Probabilitat i Estadística

FIB-UPC

Problemes d’e-status:

B1 – Càlcul de probabilitats

Qui és més emprenedor?

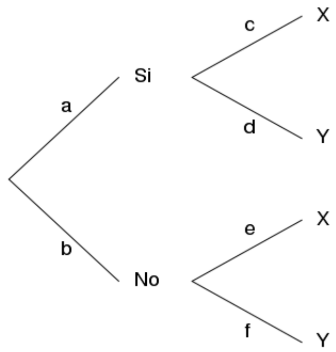
Correcció: Qui és més emprenedor?

La taula que figura a continuació conté les probabilitats conjuntes de dues famílies d'esdeveniments que, quan són observats a determinada població, es poden considerar com a aleatoris:

	Sí	No
X	0.213	0.414
Y	0.155	0.218

X i Y són universitats on els individus enquestats han estudiat, i se'ls ha demanat si havien posat en marxa algun tipus d'iniciativa privada.

L'arbre adjunt representa la mateixa experiència. Responeu a les preguntes que es mostren a continuació.



✓	1. Quin és el valor de la probabilitat que correspon a la branca marcada amb una *a*? Procureu entrar tots els valors decimals. Nota: 2	0.368
✓	2. Quin és el valor de la probabilitat que correspon a la branca marcada amb una *c*? És suficient que poseu tres decimals correctes (si heu d'arrodonir el tercer, fixeu-vos en el quart!) Nota: 2	0.579
✓	3. Quin és el valor de la probabilitat que correspon a la branca marcada amb una *e*? És suficient que poseu tres decimals correctes (si heu d'arrodonir el tercer, fixeu-vos en el quart!) Nota: 2	0.655
✓	4. Quin és el valor de la probabilitat que correspon a la branca marcada amb una *f*? Per la precisió, seguiu les mateixes instruccions que a les preguntes anteriors. Nota: 2	0.345
✓	5. Un individu ens ha dit que ha estudiat a X: probabilitat que sigui un emprenedor. Nota: 2	0.340

Resultat	
Nota	10
Puntuació	4.000

Script en R

```
xsi = 0.213; xno = 0.414; ysi = 0.155; yno = 0.218;
a = sum(xsi,ysi)
a
b = 1-a
b
c = xsi/a
c
d = 1-c
d
e = xno/sum(xno,yno)
e
f = 1-e
f
p5 = xsi/sum(xsi,xno)
p5
```

Consola de R

```
> xsi = 0.213; xno = 0.414; ysi = 0.155; yno = 0.218;
> a = sum(xsi,ysi)
> a
[1] 0.368
> b = 1-a
> b
[1] 0.632
> c = xsi/a
> c
[1] 0.5788043
> d = 1-c
> d
[1] 0.4211957
> e = xno/sum(xno,yno)
> e
[1] 0.6550633
> f = 1-e
> f
[1] 0.3449367
> p5 = xsi/sum(xsi,xno)
> p5
[1] 0.3397129
```

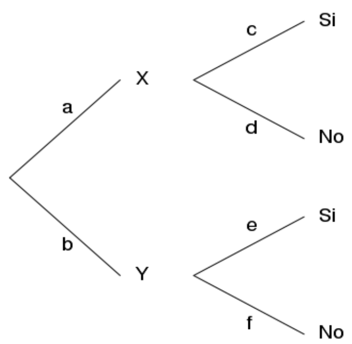
Correcció: Qui és més emprenedor?

La taula que figura a continuació conté les probabilitats conjuntes de dues famílies d'esdeveniments que, quan són observats a determinada població, es poden considerar com a aleatoris:

	Sí	**No**
X	0.366	0.347
Y	0.161	0.126

X i Y són les universitats on els individus enquestats han estudiat, i se'ls ha demanat si havien posat en marxa algun tipus d'iniciativa privada.

L'arbre adjunt representa la mateixa experiència. Responeu a les preguntes que es mostren a continuació.



✓	**1. Quin és el valor de la probabilitat que correspon a la branca marcada amb una *a*? Procureu entrar tots els valors decimals. Nota: 2**	0.713
✓	**2. Quin és el valor de la probabilitat que correspon a la branca marcada amb una *c*? És suficient que poseu tres decimals correctes (si heu d'arrodonir el tercer, fixeu-vos en el quart!) Nota: 2**	0.513
✓	**3. Quin és el valor de la probabilitat que correspon a la branca marcada amb una *e*? És suficient que poseu tres decimals correctes (si heu d'arrodonir el tercer, fixeu-vos en el quart!) Nota: 2**	0.561
✓	**4. Quin és el valor de la probabilitat que correspon a la branca marcada amb una *d*? Per la precisió, seguiu les mateixes instruccions que a les preguntes anteriors. Nota: 2**	0.487
✓	**5. Un individu ens ha dit que és un emprenedor: probabilitat que hagi estudiat a X. Nota: 2**	0.694

Resultat	
Nota	**10**
Puntuació	**4.000**

Script en R

```
xsi = 0.366; xno = 0.347; ysi = 0.161; yno = 0.126;
a = sum(xsi,xno)
a
b = 1-a
b
c = xsi/a
c
d = 1-c
d
e = ysi/sum(ysi,yno)
e
f = 1-e
f
p5 = xsi/sum(xsi,ysi)
p5
```

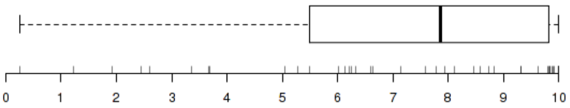
Consola de R

```
> xsi = 0.366; xno = 0.347; ysi = 0.161; yno = 0.126;
> a = sum(xsi,xno)
> a
[1] 0.713
> b = 1-a
> b
[1] 0.287
> c = xsi/a
> c
[1] 0.513324
> d = 1-c
> d
[1] 0.486676
> e = ysi/sum(ysi,yno)
> e
[1] 0.5609756
> f = 1-e
> f
[1] 0.4390244
> p5 = xsi/sum(xsi,ysi)
> p5
[1] 0.6944972
```

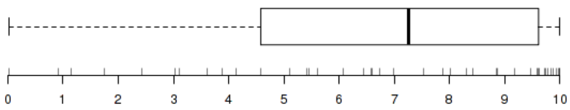
El ojímetro

Correcció: El ojímetro

En la gráfica vas a ver el boxplot de unos datos simulados según una distribución de forma diversa (asimétrica por la derecha, por la izquierda, concentrada, dispersa, ...). El único rasgo común es que los valores (las marcas dibujadas sobre el eje) fluctúan siempre entre 0 y 10.



(Los datos que figuran a continuación son los que se utilizan a partir de la pregunta 5)
[Copiar dades per enganxar a un altre programa]

	1. ¿Qué media aproximada debe tener esta muestra? **Nota: 1.125**	**7**
	2. ¿Cuánto vale la mediana de los datos? **Nota: 1.125**	**7.8**
	3. Vamos ahora a valorar la dispersión de los datos. ¿Cuánto vale el intervalo intercuartilar (distancia entre Q3 y Q1)? **Nota: 1.25**	**4.2**
	4. Estima a continuación el valor de la desviación típica de los valores de la muestra **Nota: 1.25**	**2.8**
	5. Aunque la distribución de partida es la misma que la de la muestra anterior, esta otra gráfica procede de unos datos distintos, por lo que te has de esperar indicadores similares pero no idénticos. Coge los datos adjuntos, y halla con R la media de esta segunda muestra (con tres decimales exactos).  **Nota: 1.25**	**6.755**
	6. ¿Cuanto vale la nueva mediana? **Nota: 1.25**	**7.255**
	7. Obtén los cuartiles y calcula la distancia intercuartilar. **Nota: 1.25**	**4.892**
	8. Halla con R el valor de la desviación típica de la segunda muestra (tres decimales correctos). **Nota: 1.25**	**2.974**

Resultat	
Nota	**9.75**

Script en R

```
# seleccionar los datos en e-status y Ctrl+C
datos <- read.table("clipboard", header=FALSE, col.names=c('dat'))
datos
resumen <- summary(datos$dat)
resumen
p5 = mean(datos$dat)
p5
p6 = median(datos$dat)
p6
p7 = getElement(resumen, "3rd Qu.") - getElement(resumen, "1st Qu.")
p7
p8 = sd(datos$dat)
p8
```

Consola de R

```
> datos <- read.table("clipboard", header=FALSE, col.names=c('dat'))
> datos
     dat
1   9.62
2   7.87
3   2.42
4   8.02
5   9.94
6   3.03
7   4.58
8   3.11
9   9.74
10  0.91
11  9.71
12  8.86
13  6.74
14  5.61
15  9.18
16  0.02
17  6.58
18  8.41
19  8.85
20  5.10
21  4.13
22  9.77
23 10.00
24  9.61
25  8.31
26  6.45
27  9.46
28  5.46
29  6.59
30  1.74
31  7.52
32  3.89
33  6.08
34  3.61
35  1.14
36  9.84
37  6.99
38  9.58
39  5.41
40 10.00
41  9.87
42  9.97
> resumen <- summary(datos$dat)
> resumen
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
  0.020   4.710   7.255   6.755   9.602  10.000 
> p5 = mean(datos$dat)
> p5
[1] 6.755238
> p6 = median(datos$dat)
> p6
[1] 7.255
> p7 = getElement(resumen, "3rd Qu.") - getElement(resumen, "1st Qu.")
> p7
[1] 4.892
> p8 = sd(datos$dat)
> p8
[1] 2.973508
```

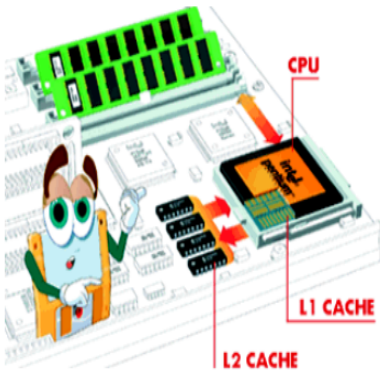
Arquitectura de Computadors

Correcció: Arquitectura de Computadors

La jerarquia de memòria d'un processador està formada per 2 nivells de cache (L1 i L2) i la memòria principal. Com ja sabeu, el funcionament és el següent: quan el processador necessita una dada, el primer lloc on va a buscar-la és a L1; si no la troba va a L2, i si tampoc hi és ha d'accedir a la memòria principal. Per a un sistema concret, les seves característiques venen determinades per aquests paràmetres:

- probabilitat d'encertar a la primera cache L1: 0.903
- probabilitat d'encertar a la segona cache L2 (si L1 falla): 0.705

Responeu a les qüestions que venen a continuació (procureu no perdre decimals en càlculs intermitjos, i responeu amb almenys 4 decimals correctes).



✓	1. Quina és la probabilitat que es trobi una dada a alguna de les caches? Nota: 2.5	0.971385
✓	2. Quina és la probabilitat d'haver trobat una dada a L2, sabent que no s'ha accedit a memòria principal? Nota: 2.5	0.07039948
✓	3. Imaginem un sistema similar, on la probabilitat d'encert de la cache L2 és la mateixa, però no coneixem la probabilitat d'encert de L1. El que sí sabem és que el 78% de les vegades que obtenim una dada sense necessitat d'accedir a memòria principal, aquesta ve concretament de la cache L2. Calculeu la probabilitat d'encert de L1. Nota: 2.5	0.1658646
✓	4. S'està treballant amb un disseny en el que les dues caches són idèntiques i amb la mateixa probabilitat de retornar la dada demanada. Quina és aquesta probabilitat si es vol que el nou disseny només falli amb probabilitat 0.055? Nota: 2.5	0.7654792

Resultat	
Nota	10

Script en R

```
L1 = 0.903
L2 = 0.705
p1 = L1 + L2 - L1*L2
p2 = ((1 - L1)*L2)/p1
p3 = (L2 - L2*0.78) / (0.78 + L2 - 0.78*L2)
p4 = 1 - sqrt(0.055)
p1
p2
p3
p4
```

Consola de R

```
> L1 = 0.903
> L2 = 0.705
> p1 = L1 + L2 - L1*L2
> p2 = ((1 - L1)*L2)/p1
> p3 = (L2 - L2*0.78) / (0.78 + L2 - 0.78*L2)
> p4 = 1 - sqrt(0.055)
> p1
[1] 0.971385
> p2
[1] 0.07039948
> p3
[1] 0.1658646
> p4
[1] 0.7654792
```

La seguretat de la botiga

Correcció: La seguretat de la botiga

En una botiga amb alarma, la probabilitat que una nit qualsevol hi hagi un intent de robatori és 0.023. Si hi ha intent de robatori, l'alarma es dispara amb probabilitat 0.849 i si no n'hi ha, la probabilitat que es dispari és de 0.28. Es demana que calculeu (amb quatre decimals exactes, almenys):

✓	1. Quina és la probabilitat que soni l'alarma una nit qualsevol? Nota: 2	0.293087
✓	2. La probabilitat que hi hagi un intent de robatori i l'alarma soni. Nota: 2	0.019527
✓	3. La probabilitat que hi hagi un intent de robatori si l'alarma sona. Nota: 2	0.06662527
✓	4. Una certa nit l'alarma no ha sonat. Probabilitat que efectivament no hi hagi hagut un intent de robatori? Nota: 2	0.9950871
✓	5. Una altra botiga porta el mateix sistema d'alarma (que es dispara d'acord amb les probabilitats especificades adalt). Aquesta alarma sona el 38.7% de les nits, amb motiu o no. Es tracta de trobar la probabilitat d'intent de robatori per a aquesta segona botiga. Nota: 2	0.1880492

Resultat	
Nota	10

Script en R

```
pr = 0.023; par = 0.849; panor = 0.28;
p1 = pr*par + (1-pr)*panor
p2 = pr*par
p3 = p2/p1
p4 = ((1-pr)*(1-panor)) / (pr*(1-par) + (1-pr)*(1-panor))
p5 = 1 - ((0.387-par) / (panor-par))
p1
p2
p3
p4
p5
```

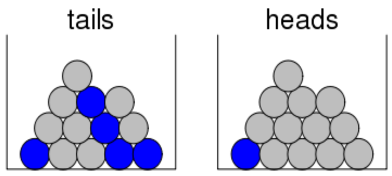
Consola de R

```
> pr = 0.023; par = 0.849; panor = 0.28;
> p1 = pr*par + (1-pr)*panor
> p2 = pr*par
> p3 = p2/p1
> p4 = ((1-pr)*(1-panor)) / (pr*(1-par) + (1-pr)*(1-panor))
> p5 = 1 - ((0.387-par) / (panor-par))
> p1
[1] 0.293087
> p2
[1] 0.019527
> p3
[1] 0.06662527
> p4
[1] 0.9950871
> p5
[1] 0.1880492
```

Pick the right color

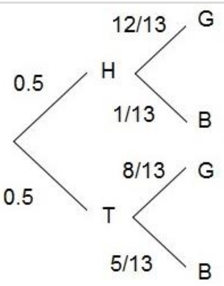
Correcció: Pick the right color

You can see on the right two urns containing grey and blue balls. When a ball is drawn from an urn, it is assumed that it has been chosen randomly. Sometimes the urn is selected before with the aid of a coin (see above the urn). Read and answer the following questions:



1. A ball is picked, it doesn't matter from which urn. What is the probability that you get a grey ball? Nota: 2.222	0.76923
2. You take one ball from each urn. What is the probability that the balls are of different colors? Nota: 2.222	0.40236686
3. You flip a coin, and you obtain heads. What is the probability of getting a blue ball? Nota: 2.222	0.076923
4. A different flip of the coin: a blue ball was drawn from the selected urn. Find the probability that the coin was heads. Nota: 2.222	0.1667
5. You are asked to choose an urn with the coin, and pick two balls (without replacement) from it. The balls happen to be both of grey color. What is the probability that the coin was tails? Nota: 1.111	0.2979

Resultat	
Nota	10



1. $P(G) = 0.5 * 12/13 + 0.5 * 8/13 = 0.76923$

2. $P(\text{dif. color}) = (P(G|H) \wedge P(B|T)) + (P(B|H) \wedge P(G|T)) = P(G|H) * P(B|T) + P(B|H) * P(G|T) =$

$= 12/13 * 5/13 + 1/13 * 8/13 = 0.40236686$

3. $P(B|H) = 1/13 = 0.076923$

4. $P(H|B) = P(H \wedge B) / P(B) = (0.5 * 1/13) / ((0.5 * 1/13) + (0.5 * 5/13)) = 0.16$

5. $P(T | G \wedge G) = (8/13 * 7/12) / ((8/13 * 7/12) + (12/13 * 11/12)) = 0.2979$

Tot viatge comença … amb una cua

Correcció: Tot viatge comença... amb una cua

Per estudiar l'eficiència en un aeroport, una primera aproximació ens porta a estudiar les probabilitats d'haver-se d'esperar a l'hora de facturar i al control de passaports (CP). Considerem el cas d'un aeroport on s'ha comprovat que per a un viatger que arriba, la probabilitat de trobar cua a facturació és 0.66. Posteriorment s'ha comprovat que les probabilitats d'esperar o no per el CP depenen del fet d'haver esperat al moment de facturar, com mostra la següent taula:

	per facturar s'ha hagut d'esperar	per facturar no s'ha hagut d'esperar
no espera	0.023	0.443
espera el primer a la cua	0.376	0.436
espera perque ja hi ha una cua	0.601	0.121

Calculeu (amb quatre decimals exactes):

1. la probabilitat de no esperar a facturació ni a CP. **Nota: 1.667**	**0.15062**
2. la probabilitat de haver d'esperar a facturació o a CP. **Nota: 1.667**	**0.84938**
3. la probabilitat de haver d'esperar a un dels llocs (però només a un). **Nota: 1.667**	**0.20456**
4. la probabilitat de ser atés immediatament al CP (és a dir, no hi ha ningú passant el control). **Nota: 1.667**	**0.1658**
5. Si un viatger arriba al CP i ha d'esperar perquè hi ha una persona passant el control, i cap més: ¿quina és la probabilitat d'haver esperat a facturació? **Nota: 1.667**	**0.6260343**
6. Els passatgers A i B arriben a l'aeroport en dos moments independents per agafar els seus vols. Trobeu la probabilitat que els dos passin directament el control de passaports sense fer cua. **Nota: 1.667**	**0.02748964**

Resultat	
Nota	**10**

Script en R

```
pf = 0.66; pnof = 1-pf;
pfep = 0.376; pfeu = 0.601; pfnoe = 0.023;
pnofep = 0.436; pnofeu = 0.121; pnofnoe = 0.443;
p1 = pnof*pnofnoe
p2 = pf + (pf*pfep + pnof*pnofep) + (pf*pfeu + pnof*pnofeu) -
(pf*pfep) - (pf*pfeu)
p3 = pf*pfnoe + pnof*(pnofep+pnofeu)
p4 = pf*pfnoe + pnof*pnofnoe
p5 = (pf*pfep) / (pf*pfep + pnof*pnofep)
p6 = p4^2
p1
p2
p3
p4
p5
p6
```

Consola de R

```
> pf = 0.66; pnof = 1-pf;
> pfep = 0.376; pfeu = 0.601; pfnoe = 0.023;
> pnofep = 0.436; pnofeu = 0.121; pnofnoe = 0.443;
> p1 = pnof*pnofnoe
> p2 = pf + (pf*pfep + pnof*pnofep) + (pf*pfeu + pnof*pnofeu) -
(pf*pfep) - (pf*pfeu)
> p3 = pf*pfnoe + pnof*(pnofep+pnofeu)
> p4 = pf*pfnoe + pnof*pnofnoe
> p5 = (pf*pfep) / (pf*pfep + pnof*pnofep)
> p6 = p4^2
> p1
[1] 0.15062
> p2
[1] 0.84938
> p3
[1] 0.20456
> p4
[1] 0.1658
> p5
[1] 0.6260343
> p6
[1] 0.02748964
```

Victimas del ladrillo

Correcció: Victimas del ladrillo

Un banco clasifica a sus clientes en 4 categorías:

1. posee tarjeta de crédito (T) y domicilia la nómina (N)
2. sin tarjeta (¬T) pero con nómina (N)
3. con tarjeta (T) pero sin nómina (¬N)
4. ni tarjeta (¬T) ni nómina (¬N): cliente poco interesante.

Interesa estudiar la solvencia de un cliente cuando solicita un crédito. Analizando datos almacenados se ha determinado la proporción de clientes solventes (70.6%) e insolventes (29.4%), así como la distribución de las cuatro categorías anteriores según la solvencia del cliente:

Solvente	Probabilidades (1, 2, 3, 4)
Sí	0.399113, 0.32078, 0.216865, 0.063242
No	0.104117, 0.210354, 0.265389, 0.42014

Copiar dades per enganxar a un altre programa

✓	1. ¿Cuál es la probabilidad de que un cliente solvente sólo tenga tarjeta? Nota: 1	0.216865
✓	2. Si analizamos un cliente insolvente, ¿con qué probabilidad éste tendría algún producto (tarjeta o nómina)? Nota: 1	0.57986
✓	3. ¿Cuál es la probabilidad de que un cliente solvente tenga tarjeta? Nota: 1	0.615978
✓	4. En porcentaje: ¿cuántos clientes son de la primera categoría? Nota: 1	31.2384176
✓	5. Ante un cliente, ¿qué probabilidad existe de que fuera insolvente y tuviera la nómina domiciliada? Nota: 1	0.09245
✓	6. Calcule la probabilidad de que un cliente cualquiera tenga nómina domiciliada y sea solvente. Nota: 1	0.508244458
✓	7. De cada 100 clientes: ¿cuántos no tienen tarjeta ni domicilian su nómina? (incluya un decimal exacto en la respuesta) Nota: 1	16.8170012
✓	8. Un cliente de la categoría 4) solicita un crédito al banco. ¿Probabilidad de que sea insolvente? Nota: 1	0.7345
✓	9. Repita la pregunta anterior para un cliente de la categoría 1). Nota: 1	0.09798959215
✓	10. Si con un cliente solvente el banco tiene un beneficio de 1000 euros al año, y pierde 600 con uno insolvente, ¿qué beneficio promedio por cliente tiene el banco? Nota: 1	529.6

Resultat	
Nota	10

Script en R

```
ps = 0.706; pnos = 1-ps;
ps1 = 0.399113; ps2 = 0.32078; ps3 = 0.216865; ps4 = 0.063242;
pnos1 = 0.104117; pnos2 = 0.210354; pnos3 = 0.265389; pnos4 = 0.42014;
p1 = ps3
p2 = pnos*(pnos1+pnos2+pnos3)/pnos
p3 = ps*(ps1+ps3)/ps
p4 = (ps*ps1 + pnos*pnos1)*100
p5 = pnos*(pnos1+pnos2)
p6 = ps*(ps1+ps2)
p7 = (ps*ps4 + pnos*pnos4)*100
p8 = (pnos*pnos4)/(p7/100)
p9 = (pnos*pnos1)/(p4/100)
p10 = ps*1000 - pnos*600
p1
p2
p3
p4
p5
p6
```

```
p7
p8
p9
p10
```

Consola de R

```
> ps = 0.706; pnos = 1-ps;
> ps1 = 0.399113; ps2 = 0.32078; ps3 = 0.216865; ps4 = 0.063242;
> pnos1 = 0.104117; pnos2 = 0.210354; pnos3 = 0.265389; pnos4 = 0.42014;
> p1 = ps3
> p2 = pnos*(pnos1+pnos2+pnos3)/pnos
> p3 = ps*(ps1+ps3)/ps
> p4 = (ps*ps1 + pnos*pnos1)*100
> p5 = pnos*(pnos1+pnos2)
> p6 = ps*(ps1+ps2)
> p7 = (ps*ps4 + pnos*pnos4)*100
> p8 = (pnos*pnos4)/(p7/100)
> p9 = (pnos*pnos1)/(p4/100)
> p10 = ps*1000 - pnos*600
> p1
[1] 0.216865
> p2
[1] 0.57986
> p3
[1] 0.615978
> p4
[1] 31.23842
> p5
[1] 0.09245447
> p6
[1] 0.5082445
> p7
[1] 16.817
> p8
[1] 0.7345017
> p9
[1] 0.09798959
> p10
[1] 529.6
```