IG12 2006/07: Exercicis puntuables part "Probabilitats + Models"

Pablo Gregori

1. Un test sobre una malaltia es comprova sobre una població d'individus dels quals es coneix a priori si estan sans o no, amb els resultats que figuren a la taula:

	Test +	Test -
Sà	3	18
Malalt	13	2

   Usant eixes dades per estimar les probabilitats de que el test done positiu (+) ó negatiu (-), condicionades a que la persona està sana o malalta, i suposant que la malaltia afecta al 3% de la població mundial, calcula la probabilitat de que una persona tinga la malaltia, una vegada el test ho ha confirmat.

---

2. Un sensor d'imatge es fabrica per reconèixer automàticament els colors del semàfor. Al fer-li proves, s'ha vist que, quan la llum estava en verd, el sensor ho detectava com groc en un 5% de les ocasions i com roig en un 1%. Quan estava en groc, el sensor ho detectava com verd en un 8% de les ocasions i com roig en el 30% de les ocasions, mentre que quan estava en roig, el sensor ho detectava com verd en un 10% de les ocasions i com groc en un 25% de les ocasions.

   Si aquest sensor es coloca a un cotxe automàtic, que a un moment donat es troba en un semàfor, i el sensor detecta que està en verd, quina és la probabilitat real de que el semàfor estiga en cadascú dels colors que pot estar?

   Ajuda: El semàfor en qüestio passa cicles de 50 s. en verd, 2 s. en groc i 30 s. en roig.

---

3. Calcula l'esperança de la variable aleatòria X = "premi aconseguit" en el concurs "Allá tú".

   En aquest concurs hi ha 20 premis diferents (veure foto), cadascú ocult a una caixa diferent, que s'assignen aleatòriament als 20 concursants. Un concursant és triat a l'atzar per jugar amb la seua caixa. El joc consisteix a anar descobrint els premis de les caixes dels altres concursants i, a cada pas, acceptar un premi inferior oferit per l'organització, o continuar el procés fins haver descobert totes les altres caixes i quedar-se amb el contingut de la seua pròpia.

   Com la situació és prou complexa, anem a simplificar-la suposant que el concursant sempre es queda fins al final amb la caixa que té assignada per atzar. Quin és el valor mitjà o esperança del premi del concurs?

   Nota: suposeu que els valor monetari dels premis "baile", "ajos" i "tutú" (veure foto) són, respectivament, 0, 0.50 i 150 euros.

---

4. Una font emet bits (valors 0 ó 1) per un canal a un receptor, de manera que l'1% dels bits emesos és invertit pel canal, i per tant, mal rebut pel receptor.

   (a) Si el receptor rep la cadena 0110100, quina és la probabilitat de que continga algun error?

   Per augmentar la fiabilitat del canal s'estableix el següent sistema per detectar errors: per cada 7 bits emesos per l'emisor s'envia un 8é bit (fictici) anomentat "bit de paritat" de manera que val '0' si s'ha enviat un número parell de bits '1', i val '1' en cas contrari. Si el receptor troba que la paritat dels 7 primers bits no casa amb el valor del 8é bit, aleshores pensa que hi ha hagut un error i torna a demanar la cadena de 8 bits a l'emisor.

Per exemple, si l'emisor vol enviar la cadena 0110100, aleshores afegiria el bit de paritat '1', i enviaria en total la cadena 01101001. Si el receptor rep la cadena 01101001, com la paritat és correcta, pensarà que la cadena està ben rebuda, però si rep la cadena 01101000, com la paritat no correspon, aleshores pensarà que hi ha hagut algun error a la transmissió.

Si el receptor rep la cadena 01101001 (l'últim bit és el de paritat, però també pot haver-se transmés amb error!), quina és la probabilitat de que els 7 primers bits continguen algun error?

---

5. Un dispositiu té una probabilitat d'avariar-se del 0.2% a cada vegada que s'utilitza.

   (a) Quin és el número esperat d'utilitzacions del dispositiu abans d'avariar-se?

   (b) Quina és la probabilitat de que s'utilitze el dispositiu en al menys 500 ocasions sense problemes? (planteja-ho i usa R si pots per al resultat numèric final)

---

6. Les alarmes que es reben a una empresa de seguretat arriben a raó de 3.4 alarmes diàries. Si l'empresa té contractats 10 vigilants (que es desplacen immediatament al lloc de l'alarma, i podrien passar en servei tot el dia),

   Calcula la probabilitat de que un dia concret es sature el servei i es queden sense vigilants.

---

7. El primer semàfor que et trobes per incorporar-te al carrer principal passa cicles de 50 s en roig i 15 en verd. Cada volta que eixes de casa i agafes el cotxe, ho fas amb una variabilitat tan gran, que pots considerar que arribes al semàfor en un moment totalment aleatori. Quin percentatge aproximat dels dies de la teua vida en que agafes el cotxe et trobaràs el semàfor en verd?

---

8. Un operador està a càrrec del servei d'atenció al client, rebent i atenent les telefonades dels usuaris, des de la seua casa (és un teletreballador). Per la seua experiència, sap que rep una mitjana de 13.2 telefonades en el seu horari de treball (de 8 hores). El telèfon de treball és fixe.

Un dia, l'operador necessita absentar-se de casa de treball per motius personals, dins de l'horari laboral, però no vol que es sàpia, i creu que amb 30 minuts podrà reincorporar-se.

Calcula la probabilitat de que s'absente eixe temps sense que cap telefonada quede sense atendre.

---

9. Als anàlisis de sang, per cada substància analitzada, figura el valor concret del pacient que es fa l'anàlisi, i un interval de referència al costat, de manera que el metge que interpreta l'informe de l'anàlisi pot alertar al pacient d'una possible anomalia si el valor està fora de l'interval.

   La manera en que es contrueix l'interval és la següent: després d'un estudi molt complet, imaginem que per certa substància concreta, s'accepta la hipòtesi de que la concentració segueix la distribució normal de mitjana 6.5 i desviació típica 0.35.

   Per sistema, en el àmbit de la medicina, es considera acceptable el 95% "central" de valors (és a dir, exceptuant el 2.5% de valors inferiors i superiors, respectivament).

   Per tant, quin interval de referència correspondria a aquesta substància concreta?

---

10. Un dispensador d'un producte té un stock de 100 unitats, que es reponen quan s'acaben totalment (automàticament un missatge s'envia al centre reponedor i en poques hores arriba un operari i ompli amb 100 noves unitats).

    Imagimen que cada ususari d'eixe dispensador demana una quantitat variable de productes que es pot modelitzar amb una variable de Poisson de mitjana 3.25 productes.

    Imaginem tu ets ú dels usuaris que va a usar eixe dispensador, que necessitaràs una quantitat de productes que segueix la mateixa llei de Poisson, i que, com no saps quan han fet la reposició, ni el número d'usuaris que ha usat el dispensador abans que tu, aleshores pots considerar que el "número d'unitats que queden al dispensador" pot ser qualsevol, de 0 a 100, amb la mateixa probabilitat.

    Quina probabilitat tens de que pugues obtindre totes les unitats que necessites en eixe moment del dispensador? (usa R per calcular el resultat numèric final)

Solucions:

1. Mirant la taula, podem deduir que $P(+|S) = 3/21$, $P(-|S) = 18/21$, $P(+|M) = 13/15$, $P(-|M) = 2/15$.

   Ens demanen $P(M|+)$, aleshores

$$P(M|+) = \frac{P(+|M) * P(M)}{P(+)} = \frac{\frac{13}{15} * 0.03}{P(+|M) * P(M) + P(+|S) * P(S)}$$
$$= \frac{\frac{13}{15} * 0.03}{\frac{13}{15} * 0.03 + \frac{3}{21} * 0.97} = 0.1579861$$

2. Usem la notació V, G, R pels esdeveniments de que el semàfor (la llum) **estiga realment** en verd, groc o roig, i sV, sG, sR pels esdeveniments de que el **sensor detecte** verd, groc o roig.

   Aleshores les dades que tenim al fer les proves són condicionades, ja que controlem la llum. Tenim amb la llum verd $P(sG|V) = 0.05$, $P(sR|V) = 0.01$ i lògicament $P(sV|V) = 0.94$. Després amb la llum groga $P(sV|G) = 0.08$, $P(sR|G) = 0.30$ i lògicament $P(sG|G) = 0.62$. Finalment amb la llum roja $P(sV|R) = 0.10$, $P(sG|R) = 0.25$ i lògicament $P(sR|R) = 0.65$.

   Al arribar el cotxe al semàfor no sabem en quin color està, però el sensor ens informa de que ha detectat verd (per tant sabem que passa sV). Ens demanen $P(V|sV)$

$$P(V|sV) = \frac{P(sV|V) * P(V)}{P(sV)}$$
$$= \frac{0.94 * P(V)}{P(sV|V) * P(V) + P(sV|G) * P(G) + P(sV|R) * P(R)}$$
$$= \frac{0.94 * (50/82)}{0.94 * (50/82) + 0.08 * (2/82) + 0.10 * (30/82)} = 0.9370016$$

   bserva que, a falta de més informació, $P(V), P(G)$ i $P(R)$ s'obtenen dels temps que passa en cada color el semàfor.

3. La variable aleatòria $X =$"premi aconseguit" no segueix cap dels models estudiats, però sabem que tots els resultats són equiprobables, i que $X \in \{0, 0.50, 1, 5, 10, 30, 50, 150, 300, 600, 1500, 3000, 6000, 12000, 18000, 24000, 30000, 60000, 12$

   Per tant $f(x_i) = P(X = x_i) = 1/20$, i l'esperança val

$$E(X) = \sum_{i=1}^{20} x_i f(x_i) = \frac{0 + 0.50 + 1 + \cdots + 240000}{20} = 25782.33$$

4. (a) Si pensem en el concepte $X =$"número d'errors que té la cadena", podem plantejar el problema com un model binomial, $X \sim Bin(n = 7, p = 0.01)$. Aleshores ens demanen $P(X > 0) = 1 - P(X = 0) = 1 - f(0) = 1 - 0.99^7 = \boxed{0.06793465}$.

   (b) Si s'afegeix el bit de paritat, i rebem la cadena 01101001, hem de pensar que pot haver error(s) en els bits normals, però tambe pot haver error en el bit de paritat, per tant, anomenant $X =$"número d'errors que té la cadena de bits normal" i $Y =$"valor del bit de paritat".

   Com que el tractament dels bits és diferent, i la paritat depén de la resta de la cadena, és convenient usar probabilitats condicionades.

   Per exemple, les probabilitats condicionades que es poden traure són:

   - $P(Y = 0|X = 0) = P(\text{'paritat errònia'}) = 0.01$, i per tant $P(Y = 1|X = 0) = 0.99$
   - $P(Y = 0|X = 1) = P(\text{'paritat ok'}) = 0.99$, i per tant $P(Y = 1|X = 0) = 0.01$
   - $P(Y = 0|X = 2) = P(\text{'paritat errònia'}) = 0.01$, i per tant $P(Y = 1|X = 0) = 0.99$
   - $P(Y = 0|X = 3) = P(\text{'paritat ok'}) = 0.99$, i per tant $P(Y = 1|X = 0) = 0.01$
   - $P(Y = 0|X = 4) = P(\text{'paritat errònia'}) = 0.01$, i per tant $P(Y = 1|X = 0) = 0.99$
   - $P(Y = 0|X = 5) = P(\text{'paritat ok'}) = 0.99$, i per tant $P(Y = 1|X = 0) = 0.01$
   - $P(Y = 0|X = 6) = P(\text{'paritat errònia'}) = 0.01$, i per tant $P(Y = 1|X = 0) = 0.99$
   - $P(Y = 0|X = 7) = P(\text{'paritat ok'}) = 0.99$, i per tant $P(Y = 1|X = 0) = 0.01$

   Així, juntant-ho tot, calculem la probabilitat de que no hi haja cap error als 7 primers bit, condicionat a que el bit de paritat que veiem es 'Y=1':

$$\begin{aligned} P(X = 0|Y = 1) &= \frac{P(Y = 1|X = 0) * P(X = 0)}{P(Y = 1)} \\ &= \frac{0.99 * 0.99^7}{\sum_{i=0}^{7} P(Y = 1|X = i) * P(X = i)} \\ &= \frac{0.99^8}{0.99 * (f(0) + f(2) + f(4) + f(6)) + 0.01 * (f(1) + f(3) + f(5) + f(7))} \end{aligned}$$

   on $f(x) = \binom{7}{x} 0.99^x 0.01^{7-x}$ per cada $x = 0, 1, \ldots, 7$. El resultat final dóna

$$P(X = 0|Y = 1) = 0.9971506$$

I per tant la probabilitat de que hi haja algun error seria $1-0.9971506 = \boxed{0.0028494}$.

En conclusió, si s'aprofita la informació del bit de paritat, passem d'un 6% aproximat d'errors a un 2‰, que és la 30-èsima part.

5. Si pensem al concepte $X =$"número d'utilitzacions del dispositiu abans d'avariar-se (incloent la de l'avaria)", es tracta del model binomial negatiu, és a dir, $X \sim BinNeg(r = 1, p = 0.002)$. Per tant

   (a) $E(X) = r/p = 500$

   (b) $P(X \geq 501) = 1 - P(X < 501) = 1 - P(X \leq 500) = 1 - F(500) = 1 - (1 - 0.002)^{500} = 0.6324887$

   Es suposa que si volem 500 utilitzacions sense problemes, com la variable conta les utilitzacions fins a la mateixa utilització de l'avaria, aleshores volem 501 ó més (contant la de l'avaria).

6. Si pensem al concepte $X =$"número d'alarmes rebuden en un dia qualsevol", aleshores $X \sim Po(\lambda = 3.4)$.

   Que es sature el servei i es queden sense vigilants implica que $X = 10$ ó més, per tant

   $P(X \geq 10) = 1 - P(X < 10) = 1 - P(X \leq 9) = 1 - F(9) = 0.002708733$

   (es molt baixa, per tant no cal preocupar-se molt en contractar més vigilants).

7. Si pensem en un cicle concret del semàfor (ja que es repeteix), es tracta d'un interval de temps de 0 a 65 segons. Si nosaltres arribem en un moment completament a l'atzar d'eixe interval, aleshores el concepte $X =$"moment de l'interval en el que arribem" segueix la distribució uniforme, és a dir, $X \sim Unif(a = 0, b = 65)$.

   De tot l'interval podem pensar que el subinterval $[0, 50]$ és de color roig i el subiunterval $[50, 65]$ de color verd. Per tant,

   $P(\text{'verd'}) = P(50 < X < 65) = F(65) - F(50) = \frac{65-0}{65-0} - \frac{50-0}{65-0} = 15/65 = 0.2307692$

   Per tant, a llarg termini, al voltant del 23% dels dies trobarem el semàfor en verd.

8. Si pensem al concepte $X =$"número de telefonades rebudes durant l'absència", aleshores $X \sim Po(\lambda = 13.2/16 = 0.825)$. Per tant, ens demanen

   $P(X = 0) = f(0) = 0.438235$

9. Es tracta de trobar els valors [A, B] de manera que $P(X \leq A) = 0.025$ i $P(X \leq B) = 0.975$.

$$
\begin{array}{cc}
P(X \leq B) = 0.975 & P(X \leq A) = 0.025 \\
\downarrow & \downarrow \\
P(\frac{X-6.5}{0.35} \leq \frac{B-6.5}{0.35}) = 0.975 & P(\frac{X-6.5}{0.35} \leq \frac{A-6.5}{0.35}) = 0.025 \\
\downarrow & \downarrow \\
F_Z(\frac{B-6.5}{0.35}) = 0.975 & F_Z(\frac{A-6.5}{0.35}) = 0.025 \\
\downarrow & \downarrow \\
\text{(taules)} & \text{(taules)} \\
\downarrow & \downarrow \\
\frac{B-6.5}{0.35} = 1.96 & \text{(0.025 no està)} \\
\downarrow & \downarrow \\
B = 7.186 & 1 - F_Z(-\frac{A-6.5}{0.35}) = 0.025 \\
 & \downarrow \\
 & F_Z(-\frac{A-6.5}{0.35}) = 0.975 \\
 & \downarrow \\
 & -\frac{A-6.5}{0.35} = 1.96 \\
 & \downarrow \\
 & A = 5.814
\end{array}
$$

10. Aquest problema el deixe sense resoldre. A l'examen no eixirà res de dificultat similar.