Dpto. Matemática Aplicada y Estadística

Grado en IIAA y Grado en IHJ
Asignatura: **Estadística Aplicada. Curso 2011-2012**
Asignatura: Estadística Aplicada. **Problemas del Modelo A.**

PROBLEMA 1.- Mucha gente manifiesta reacciones de alergia sistémica a las picaduras de insectos. Estas reacciones varían de paciente a paciente, no sólo en cuanto a gravedad, sino también en el tiempo transcurrido hasta que se inicia la reacción. Los datos siguientes representan este "tiempo transcurrido (en min.) hasta el inicio de la reacción" en 30 pacientes que experimentaron una reacción sistémica a la picadura de abeja:

3.5	5.9	6.2	7.4	7.5	7.9	8.0	8.5	8.8	9.1
9.1	9.8	9.9	10.1	10.4	10.5	10.9	11.2	11.4	11.4
11.5	11.6	11.7	12.3	12.5	12.7	12.9	13.6	14.7	16.5

Se pide:

1. Realizar el diagrama de caja y bigotes para estos datos. Colocar en cada línea del gráfico su valor numérico. Comentar las características más relevantes del diagrama. Determinar, a partir de los valores observados, entre qué valores se encuentran las observaciones que pueden ser consideradas como no atípicas
2. Agrupar los datos en intervalos de la misma amplitud y representar gráficamente la distribución de frecuencias. Describir las características más relevantes de dicha gráfica. ¿Qué medidas de posición central y de dispersión son más adecuadas para resumir los datos? Razona tu respuesta.
3. Calcular la media y la desviación típica del tiempo transcurrido hasta que se inicia la reaccióna.
4. ¿En qué intervalo se sitúa la mediana? ¿Cuál es el intervalo modal?
5. ¿Cuál es el porcentaje de valores de la muestra cuya tiempo transcurrido hasta que se inicia la reacción es superior a 11.5 minutos?

PROBLEMA 2.- En una cadena de librerías se estudia la relación entre el número de libros vendidos en cientos (l) y el precio correspondiente en cientos de euros (p). Observadas las características durante un mes, se obtuvieron los siguientes valores muestrales:

$$\bar{l} = 5 \quad s_p = 1.87$$

Como conclusión de estudio se establece el siguiente modelo lineal:

$$l = 10.4 - 1.8 * p$$

1. Obtener la recta de regresión del precio (p) en función de la cantidad de libros vendidos (l), sabiendo que el coeficiente de determinación del modelo es $R^2 = 0.81$.
2. ¿Sería posible que en este modelo el coeficiente de correlación fuese positivo? Justificar la respuesta.
3. Sabiendo que la cantidad de libros vendidos en un mes es de 700 ejemplares, ¿cuál sería el precio estimado por el modelo?

PROBLEMA 3.- Los ladrillos fabricados por una empresa pueden tener dos tipos de defectos D1 y D2. El defecto D1 ocurre con una probabilidad de 0.07, el defecto D2 con una probabilidad de 0.06 y en el 2% de los ladrillos ocurren ambos tipos de defectos. Se pide:

1. Calcular la probabilidad de que al seleccionar un ladrillo al azar de la producción no tenga ningún tipo de defecto.

2. Calcular la probabilidad de que el ladrillo selecionado tenga exactamente un único tipo de defecto.

3. Cada ladrillo se somete de manera automática a un test de ruptura, con las siguientes probabilidades:

   - Si el ladrillo tiene alguno de los defectos tiene una probabilidad de romperse del 0.85
   - Si el ladrillo no tiene ninguno de los defectos tiene una probabilidad de romperse de 0.06.

   Se pide:

   3.1. Calcular la probabilidad de que un ladrillo seleccionado al azar se rompa durante el test.

   3.2. Si un ladrillo escogido al azar no se ha roto durante el test, ¿cuál es la probabilidad de que sea defectuoso?

PRUEBA ESCRITA 1: **MODELO A**

PROBLEMA 1:

$X \equiv$ Tiempo transcurrido (en min.) hasta el ini

de la reacción.

$n = 30$

1: $$Me = \frac{x_{(15)} + x_{(16)}}{2} = \frac{10.4 + 10.5}{2} = 10.45$$

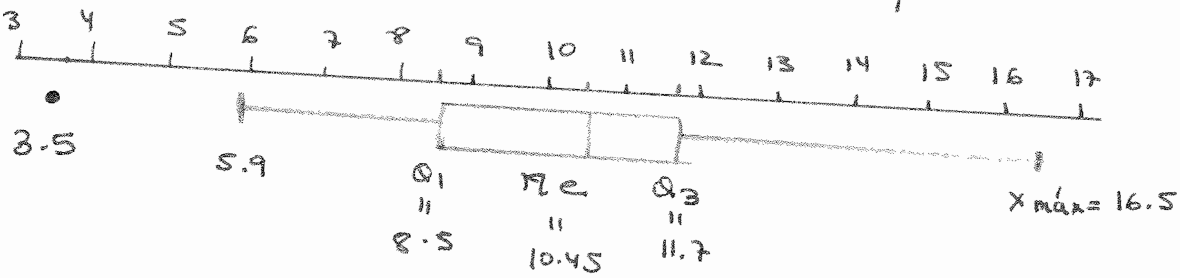
$Q_1 = 8.5$

$Q_3 = 11.7$ | $RIQ = 3.2$

$$\begin{cases} L_{INF} = 8.5 - 1.5 * 3.2 = 3.7 \\ L_{SUP} = 11.7 + 1.5 * 3.2 = 16.5 \end{cases}$$

$\Rightarrow$ $x_i \notin [3.7, 16.5]$ son datos atípicos

$\Rightarrow$ $x = 3.5$ es un dato atípico.



→ $\exists$ un dato atípico.

→ Mayor dispersión a partir de Q_3

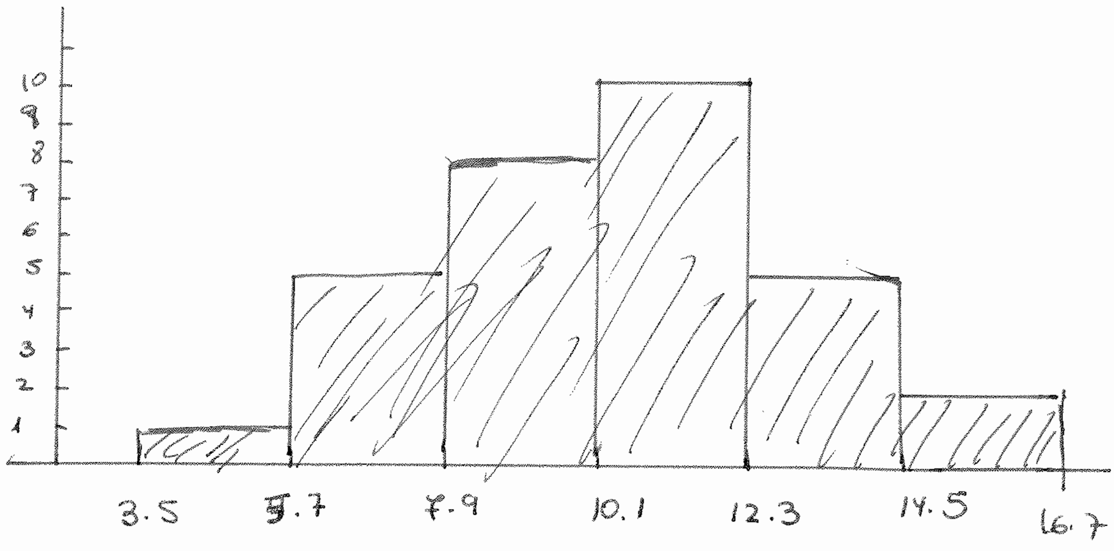
→ Las observaciones no atípicas están en el intervalo $[3.7, 16.5]$

(2) - $n = 30 \rightarrow \bullet\ K \geq \sqrt{30} = 5.5 \Rightarrow K = 6$

$$\bullet\ h = \frac{16.5 - 3.5}{6} = \frac{13}{6} = 2.17 \cong 2.2$$

límites	mi	ni	fi	Ni
[3.5, 5.7)	4.6	1	0.03	1
[5.7, 7.9)	6.8	4	0.13	5
[7.9, 10.1)	9	8	0.27	13
[10.1, 12.3)	11.2	10	0.33	23
[12.3, 14.5)	13.4	5	0.17	28
[14.5, 16.7]	15.6	2	0.07	30
TOTALES	—	30	1	—

$X > 11.5$



→ Asimétrico a la izda

→ unimodal

→ No parece existir datos atípicos pero en el apartado anterior hemos comprobado que existe un dato atípico

⇒ Utilizaremos la Me como medida de cent
y el RIQ como medida de dispersión

③ - Con DATOS AGRUPADOS:

$$\bar{x} = \frac{1}{30}\left[4.6*1 + 6.8*4 + 9*8 + 11.2*10 + 13.4*5 + 15.6*2\right] =$$

$$= \frac{314}{30} = \boxed{10.47 \text{ min.}}$$

$$S_x^2 = \frac{n}{n-1}\left(\overline{x^2} - \bar{x}^2\right) =$$

$$= \frac{30}{29}\left[\frac{4.6^2*1 + 6.8^2*4 + 9^2*8 + 11.2^2*10 + 13.4^2*5 + 15.6^2*2}{30} - (10.47)^2\right] = 7.04872 \Rightarrow S_x = \boxed{2.65 \text{ min}}$$

Con datos NO AGRUPADOS:

$$\bar{x} = \frac{3.5 + 5.9 + \dots + 13.6 + 14.7 + 16.5}{30} = \boxed{10.25 \text{ min.}}$$

$$S_x^2 = \frac{30}{29}\left(\frac{3.5^2 + 5.9^2 + \dots + 14.7^2 + 16.5^2}{30} - (10.25)^2\right) =$$

$$= 7.49 \Rightarrow S_x = \boxed{2.7369 \text{ min.}}$$

(4): Int. mediano = [10.1, 12.3)

Int. modal = [10.1, 12.3)

(5): ¿% de $X > 11.5$?

Con datos AGRUPADOS:

% de $X > 11.5 = \frac{7}{30} * 100\% = 23.33\%$

Con datos NO AGRUPADOS:

% de $X > 11.5 = \frac{9}{30} * 100\% = 30\%$

PROBLEMA 2:

$l \equiv$ Nº de libros vendidos (en 10^2)

$p \equiv$ precio (en 10^2 €)

$\leadsto \begin{cases} \bar{l} = 5 \\ s_p = 1.87 \end{cases} \Rightarrow \begin{array}{l} l = 10.4 - 1.8 * p \\ R^2 = 0.81 \end{array}$

$$-1.8 = \frac{s_{l,p}}{s^2_p} \Rightarrow \boxed{s_{l,p} = -1.8 * (1.87)^2 = -6.2944}$$

$$R^2 = 0.81 = \frac{s^2_{l,p}}{s_l^2 * s_p^2} = \frac{(-6.2944)^2}{s_l^2 * (1.87)^2}$$

$$\Rightarrow \boxed{s_l^2 = \frac{(-6.2944)^2}{0.81 * (1.87)^2} = 13.9876}$$

(a): Nos piden $p = f(l) = \alpha + \beta l$

$$\boxed{\hat{\beta} = \frac{s_{p,l}}{s^2_l} = \frac{-6.2944}{13.9876} = -0.45}$$

$$\boxed{\hat{\alpha} = \bar{p} - \hat{\beta} * \bar{l} = 3 - (-0.45) * 5 = 5.25}$$

¿$\bar{p}$? Como $10.4 = \bar{l} - (-1.8) * \bar{p} = 5 + 1.8 * \bar{p}$

$$\Rightarrow \bar{p} = \frac{10.4 - 5}{1.8} = 3$$

$$\Rightarrow p = 5.25 - 0.45 * l$$

(b): $r = \dfrac{s_{p,l}}{s_p * s_l} = \dfrac{-6.2944}{+} < 0$

$\Rightarrow$ sería negativo ya que la covarianza lo es.

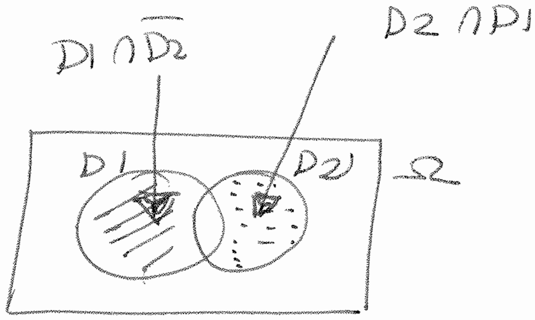
(c): si $e = 7 * 10^2 \Rightarrow \boxed{\hat{p} = 5.25 - 0.45 * 7 = = 2.1 \ (* 10^2 \ €)}$

PROBLEMA 3:

$D1 \longrightarrow p(D1) = 0.07$

$D2 \longrightarrow p(D2) = 0.06$

$p(D1 \cap D2) = 0.02$



(a): $p(\overline{D1} \cap \overline{D2}) = p(\overline{D1 \cup D2}) = 1 - p(D1 \cap D2) =$

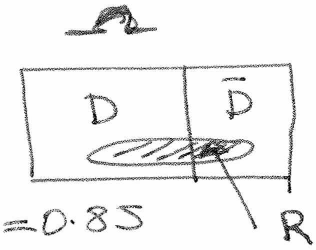
$= 1 - [p(D1) + p(D2) - p(D1 \cap D2)] =$

$= 1 - (0.07 + 0.06 - 0.02) = \boxed{0.89}$

(b): $p[(\overline{D1} \cap D2) \cup (D1 \cap \overline{D2})] = p(D1 \cup D2) - p(D1 \cap D2) =$

$= 0.11 - 0.02 = \boxed{0.09}$

(c): D ≡ ser defectuoso $\to p(D) = 0.11$, $p(\overline{D}) = 0.89$



R ≡ ladrillo se rompe $\to p(R/D) = 0.85$, $p(R/\overline{D}) = 0.06$

(c.1): $p(R) \overset{T^a \text{ Prob. Total}}{=} p(D) * p(R/D) + p(\overline{D}) * p(R/\overline{D}) =$

$= 0.11 * 0.85 + 0.89 * 0.06 = \boxed{0.1469}$

(c.2): $p(D/\overline{R}) \overset{T^a \text{ Bayes}}{=} \dfrac{p(D)\,p(\overline{R}/D)}{p(\overline{R})} = \dfrac{0.11 * (1 - 0.85)}{1 - 0.1469} =$

$= \boxed{0.0193}$