Examen de Estadística I. 23-06-2008.

1. X, horas de estudio de un examen.
Y, nota de un examen.
$\overline{x} = 14.42, s_x^2 = 20.33$
$\overline{y} = 6.57, s_y^2 = 4.53, r_{xy} = 0.96$

a) $y^* = a + bx$
$b = \frac{s_{xy}}{s_x^2} = \frac{r_{xy}\, s_x\, s_y}{s_x^2} = \frac{r_{xy}\, s_y}{s_x} = \frac{(0.96)\sqrt{4.53}}{\sqrt{20.33}} = 0.4531,\ a = \overline{y} - b\overline{x} = 6.57 - (0.4531)(14.42) = 0.0354$
$\underline{y^* = 0.0354 + 0.4531\, x}$

b) $x_0 = 20$
$y_0^* = a + b\, x_0 = 0.0354 + (0.4531)(20) = \underline{9.0986}$
$R^2 = r_{xy}^2 = (0.96)^2 = \underline{0.9216}$

c) $s_e^2 = (1 - R^2) * s_y^2 = (1 - 0.9216) * (4.53) = \underline{0.3552}$
$s_{y^*}^2 = R^2 * s_y^2 = (0.9216)(4.53) = \underline{4.1748}$

2. X, importe de una factura emitida por un taller de reparación en euros.
Y, tiempo de un servicio de un taller de reparación en horas.
$\overline{x} = 75, s_x = 25$, Mo(x) = 60, Me(x) = 65
$X = 18\,Y + 10$

a) $\overline{y} = \frac{\overline{x}-10}{18} = \frac{75-10}{18} = \underline{3.6\widehat{1}}$.

b) $\text{Mo(y)} = \frac{\text{Mo(x)}-10}{18} = \frac{60-10}{18} = \underline{2.\widehat{7}}$

c) $g_0(y) = \frac{s_y}{\overline{y}} = \frac{s_x/18}{3.6\widehat{1}} = \frac{25/18}{3.6\widehat{1}} = \underline{0.3846}$

3. Y_t, volumen de comercio electrónico en el período t en millones de euros.
$y_t^* = -66.5 + b\, t$ (t, años; origen, 2001).
$IVE(1T) = 0.8,\ IVE(2T) = 0.9,\ IVE(4T) = 1.3$

a) $s_t^2 = 4, \overline{y} = 1021, s_y^2 = 544969$
$\overline{t} = 3$, ya que la serie comprende ocho años, el período 2001-2007.
$a = \overline{y} - b\overline{t} \Leftrightarrow -66.5 = 1021 - 3b \Leftrightarrow b = \frac{1021+66.5}{3} = 362.5$
$y_7 = -66.5 + (362.5)(7) = \underline{2471}$

b) $y_t = \frac{-66.5}{4} + \frac{362.5}{4^2}\, t = -16.625 + 22.6563\, t$ (t, trimestres; origen, trimestre central 2001).
Del trimestre central de 2001 al tercero de 2008 hay $(7)(4) + 0.5 = 28.5$ trimestres.
$IVE(3T) = 4 - IVE(1T) - IVE(2T) - IVE(4T) = 1$, luego $\hat{y}_{3T08} = y_{28.5} = -16.625 + (22.6563)(28.5) = \underline{629.0781}$, es decir, como no existe componente estacional en el tercer trimestre, la predicción corregida por estacionalidad coincide con el valor de la tendencia en el trimestre.

c) $\hat{y}_5 = 1306.23 = \frac{y_5}{I_{2006}^{2001}/100} = \frac{-66.5+(362.5)(5)}{I_{2006}^{2001}/100} = \frac{174600}{I_{2006}^{2001}} \Leftrightarrow I_{2006}^{2001} = \frac{174600}{1306.23} = \underline{133.6671}$

4. X, duración de un partido de Roland Garros en horas.
$X \sim Un(a=1, b=3),\ f(x) = \frac{1}{b-a},\ x \in [a,b]$

a) $p(X > 2.5) = \int_{2.5}^{3} f(x)\,dx = \int_{2.5}^{3} \frac{1}{2}\,dx = \frac{0.5}{2} = \underline{0.25}$

b) Y, número de encuentros de la jornada, de entre 4, con duración menor de dos horas y media.
$Y \sim Bi(4,\theta),\ f(y) = \left(\begin{array}{c} 4 \\ y \end{array}\right) \theta^y (1-\theta)^{4-y},\ y = 0,1,2,3,4$
$\theta = p(X < 2.5) = 1 - 0.25 = 0.75$
$p(Y=3) = \left(\begin{array}{c} 4 \\ 3 \end{array}\right) (0.75)^3 (0.25)^1 = \underline{0.4219}$

c) $X \sim N(\mu = 2.5, \sigma^2)$
$p(X > 2.5) = \underline{0.5}$, dada la simetría de la distribución Normal con respecto a su media.

5. X, ventas diarias sección 1 en euros.
Y, ventas diarias sección 2 en euros.
$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim N\left(\begin{pmatrix} 800 \\ 1200 \end{pmatrix}, \begin{pmatrix} 200^2 & 0 \\ 0 & 100^2 \end{pmatrix}\right)$$

a) $B = 0.1X + 0.08Y$, beneficio diario global en euros.
$B \sim N(\mu_b = 0.1\mu_x + 0.08\mu_y = (0.1)(800) + (0.08)(1200), \sigma_b^2 = (0.1)^2\sigma_x^2 + (0.08)^2\sigma_y^2 = (0.1)^2(200)^2 + (0.08)^2(100)^2) = \underline{N(176, 464)}$

b) $p(B > 224) = p(Z_b > \frac{224-176}{\sqrt{464}}) = p(Z_b > 2.2283) = \underline{0.0129}$, donde $Z_b = \frac{B-\mu_b}{\sigma_b} \sim N(0,1)$.

c) $D = 0.1X - 0.08Y$, diferencia de beneficios diarios entre las secciones 1 y 2 en euros.
$D \sim N(\mu_d = 0.1\mu_x - 0.08\mu_y = (0.1)(800) - (0.08)(1200) = -16, \sigma_d^2 = (0.1)^2\sigma_x^2 + (0.08)^2\sigma_y^2 = (0.1)^2(200)^2 + (0.08)^2(100)^2 = 464)$
$p(D > 0) = p(Z_d > \frac{0-(-16)}{\sqrt{464}}) = p(Z_d > 0.7428) = 0.2281$, donde $Z_d = \frac{D-\mu_d}{\sigma_d} \sim N(0,1)$. Como es más probable que la diferencia de beneficios considerada sea negativa, $p(D < 0) = 0.7719$, entonces la sección 2 <u>es comercialmente más atractiva</u> que la sección 1.

También se puede resolver en el mismo sentido teniendo en cuenta que el beneficio de la sección 2 tiene <u>mayor media</u> que el de la sección 1, $(0.08)(1200) = 96 > 80 = (0.1)(800)$, pero <u>menor desviación típica</u>, $(0.08)(100) = 8 < 20 = (0.1)(200)$.