Estadística
Curs 2013-2014/2
Grup M3 - Professors: José Rodellar / Francesc Pozo

[Feu els Problemes en Fulls Separats]

Primer Parcial. 2/04/2014

1. `[10 punts]` Si els esdeveniments A i C satisfan que $A \cap B = B \cap C$, on B és un esdeveniment tal que $P(B) > 0$, demostreu que

$$P(A|B) = P(C|B).$$

Solució. Aplicant la definició de probabilitat condicionada:

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

tal i com volíem veure.

---

2. `[30 punts]` Una prova preliminar sobre el contingut alcohòlic en sang d'un conductor, realitzat pels Mossos d'Esquadra, és fiable en el 80% de les ocasions, és a dir, dóna un resultat correcte 8 de cada 10 vegades. Els conductors als quals el resultat de la prova d'alcoholèmia és positiu se sotmeten a una segona prova independent molt més precisa, que mai falla en un conductor sobri però que té un 10% d'error en els embriacs[1].

Sabem que el 5% dels conductors detinguts pels Mossos d'Esquadra són embriacs. Si definim els següents esdeveniments

- $\mathcal{E}$ és l'esdeveniment "ser embriac";
- $\mathcal{P}_1$ és l'esdeveniment "la primera prova dóna un resultat positiu"; i
- $\mathcal{P}_2$ és l'esdeveniment "la segona prova dóna un resultat positiu".

responeu a les següents preguntes següents:

(a) `[5 punts]` Quina proporció de conductors detinguts seran sotmesos a la segona prova? Tingueu present que aquesta proporció ve donada per la proporció de conductors als que la primera prova és positiva, és a dir, $P(\mathcal{P}_1)$.

(b) `[10 punts]` Un conductor té un resultat positiu de la primera prova. Quina és la probabilitat que hagués estat conduint amb un nivell d'alcohol en sang superior al permès legalment, és a dir, $P(\mathcal{E}|\mathcal{P}_1)$?

(c) `[5 punts]` Quina proporció de conductors tindran un resultat negatiu de la segona prova, és a dir, $P(\overline{\mathcal{P}}_2)$? Tingueu present que per fer la segona prova és necessari que hagi donat un resultat positiu de la primera prova. És a dir, $P(\overline{\mathcal{P}}_2)$ és calcula com $P(\mathcal{P}_1 \cap \overline{\mathcal{P}}_2)$, que a la seva vegada es pot calcular aplicant la fórmula de la probabilitat total.

---

[1] Segons el *Gran diccionari de la llengua catalana*, "persona que sol prendre begudes alcohòliques en quantitat excessiva".

(d) `[10 punts]` Un conductor té un resultat negatiu de la segona prova (i per tant un resultat positiu de la primera). Quina és la probabilitat que hagués estat conduint amb un nivell d'alcohol en sang superior al permès legalment, és a dir, $P(\mathcal{E}|\mathcal{P}_1 \cap \overline{\mathcal{P}}_2)$?

SOLUCIÓ.

(a) Aplicant la fórmula de la probabilitat total tenim:

$$\begin{aligned} P(\mathcal{P}_1) &= P(\mathcal{P}_1|\mathcal{E})P(\mathcal{E}) + P(\mathcal{P}_1|\overline{\mathcal{E}})P(\overline{\mathcal{E}}) \\ &= 0.80 \cdot 0.05 + 0.20 \cdot 0.95 = 0.23 \end{aligned}$$

És a dir, el 23% dels conductors detinguts serà sotmés a la segona prova.

(b) Aplicant el Teorema de Bayes, tenim que:

$$P(\mathcal{E}|\mathcal{P}_1) = \frac{P(\mathcal{P}_1|\mathcal{E})P(\mathcal{E})}{P(\mathcal{P}_1)} = \frac{0.80 \cdot 0.05}{0.23} = \frac{4}{23} \approx 0.1739130435$$

És a dir, només el 17.39% dels conductors que tenen un resultat positiu de la primera prova són embriacs.

(c) Aplicant la fórmula de la probabilitat total, tenim que:

$$\begin{aligned} P(\mathcal{P}_1 \cap \overline{\mathcal{P}}_2) &= P(\mathcal{P}_1 \cap \overline{\mathcal{P}}_2|\mathcal{E})P(\mathcal{E}) + P(\mathcal{P}_1 \cap \overline{\mathcal{P}}_2|\overline{\mathcal{E}})P(\overline{\mathcal{E}}) \\ &= P(\mathcal{P}_1|\mathcal{E})P(\overline{\mathcal{P}}_2|\mathcal{E})P(\mathcal{E}) + P(\mathcal{P}_1|\overline{\mathcal{E}})P(\overline{\mathcal{P}}_2|\overline{\mathcal{E}})P(\overline{\mathcal{E}}) \\ &= 0.80 \cdot 0.10 \cdot 0.05 + 0.20 \cdot 1 \cdot 0.95 = \frac{97}{500} \approx 0.194 \end{aligned}$$

Això significa que el 19.4% dels conductors detinguts obtenen un resultat negatiu de la segona prova.

(d) Aplicant el Teorema de Bayes, tenim que:

$$P(\mathcal{E}|\mathcal{P}_1 \cap \overline{\mathcal{P}}_2) = \frac{P(\mathcal{P}_1 \cap \overline{\mathcal{P}}_2|\mathcal{E})P(\mathcal{E})}{P(\mathcal{P}_1 \cap \overline{\mathcal{P}}_2)} = \frac{0.80 \cdot 0.10 \cdot 0.05}{0.194} = \frac{2}{97} \approx 0.02061855670$$

Aquest resultat implica que només el 2.06% dels conductors que tenen un resultat negatiu del segon test són embriacs.

---

3. `[30 punts]` Una variable aleatòria discreta X pren només els valors $\{1, 2, 3\}$. Sabem que la probabilitat de l'esdeveniment $X = 2$ és el doble de la probabilitat de l'esdeveniment $X = 3$. Sabem també que la variància és $\frac{9}{16}$.

(a) `[10 punts]` Calculeu la funció de probabilitat $f(x)$ de la variable X.

(b) `[10 punts]` Calculeu la funció de distribució de probabilitat $F(x)$ de la variable X.

(c) `[5 punts]` Calculeu la probabilitat que X prengui un valor inferior o igual a 2. Calculeu també la probabilitat que la variable prengui un valor més gran que 3.

(d) `[5 punts]` Sigui la variable $Z = 3 + X^2$. Calculeu $E(Z)$.

Solució.

(a) Suposem que $P(X = 3) = p$. Sabem, per les dades del problema, que $P(X = 2) = 2P(X = 3) = 2p$. També sabem que

$$\begin{aligned} 1 = P(\Omega) &= P((X=1) \cup (X=2) \cup (X=3)) = P(X=1) + P(X=2) + P(X=3) \\ &= P(X=1) + 2p + p = P(X=1) + 3p, \end{aligned}$$

el que implica que $P(X = 1) = 1 - 3p$. L'esperança de la variable és:

$$E(X) = \sum_{k=1}^{3} k \cdot P(X = k) = 1 \cdot (1 - 3p) + 2 \cdot 2p + 3 \cdot p = 1 + 4p$$

L'esperança del quadrat de la variable és:

$$E(X^2) = \sum_{k=1}^{3} k^2 \cdot P(X = k) = 1 \cdot (1 - 3p) + 2^2 \cdot 2p + 3^2 \cdot p = 1 + 14p$$

Aleshores, la variància és:

$$\begin{aligned} \frac{9}{16} = VAR(X) &= E(X^2) - [E(X)]^2 \\ &= 1 + 14p - (1 + 4p)^2 = -16p^2 + 6p, \end{aligned}$$

el que implica que el valor de p ha de satisfer l'equació

$$16p^2 - 6p + \frac{9}{16} = 0.$$

L'equació anterior té com a única solución (doble) el valor $p = \frac{3}{16} = 0.1875$. Aleshores, la funció de probabilitat de la variable aleatòria X és

$$f(x) = \begin{cases} 1 - 3p = \frac{7}{16}, & x = 1 \\ 2p = \frac{3}{8}, & x = 2 \\ p = \frac{3}{16}, & x = 3 \\ 0, & \text{altrament} \end{cases}$$

(b) La funció de distribució de probabilitat de la variable aleatòria X és

$$F(x) = P(X \leq x) = \begin{cases} 0, & x < 1 \\ P(X = 1) = \frac{7}{16}, & 1 \leq x < 2 \\ P(X = 1) + P(X = 2) = \frac{13}{16}, & 2 \leq x < 3 \\ P(X = 1) + P(X = 2) + P(X = 3) = 1, & x \geq 3 \end{cases}$$

(c) Les probabilitats demanades són:

$$\begin{aligned} &P(X \leq 2) = F(2) = \frac{13}{16} \\ &P(X > 3) = 1 - P(X \leq 3) = 1 - F(3) = 1 - 1 = 0 \end{aligned}$$

(d) Si X és una variable que pren els valors $\{1, 2, 3\}$ amb probabilitats $P(X = 1) = \frac{7}{16}, P(X = 2) = \frac{3}{8}$ i $P(X = 3) = \frac{3}{16}$, respectivament, aleshores la variable $Z = 3 + X^2$ és una variable aleatòria que pren valors $\{3 + 1^2, 3 + 2^2, 3 + 3^2\} = \{4, 7, 12\}$ amb probabilitats $P(Z = 4) = P(X = 1) = \frac{7}{16}, P(Z = 7) = P(X = 2) = \frac{3}{8}$ i $P(Z = 12) = P(X = 3) = \frac{3}{16}$, respectivament. D'aquesta manera, tenim que:

$$E(Z) = 4 \cdot P(Z = 4) + 7 \cdot P(Z = 7) + 12 \cdot P(Z = 12) = \frac{53}{8} = 6.625$$

---

4. `[30 punts]` Es té coneixement que el nombre de submarins que fan escala en el port de Barcelona per any s'ajusta a un model de Poisson. Sabem que el percentatge d'anys on arriba només un submarí és tres vegades el percentatge d'anys on arriben 3.

(a) `[6 punts]` Trobeu el nombre mitjà de submarins que fan escala per any.

(b) `[6 punts]` Calculeu la probabilitat que en un any no arribi cap submarí.

(c) `[6 punts]` Considerem un període de 10 anys de forma conjunta. Quin és el valor mitjà de submarins que arribaran en aquest període? Calculeu la probabilitat que en aquest període vingui al menys un submarí.

Un equip de bàsquet busca nois amb una edat inferior als 14 anys i una alçada de més de 1.90 metres. Per tal d'identificar aquests futurs jugadors, es fan trucades a números de telèfon trets aleatòriament i de forma independent d'una base de dades d'una gran ciutat, on es coneix que la proporció de nois amb aquestes característiques és del 10%.

(d) `[6 punts]` Si es fan 5 trucades, quin és el model que ens permet calcular la probabilitat d'identificar un cert nombre de jugadors? Calculeu la probabilitat que aquest número sigui almenys 1.

(e) `[6 punts]` Calculeu la probabilitat que s'hagin de fer 5 trucades per identificar 2 jugadors. Quin és el model que ens permet calcular aquesta probabilitat?

SOLUCIÓ.

(a) Sigui X la variable aleatòria que compta el nombre de submarins que fan escala en el port de Barcelona en un any. Segons l'enunciat del problema, $X \hookrightarrow \mathcal{P}(\lambda)$ on $\lambda > 0$ no és conegut. Si sabem que

$$P(X = 1) = 3P(X = 3),$$

aleshores

$$\begin{aligned} e^{-\lambda}\frac{\lambda^1}{1!} = 3e^{-\lambda}\frac{\lambda^3}{3!} &\Leftrightarrow \lambda = \frac{\lambda^3}{2} \\ &\Leftrightarrow \lambda^2 = 2 \\ &\Leftrightarrow \lambda = \sqrt{2} \end{aligned}$$

Per tant, el nombre mitjà de submarins que fan escala per any és $E(X) = \lambda = \sqrt{2}$.

(b) Les probabilitats demanades són:

$$P(X=0) = e^{-\sqrt{2}}\frac{\sqrt{2}^0}{0!} = e^{-\sqrt{2}} \approx 0.2431167345$$

(c) Sigui ara Y la variable que compta el nombre de submarins que arriba al port de Barcelona en un període de 10 anys. Sabem que aquesta variable es pot definir com

$$Y = X_1 + X_2 + \cdots + X_{10},$$

on cada $X_i \hookrightarrow \mathcal{P}(\sqrt{2})$ mesura el nombre de submarins que arriben al port l'i-èssim any. Aleshores, se sap que $Y \hookrightarrow \mathcal{P}(10\sqrt{2})$. Per tant, el nombre mitjà de submarins que arribaran en aquest període és $E(Y) = 10\sqrt{2} \approx 14.14213562$.

D'altra banda, també ens demanen la probabilitat $P(Y \geq 1)$. Aquesta probabilitat la calcularem com segueix:

$$\begin{aligned} P(Y \geq 1) = 1 - P(Y < 1) = 1 - P(Y=0) = 1 - e^{-10\sqrt{2}}\frac{\lambda^0}{0!} \\ = 1 - e^{-10\sqrt{2}} \approx 0.9999992786. \end{aligned}$$

(d) Sigui J la variable aleatòria que compta, d'entre el total de 5 trucades, en quantes s'identifica a un jugador amb les característiques desitjades. Aquesta variable segueix una distribució binomial de paràmetres $n = 5$ i $p = 0.1$, és a dir,

$$J \hookrightarrow \mathcal{B}(5, 0.1).$$

La probabilitat demanada és, doncs,

$$\begin{aligned} P(J \geq 1) = 1 - P(J < 1) = 1 - P(J = 0) \\ = 1 - \binom{5}{0} \cdot 0.1^0 \cdot 0.9^5 \approx 0.590490 \end{aligned}$$

(e) Sigui ara $K^\star$ la variable aleatòria que compta el nombre de trucades que hem de fer fins que s'identifiquen un total de dos jugadors amb les característiques desitjades. Si considerem ara la variable K que compta el nombre de trucades que fem on no identifiquem a un jugador (fins que n'hem identificat a un parell, és clar que $K = K^\star - 2$. La variable K segueix una distribució binomial negativa de paràmetres $r = 2$ i $p = 0.1$, és a dir,

$$K \hookrightarrow \mathcal{BN}(2, 0.1).$$

La probabilitat demanada és

$$P(K^\star = 5) = P(K = 3) = \binom{4}{3} \cdot 0.9^3 \cdot 0.1^2 \approx 0.02916$$

Model	$P(X=k)$	$E(X)$	$VAR(X)$
Bernoulli $b(p)$	$P(X=k)=p^k(1-p)^{1-k}$	p	pq
Binomial $\mathcal{B}(n,p)$	$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$	np	npq
Poisson $\mathcal{P}(\lambda)$	$P(X=k)=e^{-\lambda}\frac{\lambda^k}{k!}$	λ	λ
Geomètrica $\mathcal{G}(p)$	$P(X=k)=(1-p)^k p$	$\frac{1-p}{p}$	$\frac{1-p}{p^2}$
Binomial Negativa $\mathcal{BN}(r,p)$	$P(X=k)=\binom{k+r-1}{k}(1-p)^k p^r$	$\frac{r(1-p)}{p}$	$\frac{r(1-p)}{p^2}$