Estadística
Curs 2016-2017/1
Grup M1 - Professor: Francesc Pozo

[Feu els Problemes en Fulls Separats]

Primer Parcial. 09/11/2016. Versió B.

1. [20 punts]

(a) Donats A i B dos esdeveniments qualssevol, amb $P(B) > 0$, demostreu que $P(A|B) + P(\bar{A}|B) = 1$.

Solució: Considerem la següent expressió per a B:

$$B = B \cap \Omega = B \cap (A \cup \bar{A}) = (B \cap A) \cup (B \cap \bar{A}),$$

on clarament els esdeveniments $B \cap A$ i $B \cap \bar{A}$ són mútuament excloents o incompatibles. Aleshores,

$$\begin{aligned} P(A|B) + P(\bar{A}|B) &= \frac{P(A \cap B)}{P(B)} + \frac{P(\bar{A} \cap B)}{P(B)} = \frac{P(A \cap B) + P(\bar{A} \cap B)}{P(B)} \\ &= \frac{P((A \cap B) \cup (\bar{A} \cap B))}{P(B)} = \frac{P(B)}{P(B)} = 1 \end{aligned}$$

(b) Siguin A i B dos esdeveniments tal que $P(A) > 0, P(B) > 0$ i $A \cap B = \emptyset$. Són independents? Justifiqueu la vostra resposta.

Solució: Els esdeveniments A i B no poden ser independents ja que són incompatibles. Més concretament,

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(\emptyset)}{P(B)} = 0 \neq P(A) > 0.$$

Com que $P(A|B) \neq P(A)$, els dos esdeveniments no poden ser independents.

2. [20 punts] Un professor universitari té tres comptes de correu, el correu de la UPC, el correu de Google (gmail) i el correu de l'iPhone (iCloud). La majoria del seu correu, un 70%, li arriben al correu de la UPC, mentres que un 20% arriben al gmail i un 10% al iCloud. Dels missatges del correu de la UPC, només l'1% és correu brossa (*spam*). Aquests percentatges són del 2% i del 5% per als correus de gmail i iCloud, respectivament.

(a) Quina és la probabilitat que un missatge, agafat a l'atzar, sigui *spam*?

Solució: Definim els següents esdeveniments: $\mathcal{U}$ ="el correu s'envia a la UPC", $\mathcal{G}$ ="el correu s'envia a gmail", $\mathcal{C}$ ="el correu s'envia a iCloud" i $\mathcal{S}$ ="el correu és un *spam*". Ens demanen la probabilitat $P(\mathcal{S})$ que és, aplicant la fórmula de la probabilitat total, igual a

$$\begin{aligned} P(\mathcal{S}) &= P(\mathcal{S}|\mathcal{U})P(\mathcal{U}) + P(\mathcal{S}|\mathcal{G})P(\mathcal{G}) + P(\mathcal{S}|\mathcal{C})P(\mathcal{C}) \\ &= 0.01 \cdot 0.7 + 0.02 \cdot 0.2 + 0.05 \cdot 0.1 = 0.016 = 1.6\% \end{aligned}$$

(b) Obrim un missatge i és *spam*. Quina és la probabilitat que hagi estat enviat al correu de la UPC?

Solució: Hem de calcular la probabilitat $P(\mathcal{U}|\mathcal{S})$, i ho farem aplicant el teorema de Bayes

$$P(\mathcal{U}|\mathcal{S}) = \frac{P(\mathcal{S}|\mathcal{U})P(\mathcal{U})}{P(\mathcal{S})} = \frac{0.01 \cdot 0.7}{0.016} = 0.4375 = 43.75\%$$

(c) Obrim un missatge. Quina és la probabilitat que hagi estat enviat al correu de la UPC i que no sigui *spam*?

Solució: Hem de calcular la probabilitat $P(\mathcal{U} \cap \bar{\mathcal{S}})$. Sabem que, en general, si A i B són dos esdeveniments qualssevol, tenim que

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

Aleshores,

$$\begin{aligned} P(\bar{\mathcal{S}} \cap \mathcal{U}) &= P(\bar{\mathcal{S}}|\mathcal{U})P(\mathcal{U}) \\ &= (1 - P(\mathcal{S}|\mathcal{U}))\, P(\mathcal{U}) \\ &= (1 - 0.01) \cdot 0.7 = 0.693 = 69.3\% \end{aligned}$$

3. `[20 punts]` Un enquestador necessita completar cinc enquestes amb dones d'entre 25 i 35 anys amb estudis superiors universitaris que estiguin laboralment en actiu, que anomenarem *target*. La probabilitat que una persona-*target* escollida a l'atzar accepti participar en l'enquesta és $p = 0.2$. Sigui X la variable aleatòria discreta que compta el nombre de persones-*target* que no acceptaran participar en l'esquesta abans de completar les cinc enquestes que són necessàries.

(a) Quina distribució de probabilitat segueix la variable X?

Solució: La variable aleatòria X que compta el nombre de persones-*target* que no acceptaran participar en l'enquesta abans de completar les cinc enquestes que són necessàries segueix un model de distribució **binomial negatiu** de paràmetres $r = 5$ i $p = 0.2$, és a dir,

$$X \hookrightarrow \mathcal{BN}(5, 0.2)$$

(b) Quina és la probabilitat que s'hagi hagut de preguntar a un total de 15 persones-*target* per a completar les cinc enquestes?

Solució: Si hem hagut de preguntar a un total de 15 persones-*target* és perquè 10 s'han negat a participar en l'enquesta. Aleshores,

$$\begin{aligned} P(X = 10) &= \binom{10 + 5 - 1}{10} (1 - p)^{10} p^5 \\ &= \binom{14}{10} \cdot 0.8^{10} \cdot 0.2^5 \approx 0.03439 \approx 3.439\% \end{aligned}$$

(c) Quin és el nombre esperat de persones-*target* que han de ser entrevistades per a completar les cinc enquestes que són necessàries.

Solució: El nombre esperat de persones-*target* que és neguen a participar en l'enquesta ve donat per

$$E(X) = \frac{5(1-p)}{p} = \frac{5 \cdot 0.8}{0.2} = 20$$

Per tant, el nombre esperat de persones-*target* que s'ha de preguntar per a completar les cinc enquestes és

$$20 + 5 = 25.$$

4. `[20 punts]` El temps de reacció en segons, R, a un determinat estímul és una variable aleatòria contínua amb la següent funció de densitat:

$$f_R(t) = \begin{cases} \dfrac{k}{t^2}, & 1 \le t \le 3 \\ 0, & \text{altrament} \end{cases}$$

on $k \in \mathbb{R}$.

(a) Quin valor de k fa que $f_R(t)$ sigui realment una funció de densitat?

Solució: Per tal que $f_R(t)$ sigui realment una funció de densitat cal que l'àrea sota la corba definida per $f_R(t)$ sigui 1. En efecte,

$$\int_{-\infty}^{+\infty} f_R(t)dt = \int_1^3 \frac{k}{t^2}dt = \left[\frac{-k}{t}\right]_{t=1}^{t=3} = -\frac{k}{3} + k = \frac{2}{3}k$$

Per tant, si $2k/3 = 1$, aleshores, $k = 3/2$.

(b) Calculeu $E(R)$.

SOLUCIÓ: Aplicant la fórmula de l'esperança, tenim que

$$E(R) = \int_{-\infty}^{+\infty} t f_R(t)dt = \int_1^3 t \cdot \frac{3}{2t^2}dt = \int_1^3 \frac{3}{2t}dt = \left[\frac{3}{2}\ln|x|\right]_{t=1}^{t=3} = \frac{3}{2}\ln(3) \approx 1.6479$$

(c) Calculeu $VAR(R)$.

SOLUCIÓ: Per al càlcul de la variància aplicarem la fórmula de Steiner, és a dir,

$$VAR(R) = E(R^2) - E(R)^2$$

En efecte,

$$E(R^2) = \int_{-\infty}^{+\infty} t^2 f_R(t)dt = \int_1^3 t^2 \cdot \frac{3}{2t^2}dt = \int_1^3 \frac{3}{2}dt = \left[\frac{3}{2}t\right]_{t=1}^{t=3} = \frac{3}{2}(3-1) = 3$$

Per tant,

$$VAR(R) = E(R^2) - E(R)^2 = 3 - \left(\frac{3}{2}\ln(3)\right)^2 = 3 - \frac{9}{4}\ln^2(3) \approx 0.2844$$

5. `[20 punts]` Si la longitud, L de la rosca d'un pern[1] està normalment distribuïda, és a dir,

$$L \hookrightarrow N(\mu, \sigma).$$

(a) Quina és la probabilitat que la longitud de la rosca d'un pern agafat a l'atzar es trobi, com a molt, a una desviació tipus i mitja del valor mitjà? És a dir, $P(\mu - 1.5\sigma \leq L \leq \mu + 1.5\sigma)$.

SOLUCIÓ: Per al càlcul de la probabilitat $P(\mu - 1.5\sigma \leq L \leq \mu + 1.5\sigma)$ tipificarem la variable L. En efecte,

$$\begin{aligned} P(\mu - 1.5\sigma \leq L \leq \mu + 1.5\sigma) &= P(-1.5\sigma \leq L - \mu \leq +1.5\sigma) = P\left(-1.5 \leq \frac{L-\mu}{\sigma} \leq 1.5\right) \\ &= \phi(1.5) - \phi(-1.5) = \phi(1.5) - (1 - \phi(1.5)) \\ &= 2\phi(1.5) - 1 = 2 \cdot 0.93319 - 1 = 0.86638 \end{aligned}$$

(b) Quina és la probabilitat que la longitud de la rosca d'un pern agafat a l'atzar sigui superior a dues desviacions tipus i mitja del valor mitjà? És a dir, $P(|L - \mu| \geq 2.5\sigma)$. Tingueu present que, $P(|X| \geq a)$ és equivalent a $1 - P(|X| < a)$. Alhora, $P(|X| < a)$ és també equivalent a $P(-a < X < a)$.

SOLUCIÓ: Per al càlcul de la probabilitat $P(|L - \mu| \geq 2.5\sigma)$ tipificarem la variable L. En efecte,

$$\begin{aligned} P(|L-\mu| \geq 2.5\sigma) &= 1 - P(|L-\mu| < 2.5\sigma) = 1 - P(-2.5\sigma < L - \mu < 2.5\sigma) \\ &= 1 - P\left(-2.5 < \frac{L-\mu}{\sigma} < 2.5\right) \\ &= 1 - (\phi(2.5) - \phi(-2.5)) \\ &= 1 - (\phi(2.5) - (1 - \phi(2.5))) \\ &= 1 - \phi(2.5) + 1 - \phi(2.5) = 2 - 2\phi(2.5) \\ &= 2 - 2 \cdot 0.99379 = 0.01242 \end{aligned}$$

[1] Peça metàl·lica, normalment d'acer o ferro, llarga, cilíndrica, semblant a un cargol però de dimensions més grans, amb un extrem de cap rodona i un altre extrem que acostuma a ser roscatge. En aquest extrem es cargola una xaveta, femella, o rebló, i permet de subjectar una o més peces a una estructura, per regla general de gran volum.