Dpto. Matemática Aplicada y Estadística

Grado en IIAA y Grado en IHJ
Asignatura: **Estadística Aplicada. Curso 2012-2013**
SEPTIEMBRE 2013

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

1. [**0.5 puntos**] Un ingeniero estudia la relación existente entre la proliferación de una cierta especie de planta silvestre (Y) y la cantidad de lluvia caída (X). Dispone de los valores de Y para 30 valores de X, y obtiene las medidas descriptivas siguientes:

$$\overline{x} = 417.7 \qquad \overline{y} = 1160.6$$
$$s_x = 154.8 \qquad s_y = 395.1$$

Al final del estudio se decide ajustar una recta de regresión de Y sobre X con la ayuda del programa Rcmdr:

$$Y = 2.52 * X + b$$

   (a) Determinar el valor de la ordenada en el origen, esto es, del parámetro b. (**0.10 puntos**)
   (b) Calcular el coeficiente de determinación y comentar la bondad del ajuste realizado (**0.15 puntos**)
   (c) Determinar la ecuación de la recta de regresión de Y sobre X si se le suma 5 a todos los valores de X. (**0.25 puntos**)

2. [**0.75 puntos**] En una determinada explotación agraria de 7000 hectáreas se cultivan 3 tipos de frutales. En 3000 hectáreas se cultivan melocotones, en 2500 hectáreas se cultivan albaricoques y en el resto se cultivan ciruelas. Por experiencias previas, se sabe que aproximadamente el 5% de los melocotones se encuentra en mal estado, el 4% de los albaricoques se encuentran en mal estado, mientras que el 0.63% de la fruta de la explotación es ciruela y se encunetra en mal estado. Se elige una pieza de fruta al azar de la explotación agrícola, se pide:

   (a) Calcular la probabilidad de que no se encuentre en mal estado. (**0.25 puntos**)
   (b) Si la pieza de fruta se encuentra en mal estado, determinar la probabilidad de que esa fruta sea una ciruela. (**0.25 puntos**)
   (c) Determinar la probabilidad de que la pieza de fruta no se encuentre en mal estado y sea ciruela o albaricoque. (**0.25 puntos**)

3. [**1 punto**] La proporción de correos electrónicos SPAM que recibe cada usuario de internet se comporta como una variable aleatoria con función de densidad dada por:

$$f(x) = \begin{cases} \frac{3}{2}x^{\alpha}, & x \in (0,1) \\ 0, & \text{en otro caso} \end{cases}$$

Se pide:

   (a) Determinar el valor de α para que $f(x)$ sea verdaderamente una función de densidad. (**0.25 puntos**)
   (b) Determinar la función de distribución de la variable aleatoria X y la proporción de correos SPAM que se reciben por término medio. (**0.50 puntos**)
   (c) Sabiendo que un usuario ha recibido más de un 10% de correos SPAM, determinar la probabilidad de que no reciba más de un 25% de correos SPAM.(**0.25 puntos**)

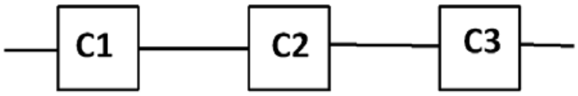
4. [**1 punto**] De un estudio sobre el uso de la plataforma Moodle por los estudiantes de la Universidad Politécnica de Cartagena se sabe que el número diario de visitas que realiza un estudiante se puede modelizar como una distribución de Poisson de varianza 5. Se pide:

   (a) Determinar la porobabilidad de que en un día determinado un estudiante realice a lo sumo 4 visitas.(**0.25 puntos**)

(b) Determinar la probabilidad de que en dos días, un estudiante realice exactamente 7 visitas a la plataforma Moodle.(**0.25 puntos**)

(c) Si se selecciona una semana al azar (= 7 días), determinar la probabilidad de que en más de 2 días de la misma se realicen a lo sumo 4 visitas a la plataforma Moodle. (**0.5 puntos**)

5. [**1.5 puntos**] Una máquina consta de tres componentes que funcionan de manera **independiente** entre sí según el diagrama siguiente:



Sabiendo que el tiempo de vida de cada componente sigue una distribución exponencial de media 500 horas, se pide:

(a) Determinar la probabilidad de que un componente funcione al menos 250 horas. (**0.25 puntos**)

(b) Si han transcurrido 200 horas y la componente C1 sigue funcionando, determinar la probabilidad de que la componente C1 no se averíe antes de las 50 horas siguientes. (**0.25 puntos**)

(c) Determinar la probabilidad de que la máquina funcione al menos 250 horas. (**0.25 puntos**)

(d) La máquina de la que hablamos fabrica piezas cuya longitud se distribuye según una variable aleatoria de media 40 mm. y de desviación típica 0.5 mm. Estas piezas se consideran aceptables si su longitud se aleja de su longitud media a lo sumo en 0.6 mm. ¿Podrías dar una cota de la proporción de piezas no aceptables que fabrica la citada máquina? Razonar la respuesta. (**0.25 puntos**)

(e) Supongamos a partir de ahora, que la longitud de las piezas se distribuye según una variable aleatoria normal de media 40 mm. y de desviación típica 0.5 mm. Determinar la probabilidad de que en un lote de 400 piezas fabricadas por esta máquina contenga al menos 300 piezas aceptables. (**0.5 puntos**)

6. [**1.25 puntos**] La composición de la leche varía considerablemente con la raza de la vaca, el estado de lactancia, alimento, época del año y muchos otros factores. Aún así, algunas de las relaciones entre componentes son muy estables y pueden ser utilizados para indicar si ha ocurrido alguna adulteración en su composición. Por ejemplo, la leche con una composición no adulterada posee una gravedad específica media igual a 1.0315 (a 20^{o}C). Cualquier alteración, por agregado de agua, puede ser facilmente identificada debido a que esta característica de la leche no tendrá ese valor. Con el fin de validar las entregas de leche de una determinada granja, se realizaron las pruebas sobre gravedad de la leche a las últimas 25 entregas de la granja, dando lugar los resultados que vienen en la tabla siguiente:

$$\sum_{i=1}^{25} x_i = 25.8225$$

$$\sum_{i=1}^{25} x_i^2 = 26.673$$

Admitiendo la hipótesis de normalidad para la gravedad específica de la leche de vaca a una temperatura de 20^{o}C, se pide:

(a) Construir de manera detallada un intervalo de confianza al 98% para la gravedad específica media de la leche de vaca a 20^{o}C. Indicar el margen de error de dicho intervalo. **(0.5 ptos.)**

(b) A partir de los datos muestrales, ¿qué conclusiones podemos sacar sobre la adulteración o no de las entregas de leche de dicha granja, esto es, podemos concluir que la gravedad específica media de la leche a 20^{o}C es igual a 1.0315? Plantear y llevar a cabo el contraste adecuado para responder a esta pregunta. Calcular de manera aproximada el $p - valor$ de la prueba.**(0.5 ptos.)**

(c) Supongamos que la desviación típica de la variable aleatoria "gravedad específica de la leche de vaca a una temperatura de 20^{o}C" es $\sigma = 0.01$, determinar qué tamaño de la muestra será necesario utilizar si queremos que el margen de error sea menor a una milésima al 98% de confianza. **(0.25 ptos.)**

Septiembre 2013 : GRADO de AGRÍCOLAS

P. 1: Y = Proliferación de la planta.

X = Cantidad de lluvia caída.

$n = 30 \rightsquigarrow \bar{x} = 417.7$, $s_x = 154.8$ | $\bar{y} = 1160.6$, $s_y = 395.1$

$$Y = 2.52 * X + b$$

(a): $\hat{b} = \bar{y} - \hat{a} \cdot \bar{x} = 1160.6 - 2.52 * 417.7 = \boxed{107.996}$

(b): $R^2 = \dfrac{s_{xy}^2}{s_x^2 * s_y^2} =$

Como $\hat{a} = \dfrac{s_{xy}}{s_x^2} = 2.52 \Rightarrow \boxed{s_{xy} = 2.52 * (154.8)^2 = 60386.8608}$

$$= \frac{(60386.8608)^2}{154.8^2 * 395.1^2} = \boxed{0.975}$$

Ajuste muy bueno

(c): • Si se le suma 5 a todos los valores de $X \Rightarrow Z = X + 5$ (≡ TRANSLACIÓN)

De donde, $\bar{z} = \bar{x} + 5 = 422.7$

$s_z^2 = s_x^2 = (154.8)^2$.

$$S_{zy} = \frac{1}{30}\sum_{i=1}^{30}(z_i - \bar{z})(y_i - \bar{y}) =$$

$$= \frac{1}{30}\sum_{i=1}^{30}(x_i + \cancel{5} - (\bar{x} + \cancel{5}))(y_i - \bar{y}) = S_{xy} =$$

$$= 60386.8608$$

entonces $y = c + dz$

$$\hat{d} = \frac{S_{yz}}{S_z^2} = \frac{S_{yx}}{S_x^2} = 2.52$$

$$\hat{c} = \bar{y} - \hat{d}\cdot\bar{z} = 1160.6 - 2.52 * 422.7 =$$

$$= 95.396 \Rightarrow \boxed{y = 2.52 * z + 95.396}$$

P. 2 :

$$\begin{cases} M = \text{melocotones} \rightarrow p(M) = \frac{3000}{7000} = \frac{3}{7} = \frac{6}{14} \\ A \equiv \text{albaricoques} \rightarrow p(A) = \frac{2500}{7000} = \frac{5}{14} \\ C \equiv \text{ciruelas} \rightarrow p(C) = \frac{1500}{7000} = \frac{3}{14} \end{cases}$$

$Q \equiv$ ESTAR en MAL ESTADO

$p(Q \mid M) = 0.05$

$p(Q \mid A) = 0.04$

$p(Q \cap C) = 0.0063 = p(C)\, p(Q \mid C)$

$$\Rightarrow p(Q \mid C) = \frac{0.0063}{3/14} = 0.0294$$

(a) :- $p(\bar{Q}) = 1 - p(Q) = \boxed{0.958}$

$$p(Q) = p(m)\,p(Q/m) + p(A)\,p(Q/A) + \overbrace{p(C)\,p(Q/C)}^{p(Q\cap C)} =$$

$$= \frac{6}{14} * 0.05 + \frac{5}{14} * 0.04 + 0.0063 = 0.042$$

(b) :- $p(C/Q) = \dfrac{p(C\cap Q)}{p(Q)} = \dfrac{0.0063}{0.042} =$

$$= \boxed{0.15}$$

(c) :- $p\left[\bar{Q} \cap (C \cup A)\right] = p\left[(\bar{Q}\cap C) \cup (\bar{Q}\cap A)\right] =$

incompatibles

$$= p(\bar{Q}\cap C) + p(\bar{Q}\cap A) =$$

$$= p(C)\,p(\bar{Q}/C) + p(A)\,p(\bar{Q}/A) =$$

$$= \frac{3}{14} * 0.9706 + \frac{5}{14} * 0.96 = \boxed{0.5509}$$

P.3:- X = Proporción de correos SPAM

$$f(x) = \begin{cases} \frac{3}{2} x^{\alpha}, & x \in (0,1) \\ 0, & \text{---} \end{cases}$$

(a):- • $f(x) = \frac{3}{2} x^{\alpha} > 0 \quad \forall x \in (0,1)$

• $1 = \int_{\mathbb{R}} f(x)\,dx = \frac{3}{2} \int_0^1 x^{\alpha}\,dx =$

$= \frac{3}{2} \left. \frac{x^{\alpha+1}}{(\alpha+1)} \right|_0^1 = \frac{3}{2} \frac{1^{\alpha+1}}{(\alpha+1)} =$

$= \frac{3}{2\alpha+2} \Leftrightarrow 3 = 2\alpha + 2 \Leftrightarrow 2\alpha = 1$

$\Leftrightarrow \boxed{\alpha = 1/2}$

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

$0 \qquad 1$

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

Si $0 < x < 1 \Rightarrow F(x) = \frac{3}{2} \int_{-\infty}^{x} t^{0.5}\,dt =$

$= \frac{3}{2} \int_0^x t^{0.5}\,dt = \cancel{\frac{3}{2}} \left. \frac{t^{0.5+1}}{\cancel{(0.5+1)}} \right|_0^x =$

$$= x^{1.5}$$

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

$$\Rightarrow \boxed{F(x) = \begin{cases} 0, & x \leq 0 \\ x^{1.5}, & 0 < x < 1 \\ 1, & x \geq 1 \end{cases}}$$

$$\Rightarrow E(X) = \int_{\mathbb{R}} x f(x)\,dx = \frac{3}{2}\int_0^1 x \cdot x^{0.5}\,dx =$$

$$= \frac{3}{2}\int_0^1 x^{1.5}\,dx = \frac{3}{2}\left.\frac{x^{1.5+1}}{1.5+1}\right|_0^1 = \frac{3}{2}\,\frac{1^{2.5}}{2.5} =$$

$$= \frac{1.5}{2.5} = \boxed{0.6}$$

(c): $P(X \leq 0.25 / X > 0.1) = \dfrac{P(0.5 < X \leq 0.25)}{P(X > 0.5)} =$

$$= \frac{P(0.1 < X \leq 0.25)}{1 - P(X \leq 0.1)} = \frac{F(0.25) - F(0.1)}{1 - F(0.1)} =$$

$$= \frac{(0.25)^{1.5} - (0.1)^{1.5}}{1 - (0.1)^{1.5}} = \boxed{0.0964}$$

P.4: $X \equiv$ N.º de visitas que realiza un estudiante al día al AV

$\sim Po(\lambda) = Po(\lambda = 5)$

$Var(X) = \lambda = 5$

$\Rightarrow f_X(x) = e^{-5} \dfrac{5^x}{x!}$, $x = 0,1,2,3,4,5,6,\dots$

(a):

$$P(X \leq 4) = e^{-5}\left(\frac{5^0}{0!} + \frac{5^1}{1!} + \frac{5^2}{2!} + \frac{5^3}{3!} + \frac{5^4}{4!}\right) =$$

$$= e^{-5}\left(1 + 5 + \frac{25}{2} + \frac{125}{6} + \frac{625}{24}\right) =$$

$$= e^{-5} * \frac{(144+300+500+625)}{24} = 65.375 \cdot e^{-5} = \boxed{0.4405}$$

(b): $Y \equiv$ N.º de visitas en dos días =

$= X_1 + X_2 \sim Po(\lambda = 5+5 = 10)$

$\left(\begin{array}{c} X_i \sim Po(\lambda = 5),\ i = 1,2 \\ \text{IND} \end{array}\right)$

$$f_Y(y) = e^{-10}\frac{10^y}{y!}, \quad y = 0,1,2\dots,6,7,8,\dots$$

$$P(Y = 7) = e^{-10}\frac{10^7}{7!} = \boxed{0.09}$$

(c): $T \equiv$ Nº de días en los que se realizan a lo sumo 4 visitas a la plataforma m. en un total de 7 días (éxito) $\sim B(n=7, p=0.4405)$

$$\begin{cases} f_T(t) = \binom{7}{t} 0.4405^t \; 0.5595^{7-t} \\ t = 0,1,2,3,4,5,6,7 \end{cases}$$

$$P(T>2) = P(T \geq 3) =$$

$$= 1 - P(T \leq 2) = 1 - \left[\binom{7}{0} 0.4405^0 \; 0.5595^7 + \right.$$

$$\left. + \binom{7}{1} 0.4405^1 \; 0.5595^6 + \binom{7}{2} 0.4405^2 \; 0.5595^5 \right] =$$

$$= 1 - \left(0.5595^7 + 7 * 0.4405 * 0.5595^6 + \right.$$

$$\left. + 21 * 0.4405^2 * 0.5595^5 \right) =$$

$$1 - 0.5595^5 \left(0.5595^2 + 7 * 0.4405 * 0.5595 + \right.$$

$$\left. + 21 * 0.4405^2 \right) = 1 - 0.3352 = \boxed{0.6648}$$

P.5: $X \equiv$ Tiempo de vida de cada componente

$\sim Exp(\lambda = 1/500)$

$E(X) = \frac{1}{\lambda} = 500 \Rightarrow \lambda = 1/500$

$$\underbrace{F_X(x) = \begin{cases} 0, & x \leq 0 \\ 1 - e^{-\frac{x}{500}}, & x > 0 \end{cases}}_{\text{F. DISTRIBUCIÓN de } X}$$

(a): $P(X \geq 250) = e^{-\frac{250}{500}} = e^{-0.5} = \underline{0.6065}$

(b): $P(X > 250 / X \geq 200) = \frac{P(X>250)}{P(X>200)} =$

$= \frac{e^{-\frac{250}{500}}}{e^{-\frac{200}{500}}} = \frac{e^{-0.5}}{e^{-0.4}} = e^{-0.5+0.4} = e^{-0.1} =$

$= \boxed{0.9048}$

(c):- p(máquina funcione al menos 250 h) =

$$p\left[(C1 \geq 250) \cap (C2 \geq 250) \cap (C3 \geq 250)\right] \underset{IND}{=}$$

$$= p(C1 \geq 250) * p(C2 \geq 250) * p(C3 \geq 250) =$$

$$= (0.6065)^3 = \boxed{0.2231}$$

(d):- L = Longitud de las piezas

$E(L) = 40$ mm.

$\sigma_L = 0.5$ mm.

Pieza ACEPTABLE si $|L - E(L)| \leq 0.6$

Como no conocemos la distribución de prob. de L vamos a utilizar la Desigualdad de Tchebychev:

$$p(\text{Aceptable}) = p(|L - E(L)| \leq 0.6) \geq$$

$$\left(0.6 = k * \sigma_L = k * 0.5 \Rightarrow k = \frac{0.6}{0.5} = 1.2\right)$$

$$\geq 1 - \frac{1}{(1.2)^2} = \boxed{0.3056}$$

$\Rightarrow$ Al menos el 30.56 % de las piezas son aceptables.

(e): Sup. $L \sim N(\mu = 40, \sigma_L = 0.5)$

$p(\text{aceptable}) = p(|L - E(L)| \leq 0.6) =$

$= p(-0.6 \leq L - 40 \leq 0.6) =$

$\phi\left(\frac{0.6}{0.5}\right) - \phi\left(\frac{-0.6}{0.5}\right) = 2\underbrace{\phi(1.2)}_{0.8849} - 1 =$

$= \boxed{0.7698}$ ← Prob. exacta de que una pieza sea aceptable.

$Y \equiv$ Nº de piezas aceptables en un lote de 400 piezas $\sim B(n = 400, p = 0.7698)$

$p(Y \geq 300) =$

$np = 400 * 0.7698 = 307.92$

$np(1-p) = 70.883184$

Si $W \sim N(307.92, \sigma_W = 8.42)$

$\Rightarrow Y \approx W$

$$P(Y \geq 300) \simeq P(W \geq 299.5) =$$

$$= 1 - \phi\left(\frac{299.5 - 307.92}{8.42}\right) =$$

$$= 1 - \underbrace{\phi(-1)}_{1-\phi(1)} = \phi(1) = \boxed{0.8413}$$

P.6: $X \equiv$ Gravedad específica de la leche de vaca a 20°C

$\sim N(\mu, \sigma)$

Datos muestrales:

$n = 25 \longrightarrow \sum x_i = 25.8225$

$\sum x_i^2 = 26.673$

$$\rightsquigarrow \begin{cases} \bar{x} = 1.0329 \\ s_x^2 = \frac{25}{24}\left(\frac{26.673}{25} - 1.0329^2\right) = 0.000039156 \\ \rightsquigarrow s_x = 0.0063 \end{cases}$$

(a): IC al 98% para μ:

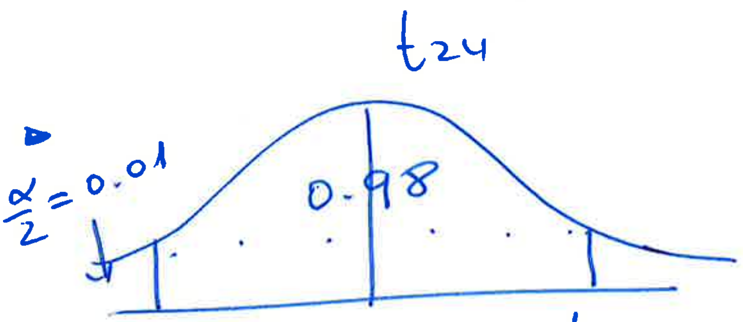
Tomamos $\bar{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right) \Rightarrow$

$$\Rightarrow Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \sim N(0, 1)$$

Como σ es desconocida, la estimamos mediante la S = desviación típica muestral

De donde, la v.a.

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t_{n-1} \quad (= t_{24})$$



$t_{24;0.99} = 2.49$

$-t_{24;0.99} = -2.49$

entonces $0.98 = P\left(-2.49 \leq \dfrac{\bar{X} - \mu}{S/\sqrt{n}} \leq 2.49\right)$

despejamos μ

$$= P\left(\bar{X} - 2.49 * \frac{S}{\sqrt{n}} \leq \mu \leq \bar{X} + 2.49 * \frac{S}{\sqrt{n}}\right)$$

$\Rightarrow$ el IC al 98% para μ sería:

$$\bar{X} \pm t_{n-1;0.99} * \frac{S}{\sqrt{n}} =$$

$$= \bar{X} \pm \underbrace{t_{24;0.99}}_{2.49} * \frac{S}{\sqrt{n}}$$

Para nuestra muestra tenemos:

$$\bar{x} \pm 2.49 * \frac{S}{\sqrt{n}} = 1.0329 \pm 2.49 * \frac{0.0063}{\sqrt{25}} =$$

$$= 1.0329 \pm 0.0031 = \boxed{(1.0298, 1.0360)}$$

Margen de error del IC $= \pm 0.0031$

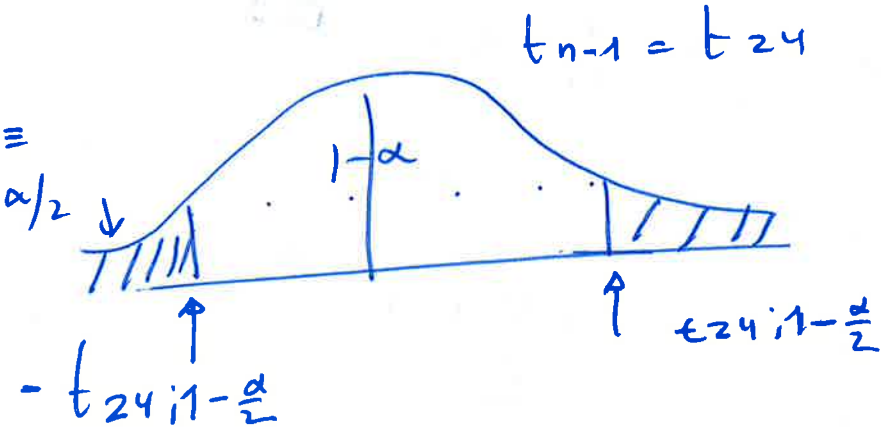
(b) :

$$\left.\begin{array}{l} H_0: \mu = 1.0315 \\ H_1: \mu \neq 1.0315 \end{array}\right\}$$

- α
- EST del CONTRASTE bajo H_0:

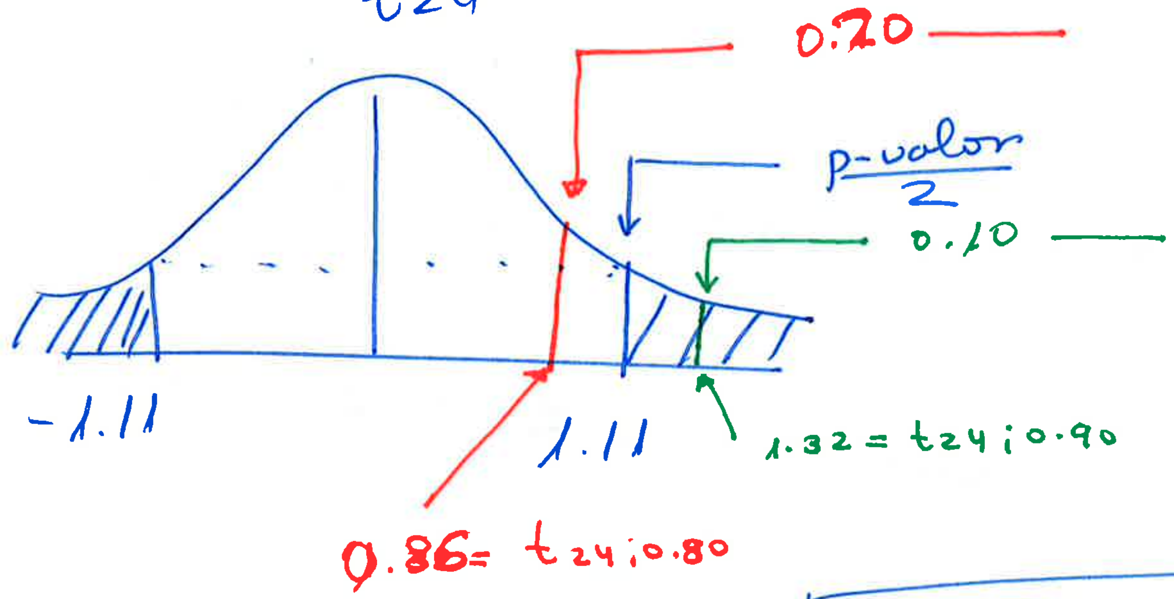
$$T \equiv \frac{\bar{X} - \mu}{S / \sqrt{n}} \sim t_{n-1} \ (\equiv t_{24})$$

- $R \equiv$



- $$t_0 \equiv \frac{1.0329 - 1.0315}{0.0063 / \sqrt{25}} = 1.11$$

▸ $p\text{-valor} = 2\,p(t_{24} > 1.11) =$



$\Rightarrow \quad 0.10 < \dfrac{p\text{-valor}}{2} < 0.20 \Rightarrow \boxed{0.20 < p\text{-valor} < 0.40}$

Aceptamos H_0 con plena confianza

⇒ las entregas de leche de dicha granja no están adulteradas.

(c): sup $X \sim N(\mu, \sigma = 0.01)$

$¿n \,/\, z_{1-\frac{\alpha}{2}} * \dfrac{\sigma}{\sqrt{n}} < 0.001$ al 98%

$\Rightarrow \quad z_{0.99} * \dfrac{\sigma}{\sqrt{n}} < 0.001 \Rightarrow 2.33 * \dfrac{0.01}{\sqrt{n}} < 0.001$

$$\Rightarrow \quad n > \left(\frac{2.33 * 0.01}{0.001}\right)^2 = 542.89$$

Necesitamos al menos 543 entregas.