Solución Examen Final Introducción a la Estadística
13 de Septiembre 2008

Problema 1.

a) La tabla dada muestra las frecuencias absolutas, $n_{i,j}$. Para obtener las frecuencias relativas basta con calcular $f_{i,j} = n_{i,j}/n$:

Cantidad a pagar	Parte correcto	Parte fraudulento
(0, 100)€	0,45	0,15
[100, 300)€	0,20	0,05
[300, 600]€	0,05	0,10

b) Dada la distribución conjunta obtenida en el apartado (a), tenemos que la distribución marginal de la variable “Cantidad a pagar” es:

Cantidad a pagar	f_i
(0, 100) €	0,60
[100, 300) €	0,25
[300, 600] €	0,15

y la distribución marginal de la variable “Tipo de parte” es:

Tipo de parte	f_j
Parte correcto	0,70
Parte fraudulento	0,30

c) La distribución de la variable “Tipo de parte” cuando la “Cantidad a pagar” es menor que 100 €es:

Tipo de parte	$f_{j\|i}$
Parte correcto	$0{,}45/0{,}60 = 0{,}75$
Parte fraudulento	$0{,}15/0{,}60 = 0{,}25$

y la distribución de la variable “Tipo de parte” cuando la “Cantidad a pagar” es mayor que 300 €es:

Tipo de parte	$f_{j\|i}$
Parte correcto	$0{,}05/0{,}15 \approx 0{,}33$
Parte fraudulento	$0{,}10/0{,}15 \approx 0{,}67$

Se observa que la proporción de partes fraudulentos es mayor cuando la cantidad a pagar es mayor o igual que 300 €.

d) Denotemos por X la cantidad a pagar por un parte. La cuantía media a pagar por un parte es:

$$50 \times 0,60 + 200 \times 0,25 + 450 \times 0,15 = 147,5.$$

Problema 2.

1. (a) Sea A el suceso de que el conductor es de alto riesgo y B de bajo riesgo. Escribimos acc si el conductor tiene un accidente. Tenemos:

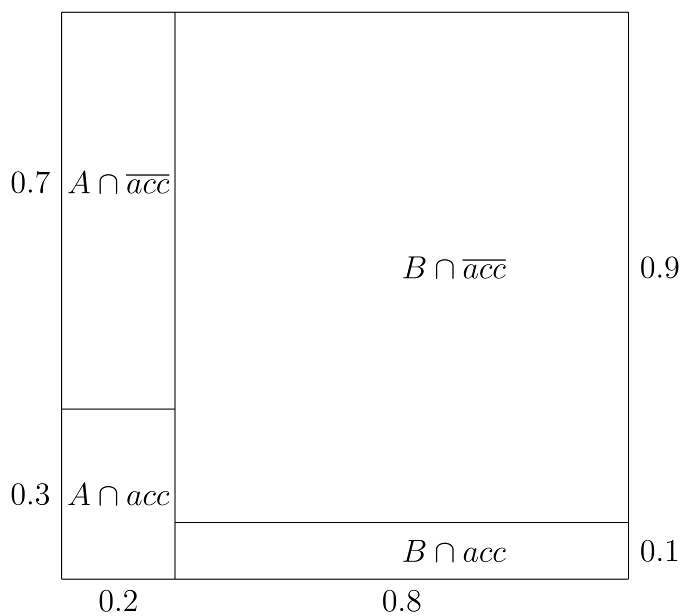
$$P(A) = 0.2, \quad P(B) = 1 - P(A) = 0.8, \quad P(acc|A) = 0.3, P(acc|B) = 0.1$$

Ahora

$$\begin{aligned} P(acc) &= P(acc|A)P(A) + P(acc|B)P(B) \quad \text{utilizando la ley de la probabilidad total} \\ &= 0.3 \times 0.2 + 0.1 \times 0.8 \\ &= 0.14 \end{aligned}$$

[.5]

Otra manera de hacerlo



Ahora $acc = (A \cap acc) \cup (B \cap acc)$ y calculando las áreas se tiene $P(acc) = 0.3 \times 0.2 + 0.1 \times 0.8 = 0.14$.

(b)

$$\begin{aligned} P(A|acc) &= \frac{P(acc|A)P(A)}{P(acc)} \quad \text{por el teorema de Bayes} \\ &= \frac{0.3 \times 0.2}{.14} \\ &= \frac{3}{7} \approx 0.4286 \end{aligned}$$

Igualmente se puede hacer el cálculo a través de las áreas, dividiendo el área de $A \cap acc$ por el área de acc. [.75]

(c) Se tiene

$$\begin{aligned} P(\text{3'era poliza es la primera de un accidentado}) &= P(\overline{acc}, \overline{acc}, acc) \\ &= P(\overline{acc})^2 P(acc) \\ &= 0.86^2 \times 0.14 \\ &= 0.103544 \end{aligned}$$

[.5]

(d) Sea X el número de conductores en el grupo de riesgo alto. Entonces $X \sim \mathcal{B}\left(140, \frac{3}{7}\right)$. Luego

$$E[X] = 140 \times \frac{3}{7} = 60 \quad V[X] = 140 \times \frac{3}{7} \times (1 - \frac{3}{7}) = 34.286$$

$$\begin{aligned} P(X \geq 60) &= P(X > 59) \\ &= P(X > 59.5) \\ &= P\left(\frac{X-60}{\sqrt{34.286}} > \frac{59.5-60}{\sqrt{34.286}}\right) \\ &= P\left(\frac{X-60}{\sqrt{34.286}} > -0.0854\right) \\ &\approx P(Z > -0.0854) \quad \text{donde } Z \sim \mathcal{N}(0,1) \\ &= P(Z < 0.0854) \\ &\approx 0.534 \quad \text{utilizando las tablas} \end{aligned}$$

[.75]

Problema 3.ea $X =$ *demanda diaria de gasoil, en miles de litros.*

a) Para que $f(x)$ sea una función de densidad tiene que verificar que $f(x) \geq 0 \; \forall x \in \mathbb{R}$ y $\int_{-\infty}^{\infty} f(x)\, dx = 1$.
Si $0 \leq x \leq 1$ entonces $x \geq 0$, y si $1 < x \leq 2$ entonces $2 - x \geq 0$, de donde se deduce que $k \geq 0$. Además

$$\begin{aligned} 1 &= \int_{-\infty}^{\infty} f(x)\, dx = \int_0^1 k \cdot x\, dx + \int_1^2 k \cdot (2-x)\, dx \\ &= k \cdot \left[\tfrac{x^2}{2}\right]_0^1 + k \cdot \left[2 \cdot x - \tfrac{x^2}{2}\right]_1^2 \\ &= k \cdot \left(\tfrac{1^2}{2}\right) + k \cdot \left(2 \cdot 2 - \tfrac{2^2}{2} - \left(2 \cdot 1 - \tfrac{1^2}{2}\right)\right) = k \end{aligned}$$

Por tanto $k = 1$.

b)

$$
\begin{aligned}
E[X] &= \int_0^1 x\cdot x\,dx + \int_1^2 (2-x)\cdot x\,dx \\
&= \left[\frac{x^3}{3}\right]_0^1 + \left[2\cdot\frac{x^2}{2} - \frac{x^3}{3}\right]_1^2 \\
&= \left(\frac{1^3}{3}\right) + \left(2\cdot\frac{2^2}{2} - \frac{2^3}{3} - \left(2\cdot\frac{1^2}{2} - \frac{1^3}{3}\right)\right) = 1 \text{ litro}
\end{aligned}
$$

$$
\begin{aligned}
Var[X] &= E[X^2] - E[X]^2 = \int_0^1 x\cdot x^2\,dx + \int_1^2 (2-x)\cdot x^2\,dx \; - 1 \\
&= \left[\frac{x^4}{4}\right]_0^1 + \left[2\cdot\frac{x^3}{3} - \frac{x^4}{4}\right]_1^2 - 1 \\
&= \left(\frac{1^4}{4}\right) + \left(2\cdot\frac{2^3}{3} - \frac{2^4}{4} - \left(2\cdot\frac{1^3}{3} - \frac{1^4}{4}\right)\right) \; - 1 = \frac{1}{6}
\end{aligned}
$$

$$
DT[X] = \sqrt{1/6} = 0.4082 \text{ litros}
$$

c) La probabilidad de que en un día la demanda sea superior a 1500 litros es:

$$
\begin{aligned}
P(X > 1.5) &= \int_{1.5}^{\infty} f(x)\,dx = \int_{1.5}^{2} (2-x)\,dx = \left[2\cdot x - \frac{x^2}{2}\right]_{1.5}^{2} \\
&= 2\cdot 2 - \frac{2^2}{2} - \left(2\cdot 1.5 - \frac{1.5^2}{2}\right) = 0.125.
\end{aligned}
$$

Problema 4. a) Al tratarse de una población normal, una de sus propiedades fundamentales es la simetría, y así media y mediana coinciden, por lo que aqu í: $\mu = 22,32$ años.

Por otro lado, el percentil 33 da el valor exacto de $Z = -0,44$ en la tabla de la normal tipificada $Z \equiv N(0;1)$, por lo que se tiene que: $Z = \frac{X-\mu}{\sigma}$ siempre, y en nuestro caso: $-0,44 = \frac{21-22,32}{\sigma} \Longrightarrow \sigma = \frac{-1,32}{-0,44} = 3$ años $\Longrightarrow \sigma^2 = 9$.

b) Teniéndose ahora una distribución continua de tipo normal de la forma: $X \simeq N(22,32;3)$, y nos piden entonces: $Prob(X > 25) = 1 - Prob(X \leq 25) = 1 - Prob(\frac{X-\mu}{\sigma} \leq \frac{25-22,32}{3}) = 1 - Prob(Z \leq 0,89) = 1 - 0,8133 = 0,1867$; sin más que tipificando y mirando en la tabla.

c) Nos piden ahora la cantidad exacta de alumnos matriculados que tienen entre 18 y 24 años; por lo tanto, como antes, será: $Prob(18 \leq X \leq 24) = Prob(X \leq 24) - Prob(X \leq 18) = Prob(\frac{X-\mu}{\sigma} \leq \frac{24-22,32}{3}) - Prob(\frac{X-\mu}{\sigma} \leq \frac{18-22,32}{3}) = Prob(Z \leq 0,56) - Prob(Z \leq -1,44) = 0,7123 - 0,0749 = 0,6374$.

Es decir, el 63,74 % de todos los alumnos matriculados en dicha facultad tiene entre 18 y 24 años, lo que supone exactamente 800 alumnos.

d)Hay que buscar aquí el valor de la variable Z= N(0;1) que deja por encima de sí el 2,5 % de los datos (es decir: $Z_{0,975}$), y mirando en la tabla correspondiente, se tiene que dicho valor es : $Z_{0,975} = 1,96$,y así: $1,96 = \frac{X-22,32}{3} \Longrightarrow X = 3 * (1,96) + 22,32 = 28,20$ años y será a partir de esta edad, medida en años, que queda por encima el 2,5% de los alumnos más veteranos de la facultad.