Dpto. Matemática Aplicada y Estadística

Grado en IIAA y Grado en IHJ
Asignatura: **Estadística Aplicada. Curso 2013-2014**
FEBRERO 2014

NOMBRE:..**APELLIDOS:**..
ESPECIALIDAD:...

1. **[0.5 puntos]** Cierto establecimiento de un centro comercial está estudiando la relación existente entre el número de clientes que diariamente entran en el mismo (x) y las ventas realizadas (y, en cientos de euros). Para ello, durante 25 días seleccionados al azar de los últimos dos meses se han observado ambas variables obteniendo la siguiente información muestral:

$$\sum x_i = 3950 \qquad \sum x_i^2 = 659500 \qquad \sum y_i = 2700 \qquad \sum y_i^2 = 306000 \qquad \sum x_i \cdot y_i = 448000$$

   A partir de estos datos se pide:

   (a) Determinar la recta de regresión de mínimos cuadrados que permite explicar las ventas del comercio en función del número de clientes que entran en el mismo al día. Dar una medida de la bondad del ajuste realizado.

   (b) Si aumenta el número de clientes que diariamente entran en el establecimiento, ¿cual es el efecto en las ventas realizadas? Justificar la respuesta en función del resultado obtenido.

2. **[0.75 puntos]** En una determinada cooperativa se comercializa la producción de brócoli de la zona. Teniendo en cuenta el ciclo de formación de la pella desde la siembra a la madurez, se dividen en las variedades temprana, intermedia y tardía. El 38% de la producción es de la variedad temprana y la producción de la variedad tardía es un cuarto de la producción de la variedad intermedia. Por experiencias previas, se sabe que la probabilidad de que en la variedad temprana aparezcan manchas en las hojas de la planta es del 0.10 y la probabilidad de que en la variedad intermedia no aparezcan manchas en las hojas de la planta es del 85%. Además, se sabe que el 3.72% de la producción es de la variedad tardía y presenta manchas en las hojas de la planta. Se pide:

   (a) Calcular la probabilidad de que elegida una planta de brócoli al azar no presente manchas en las hojas de la planta.

   (b) Si una planta de brócoli presenta manchas en las hojas, determinar la probabilidad de que sea de la variedad temprana o tardía.

   (c) Determinar la probabilidad de que la planta de brócoli sea de la variedad temprana o tardía y no presente manchas en las hojas.

3. **[1 punto]** El tiempo de vida de la batería de una cosechadora se comporta como una variable aleatoria X con función de densidad:

$$f(x) = \begin{cases} k(3-x)^2x^2, & \text{si } x \in (0,3) \\ 0, & \text{en otro caso} \end{cases}$$

   donde x representa los años. Se pìde:

   (a) Determinar el valor de k para que $f(x)$ sea verdaderamente una función de densidad. **(0.25 puntos)**

   (b) Determinar la función de distribución de la v.a. X. **(0.5 puntos)**

   (c) Determinar la probabilidad de que la batería dure al menos 2 años. **(0.25 puntos)**

4. [1.**25 puntos**] En una línea de producción, en ocasiones es necesario interrumpir el proceso productivo debido a fallos de la maquinaria, de los operarios, de la mataria prima, etc. Se sabe que el número de interrupciones del proceso productivo que se da a lo largo de una semana (5 días) sigue una distribución de Poisson de media 4 y, por lo tanto, el tiempo (en días) entre interrupciones consecutivas es una exponencial de media igual a 1.25. Se pide:

   (a) Probabilidad de que en una semana se produzca más de una interrupción del proceso productivo. **(0.25 puntos)**

   (b) Calcular el número medio de interrupciones que se produce del proceso productivo a lo largo de un mes (4 semanas). **(0.25 puntos)**

   (c) Probabilidad de que el tiempo entre dos interrupciones consecutivas sea mayor a 4 días. (**0.25 puntos**)

   (d) Si han pasado 2 días y el proceso no se ha interrumpido todavía, ¿cuál es la probabilidad de tener que esperar al menos otros 2 días más para que se produzca la nueva parada del proceso productivo? **(0.5 puntos)**

5. [**1.25 puntos**] El diámetro del grano de maíz de cierta variedad se distribuye aleatoriamente con media igua a 9 mm. y con desviación típica igual a 1.2 mm.

   (a) ¿Podrías dar una cota de la probabilidad de que el diámetro del grano de maíz no se encuentre en el intervalo [7.32,10.68]. Razonar la respuesta.(**0.25 puntos**)

   (b) Supongamos a partir de ahora, que el diámetro del grano de maíz de esta variedad se distribuye según una variable aleatoria normal de media 9 mm. y de desviación típica 1.2 mm.

      (b1) Determinar la probabilidad de que el grano de maíz de esta variedad pase a través de un tamiz de malla cuyo diámetro es igual a 8.5 mm.(**0.25 puntos**)

      (b2) Si por medio de dicho tamiz zarandeamos 2000 granos de maíz elegidos al azar, determinar la probabilidad de que pasen a traves del tamiz al menos 650 granos.(**0.5 puntos**)

      (b3) ¿Cuánto tiene que valer el diámetro de la malla del tamiz para que al menos retenga el 90% de los granos de esta variedad de maíz? (**0.25 puntos**)

6. [**1.25 puntos**] El espárrago es una planta perenne cuyo cultivo comercial puede tener una duración de 15 años y su implantación es costosa. Dada la extensión del sistema radicular, la profundidad del suelo es fundamental, considerándose indispensable contar con un promedio superior a 80 cm. de sustrato permeable. Para validar la aptitud de cierto campo de cultivo, se realizan 32 determinaciones de la profundidad del sustrato permeable (en cm.) en puntos tomados al azar de dicho campo obteniéndose una profundidad promedio del sustrato permeable en esas 32 observaciones del campo igual a 85.43 cm. De la experiencia previa con este tipo de campos de cultivo, podemos asumir que la desviación típica asociada a la variable "profundidad del sustrato permeable de dicho campo" es igual a $\sigma = 10.5$ cm. Se pide:

   (a) Construir de manera detallada un intervalo de confianza al 99% para la profundidad promedio del sustrato permeable de dicho campo. Indicar todos los resultados teóricos necesarios para obtener el citado intervalo. **(0.5 ptos.)**

   (b) A partir de los datos muestrales, ¿podemos asumir al 99% que dicho campo es apto para el cultivo del espárrago, esto es, podemos concluir que la profundidad promedio del sustrato permeable de dicho campo es significativamente mayor a 80 cm.? Plantear y llevar a cabo el contraste adecuado para responder a esta pregunta. Calcular el $p - valor$ de la prueba.**(0.5 ptos.)**

   (c) Determinar qué tamaño de la muestra será necesario utilizar si queremos estimar la profundidad promedio del sustrato permeable de dicho campo con un error menor a 3 cm. y garantizando una confianza del 99%. **(0.25 ptos.)**

Grado de Agrícolas: FEBRERO 2014

1/11

P.1:

(a): $X \equiv$ Nº de clientes

$Y \equiv$ Ventas en 10^2 €

$y = b + aX$

$n = 25 \rightarrow$

$\sum x_i = 3950 \rightarrow \bar{x} = \frac{3950}{25} = 158$

$\sum y_i = 2700 \rightarrow \bar{y} = \frac{2700}{25} = 108$

$\sum x_i^2 = 659500 \rightarrow sx^2 = \frac{25}{24}\left(\frac{659500}{25} - 158^2\right) = 1475$

$\sum y_i^2 = 306000 \rightarrow sy^2 = \frac{25}{24}\left(\frac{306000}{25} - 108^2\right) = 600$

$\sum x_i y_i = 448000 \rightarrow sxy = \frac{25}{24}\left(\frac{448000}{25} - 158 * 108\right) =$

$= 891.67$

$$\Rightarrow \quad \hat{a} = \frac{sxy}{sx^2} = \frac{891.67}{1475} = 0.6045$$

$$\hat{b} = \bar{y} - \hat{a} \cdot \bar{x} = 108 - 0.6045 * 158 = 12.4859$$

De donde, $\boxed{y = 12.4859 + 0.6045 * X}$

$$R^2 = \frac{sxy^2}{sx^2 \cdot sy^2} = \frac{891.67^2}{1475 * 600} = 0.898$$ Ajuste bueno.

(b): si $\uparrow x \Rightarrow$ Aumentan las ventas ya que la pendiente es positiva, lo que implica que la dependencia entre X e Y lo es.

P.2:

$$\begin{cases} TE \rightsquigarrow p(TE) = 0.38 \\ I \rightsquigarrow p(I) = \alpha \\ TA \rightsquigarrow p(TA) = \beta \end{cases}$$

2/11

$$\left.\begin{array}{l} 0.38+\alpha+\beta = 1 \\ \beta = \frac{1}{4}\alpha \end{array}\right\} \Rightarrow$$

$$\Rightarrow \quad 0.38 + \alpha + \frac{1}{4}\alpha = 1 \Rightarrow \frac{5}{4}\alpha = 0.62 \Rightarrow \alpha = \frac{0.62 * 4}{5} =$$

$$= 0.496 \quad \Rightarrow \quad \beta = 0.124$$

$$\Rightarrow \left.\begin{array}{l} p(TE) = 0.38 \\ p(I) = 0.496 \\ p(TA) = 0.124 \end{array}\right\}$$

$M \equiv$ Aparecer manchas en las hojas de la planta

$$\Rightarrow \begin{cases} \bullet\, p(M \mid TE) = 0.10 \\ \bullet\, p(\overline{M} \mid I) = 0.85 \longleftrightarrow p(M \mid I) = 0.15 \\ p(TA \cap M) = 0.0372 = p(TA)\, p(M \mid TA) \\ \Rightarrow \bullet\, p(M \mid TA) = \frac{0.0372}{0.124} = 0.30 \end{cases}$$

(a): $p(\overline{M}) = 1 - p(M) = 1 - 0.1496 = \boxed{0.8504}$

$$p(M) = p(TE)\, p(M \mid TE) + p(I)\, p(M \mid I) + p(TA)\, p(M \mid TA) =$$

$$= 0.38 * 0.10 + 0.496 * 0.15 + 0.124 * 0.30 =$$

$$= 0.038 + 0.0744 + 0.0372 = 0.1496$$

(b): $p(TE \cup TA \mid \bar{M}) = \dfrac{p[(TE \cup TA) \cap \bar{M}]}{p(\bar{M})} =$ 3/11

$= \dfrac{p[(TE \cap \bar{M}) \cup (TA \cap \bar{M})]}{p(\bar{M})} = \dfrac{p(TE)\, p(\bar{M}/TE) + p(TA)\, p(\bar{M}/TA)}{p(\bar{M})}$

$= \dfrac{0.38 * 0.10 + 0.124 * 0.30}{0.1496} = \dfrac{0.038 + 0.0372}{0.1496} = \dfrac{0.0752}{0.1496} =$

$= \boxed{0.5027}$

(c): $p[(TE \cup TA) \cap \bar{M}] = p[(TE \cap \bar{M}) \cup (TA \cap \bar{M})] =$

$= p(TE \cap \bar{M}) + p(TA \cap \bar{M}) = p(TE)\, p(\bar{M} \mid TE) + p(TA)\, p(\bar{M} \mid TA) =$

$= 0.38 * 0.90 + 0.124 * 0.70 = 0.342 + 0.0868 =$

$= \boxed{0.4288}$

P.3: $X \equiv$ Tiempo de vida de la batería

$$f(x) = \begin{cases} K\underbrace{(3-x)^2 x^2}_{= x^4 - 6x^3 + 9x^2}, & x \in (0,3) \\ 0, & \text{———} \end{cases}$$

(a): ▸ $f(x) = K\underbrace{(3-x)^2 x^2}_{\geq 0 \text{ en todo } \mathbb{R}\text{, en particular en } (0,3)} \geq 0 \Leftrightarrow K \geq 0.$

▸ $1 = \int_{\mathbb{R}} f(x)\,dx = K \int_0^3 (x^4 - 6x^3 + 9x^2)\,dx =$

$$= k\left(\frac{x^5}{5}\Big|_0^3 - \cancel{6}^{3} \frac{x^4}{\cancel{4}_2}\Big|_0^3 + \cancel{9}^{3} \frac{x^3}{\cancel{3}}\Big|_0^3 \right) =$$

4/11

$$= k\left(\frac{3^5}{5} - 3\cdot\frac{3^4}{2} + 3\cdot 3^3\right) = k\, 3^4 \underbrace{\left(\frac{3}{5} - \frac{3}{2} + 1\right)}_{\frac{6-15+10}{10} = \frac{16-15}{10} = \frac{1}{10}} =$$

$$= \frac{81 k}{10} \Leftrightarrow \boxed{k = \frac{10}{81}}$$

$$\Rightarrow \boxed{f(x) = \begin{cases} \frac{10}{81}(x^4 - 6x^3 + 9x^2), & x \in (0,3) \\ 0, & \text{———} \end{cases}}$$

(b): $\forall x \in \mathbb{R},\ F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt$

Si $x \leq 0 \Rightarrow F(x) = 0$

Si $0 < x < 3 \Rightarrow F(x) = \frac{10}{81}\int_0^x (t^4 - 6t^3 + 9t^2)\,dt =$

$$= \frac{10}{81}\left[\frac{t^5}{5}\Big|_0^x - \cancel{6}^{3} \frac{t^4}{\cancel{4}_2}\Big|_0^x + \cancel{9}^{3} \frac{t^3}{\cancel{3}}\Big|_0^x \right] =$$

$$= \frac{10}{81}\left(\frac{x^5}{5} - \frac{3x^4}{2} + 3x^3\right)$$

Si $x \geq 3 \Rightarrow F(x) = 1$

$$\Rightarrow \boxed{F(x) = \begin{cases} 0, & x \leq 0 \\ \frac{10}{81}\left(\frac{x^5}{5} - \frac{3x^4}{2} + 3x^3\right), & 0 < x < 3 \\ 1, & x \geq 3 \end{cases}}$$

(b): $P(X \geq 2) = 1 - P(X < 2) = 1 - F(2) =$

$$= 1 - \frac{10}{81}\left(\frac{2^5}{5} - \frac{3\cdot 2^{\cancel{4}\,3}}{\cancel{4}} + 3\,2^3\right) =$$

$$= 1 - \frac{10}{81}\,2^3\left(\frac{2^2}{5} - \cancel{3} + \cancel{3}\right) = 1 - \frac{\cancel{10}^{\,2}}{81}\,\frac{2^5}{\cancel{5}} =$$

$$= 1 - \frac{2^6}{81} = 1 - \frac{64}{81} = \frac{17}{81} = \boxed{0.2099}$$

P.4: • $X \equiv$ Nº de interrupciones a lo largo de $\underbrace{5 \text{ días}}_{=1 \text{ semana}}$ $\sim Po(\lambda = 4)$

$$\rightarrow f_X(x) = e^{-4}\,\frac{4^x}{x!}, \quad x = 0, 1, 2, 3, 4, \ldots$$

• $T \equiv$ Tiempo, en días, entre interrupciones consecutivas $\sim Exp(\lambda = 0.8)$

$$E(T) = \frac{1}{\lambda} = 1.25 \Rightarrow \lambda = \frac{1}{1.25} = 0.8$$

$$\rightarrow f_T(t) = \begin{cases} 0, & t \leq 0 \\ 0.8\,e^{-0.8t}, & t > 0 \end{cases} = \text{f. densidad de } T$$

y

$$\underbrace{F_T(t) = \begin{cases} 0, & t \leq 0 \\ 1 - e^{-0.8t}, & t > 0 \end{cases}}_{\text{f. DISTRIBUCIÓN de } T}$$

(a): $P(X > 1) = P(X \geq 2) = 1 - P(X \leq 1) =$

$= 1 - e^{-4}\left(\frac{4^0}{0!} + \frac{4^1}{1!}\right) = 1 - 5e^{-4} = \boxed{0.9084}$

(b): $Y \equiv$ Nº de interrupciones a lo largo de 4 semanas $= X_1 + X_2 + X_3 + X_4 \sim Po(\lambda = 4+4+4+4 = 16)$

con $X_i \sim P(\lambda = 4)$, $i = 1, 2, 3, 4$

$\Rightarrow \boxed{E(Y) = 16 \text{ interrupciones}}$

(c): $P(T > 4) = 1 - P(T \leq 4) = 1 - F_T(4) =$

$= 1 - \left(1 - e^{-0.8*4}\right) = e^{-3.2} = \boxed{0.0408}$

(d): $P(T \quad 4 / T > 2) = \frac{P[(T \geq 4) \cap (T \geq 2)]}{P(T > 2)} =$

$= \frac{P(T \geq 4)}{P(T > 2)} = \frac{1 - P(T < 4)}{1 - P(T \leq 2)} = \frac{1 - F_T(4)}{1 - F_T(2)} =$

$= \frac{1 - (1 - e^{-0.8*4})}{1 - (1 - e^{-0.8*2})} = \frac{e^{-3.2}}{e^{-1.6}} = e^{-3.2+1.6} = e^{-1.6} =$

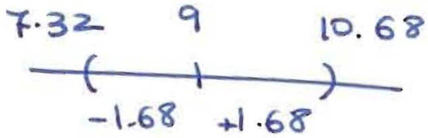
$= \boxed{0.2019}$

P.5: $X \equiv$ Diámetro del grano de maíz 7/1

$E(X) = \mu = 9$ mm.

$\sigma_X \equiv \sigma = 1.2$ mm.

(a): $p\left(X \notin [7.32,\ 10.68]\right) = p\left(|X-9| > 1.68\right) =$



$\leq p\left(|X-9| \geq \underbrace{1.68}\right) \leq \dfrac{1}{(1.4)^2} = \boxed{0.5102}$

$\left(\begin{array}{l} 1.68 = K * \sigma = K * 1.2 \\ \Rightarrow K = \dfrac{1.68}{1.2} = 1.4 \end{array}\right)$

Aplicamos la D. de Tchebychev pues no conocemos la dist. de prob. de la v.a. X

(b): Sup $X \sim N(\mu = 9,\ \sigma = 1.2)$

(b.1): $p($pase a través del tamiz$) =$

$= p(X \leq 8.5) = \phi\left(\dfrac{8.5-9}{1.2}\right) = \phi(-0.42) =$

$= 1 - \phi(0.42) = 1 - 0.6628 = \boxed{0.3372}$

(b.2): $Y \equiv$ Nº de granos que pasan a través del tamiz en un total de 2000 granos $\sim B(n=2000, p=0.3372)$ 8/11

$P(Y \geq 650) =$

$np = 674.4$

$np(1-p) = 446.99232 \rightarrow \sigma_W = 21.14$

Si $W \sim N(674.4, \sigma_W = 21.14)$, entonces $Y \approx W$

$\cong P(W \geq 650 - 0.5) = 1 - \phi\left(\frac{649.5 - 674.4}{21.14}\right) =$

$= 1 - \phi(-1.18) = \phi(1.18) = \boxed{0.8810}$

(b.3): ¿ α / $P(X > \alpha) = 0.90$?

$0.90 = P(X > \alpha) = 1 - \phi\left(\frac{\alpha - 9}{1.2}\right) \Rightarrow \phi\left(\frac{\alpha-9}{1.2}\right) = 0.10$

$\Rightarrow \frac{\alpha - 9}{1.2} = Z_{0.10} = -Z_{0.90} = -1.29$

$\Rightarrow \alpha = 9 - 1.2 * 1.29 = \boxed{7.456 \text{ mm.}}$

P.6: $X \equiv$ Profundidad del sustrato permeable de dicho campo $\leadsto E(X) = \mu$ y $\sigma_X = 10.5$ 9/

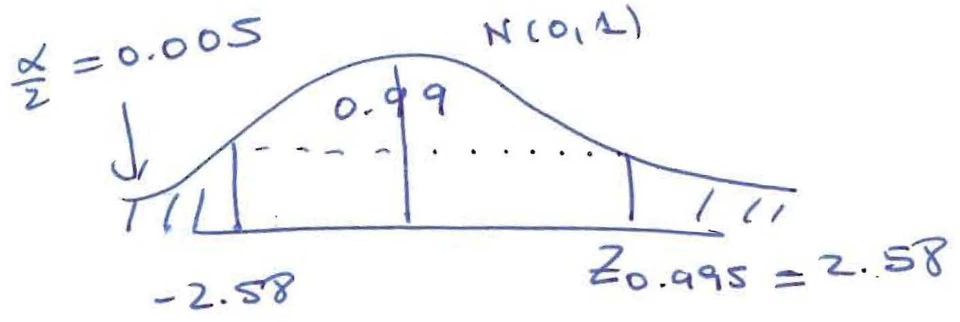
(Datos muestrales: $n = 32 \leadsto \bar{x} = 85.43$)

(a): IC al 99% para μ:

Tomamos $\bar{X} \underset{n \geq 30}{\approx} N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$ aplicando el T.ª Central del límite ya que no conocemos la dist. de prob. de la v.a. X (notar que $n \geq 30$).

$$\Rightarrow \quad Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \underset{n \geq 30}{\approx} N(0,1)$$

En la $N(0,1)$ tomamos los puntos:



De donde, $0.99 = p\left(-2.58 \leq \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq 2.58\right) =$

Dejando adecuadamente μ

$$= p\left(\bar{X} - 2.58 * \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + 2.58 * \frac{\sigma}{\sqrt{n}}\right)$$

$\Rightarrow$ El IC al 99% para μ quedaría como sigue $\bar{X} \pm 2.58 * \frac{\sigma}{\sqrt{n}}$

10/11

Para nuestra muestra buscaría la fórmula:

$$\bar{X} \pm 2.58 * \frac{\sigma}{\sqrt{32}} = 85.43 \pm \underbrace{2.58 * \frac{10.5}{\sqrt{32}}}_{= 4.79} =$$

$$\boxed{= (80.64,\ 90.22)}$$

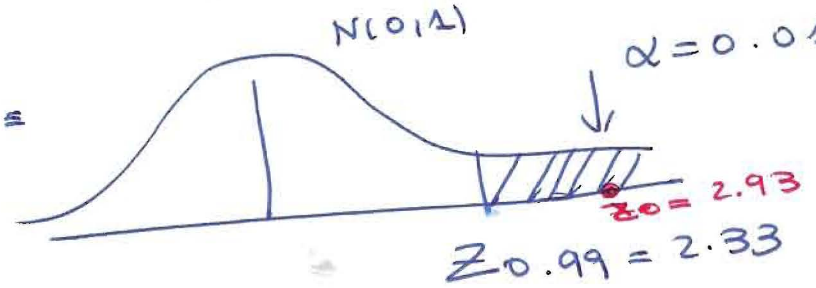
(b) :

$$\left.\begin{array}{l} H_0: \mu = 80 \\ H_1: \mu > 80 \end{array}\right]$$

- $\alpha = 0.01$
- Est. del contraste bajo H_0.

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} \underset{n \geq 32}{\approx} N(0,1)$$

- $R =$



- $Z_0 = \frac{85.43 - 80}{10.5 / \sqrt{32}} = 2.93 \in R \Rightarrow$ Rechazamos H_0 al 99% de confianza

$\Rightarrow$ El sustrato permeable promedio es significativamente mayor a 80cm, esto es, el suelo es apto

11/1

- p-valor $= p(Z > 2.93) = 1 - \phi(2.93) =$

$= 1 - 0.9983 = \boxed{0.0017}$ muy pequeño, lo que confirma el rechazo de H_0.

(c): ¿n / margen de error < 3 al 99%

$$\Rightarrow Z_{1-\frac{\alpha}{2}} * \frac{\sigma}{\sqrt{n}} < 3 \Rightarrow Z_{0.995} * \frac{\sigma}{\sqrt{n}} < 3$$

$$\Rightarrow 2.58 * \frac{10.5}{\sqrt{n}} < 3 \Rightarrow n > \underbrace{\left(\frac{2.58 * 10.5}{3}\right)^2}_{81.54}$$

$\Rightarrow$ Al menos necesitamos $\boxed{\text{82 observaciones}}$.