostrar algebraicamente que el estimador jackknife de la varianza de la media muestral $\bar{x}_n=(1/n)\sum_{i=1}^{n}x_i$, es:

$$\widehat{V}_J\left(\bar{x}_n\right)=\frac{1}{n\left(n-1\right)}\sum_{i=1}^{n}\left(x_i-\bar{x}_n\right)^2.$$

M. Ruíz Espejo
U.N.E.D.