Dpto. Matemática Aplicada y Estadística

Grado en IIAA y Grado en IHJ
Asignatura: **Estadística Aplicada. Curso 2011-2012**
JUNIO 2012

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

1. [**0.75 puntos**] Una empresa productora de cítricos constata que las producciones de dos fincas son de calidad distinta aunque utilizan la misma especie de árboles. Después de notar que los dos tipos de suelo son distintos (la finca A está constituida de suelo fino y homogéneo mientras que la finca B tiene una mezcla de suelo fino y suelo más grueso) decide hacer mediciones de la densidad del suelo en las dos fincas:

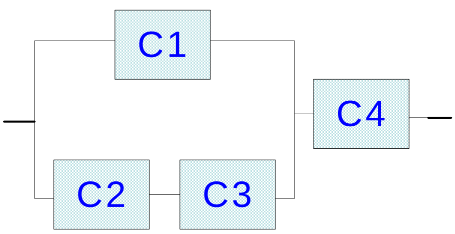
FINCA A						**FINCA B**					
87	89	86	89	94	92	92	87	94	95	92	95

   (a) Determinar la media y la desviación típica de la densidad del suelo en cada una de las dos fincas.
   (b) ¿En cuál de las dos fincas la densidad presenta menor dispersión relativa?
   (c) Si se cambia el dato $x = 86$ de la FINCA A por un dato arbitráriamente pequeño, de las siguientes afirmaciones ¿cuáles son verdaderas y cuáles falsas? Justificar la respuesta.
      i. La media disminuye y la varianza se mantiene constante.
      ii. La mediana no varía y la varianza crece.
      iii. El rango intercuartílico disminuye y la mediana permanece constante.

2. [**0.75 puntos**]Un agricultor dispone de tres productos para fumigar las parcelas dedicadas al cultivo de trigo. El producto P1 se utiliza en el 34% de las parcelas y los productos P2 y P3 tienen la misma posibilidad de ser utilizados para fumigar las parcelas. La probabilidad de que el rendimiento medio de la semilla sea superior al 60% es igual a 0.20 para las parcelas tratadas con el producto P1, a 0.90 para las parcelas tratadas con el producto P2 y a 0.45 para las parcelas tratadas con el producto P3. Se selecciona una parcela al azar, se pide:
   (a) La probabilidad de que tenga un rendimiento medio de la semilla superior al 60%.
   (b) Sabiendo que la semilla en dicha parcela ha tenido un rendimiento medio superior al 60%, ¿cuál es la probabilidad de que se haya tratado con el producto P2 o con el producto P3?.
   (c) La probabilidad de que haya sido tratada con el producto P3 si la parcela no tuvo un rendimiento medio de la semilla superior al 60%.

3. [**1 punto**] El número de accidentes de trabajo que se producen en una fábrica por semana sigue una distribución de Poisson de media 0.5 accidentes. Se pide:
   (a) Determinar la probabilidad de que en una semana haya a lo sumo un accidente. (**0.25 puntos**)
   (b) Determinar la probabilidad de que en un mes el número de accidentes que se producen se encuentre entre 4 y 7 (ambos incluidos). (**0.25 puntos**)
   (c) La Dirección General de Trabajo decide declarar *semanas laborables blancas* aquellas en las que a lo sumo se produce un accidente. ¿Cuál sería la probabilidad de que en año y medio el número de *semanas laborables blancas* esté por encima de 70? (**0.5 puntos**)

4. [**1 punto**] El diagrama siguiente corresponde a un dispositivo de una estación agroclimatológica formado por cuatro componentes eléctricas que funcionan de manera **independiente** entre sí:



El tiempo de vida de cada una de las componentes eléctricas sigue una distribución exponencial de media 500 horas. Con el fin de garantizar el tiempo de vida del dispositivo, el responsable de la estación exige que la duración de cada componente eléctrica sea superior a 200 horas desechando las que no cumplan la norma de calidad. A partir de estos datos, se pide:

(a) Determinar la proporción de componentes eléctricas que serán desechadas por el responsable de la estación. Si una componente eléctrica se ha comprobado 200 horas y no ha fallado, determinar la probabilidad de que no falle en las 100 horas siguientes. (**0.5 ptos.**).

(b) Determinar la probabilidad de que el dispositivo completo funcione al menos 200 horas. (**0.5 puntos**)

5. [**1.25 puntos**] Un fitomejorador desea controlar la variabilidad de los brotes comerciales de espárrago, ya que la norma de embalaje establece una longitud mínima de 19.2 cm.y una longitud máxima de 22.8 cm. La longitud de los brotes de este cultivo varía aleatoriamente con una media de 21 cm. y una desviación típica de 1.5 cm.Se pide:

   (a) ¿Podrías proporcionar una cota de la proporción espárragos que cumplen la norma de embalaje? Razonar la respuesta. (**0.25 puntos**)

   (b) Supongamos, a partir de este momento, que la longitud de los brotes de espárragos se puede modelizar como una distribución normal de media 21 cm. y de desviación típica 1.5 cm., dar respuesta a las siguientes cuestiones

      (b1) Determinar el porcentaje exacto de brotes de espárragos que cumplen la norma de embalaje. (**0.25 puntos**)

      (b2) Si el fitomejorador elige al azar 6 espárragos, determinar la probabilidad de que a lo sumo uno no verifique la norma de embalaje. (**0.5 puntos**)

      (b3) ¿Cuál debería ser el valor de la desviación típica de la longitud de los brotes de espárragos para que el porcentaje de espárragos que no cumple la norma de embalaje se reduzca al 5%? (**0.25 puntos**)

6. [**1.25 puntos**] En un criadero de semillas se está probando una nueva variedad de maiz que saldrá a la venta cuando el rendimiento promedio por hectárea sea superior a 267 Tm.. Para llevar a cabo el estudio, se plantó la nueva variedad de maiz en 19 parcelas experimentales de igual superficie y se anotó el rendimiento observado. Los resultados obtenidos vienen en la tabla siguiente:

rendimiento promedio por hectárea observado en la muestra: 270.5
varianza muestral : 56.25

Admitiendo la hipótesis de normalidad para el rendimiento de la nueva variedad de maiz por hectárea, se pide:

   (a) Construir de manera detallada un intervalo de confianza al 95% para el rendimiento medio de la nueva variedad de maiz por hectárea. ¿Cuál es el efecto del aumento del nivel de confianza en el intervalo de confianza? **(0.5 ptos.)**

   (b) A partir de los datos muestrales, ¿se puede concluir que la nueva variedad de maiz podrá salir a la venta, esto es, que el rendimiento promedio por hectárea es mayor a 267 Tm. ? Plantear y llevar a cabo el contraste adecuado para responder a esta pregunta tomando $\alpha = 0.05$ ¿Cuál sería la decisión al 99% de confianza? **(0.5 ptos.)**

   (c) Supongamos que la desviación típica de la v.a. rendimiento de la nueva variedad de maiz por hectárea es igual a $\sigma = 7.1$ Tm, determinar qué tamaño de la muestra será necesario utilizar si se desea estimar el rendimiento promedio de la nueva variedad de maiz por hectárea con un error de estimación menor de 1.8 Tm. al 95% de confianza. **(0.25 ptos.)**

GRADO de AGRÍCOLAS: Junio 2012

P. 1º

FINCA A: 87, 89, 86, 89, 94, 92 $n=6$

FINCA B: 92, 87, 94, 95 92 95

(a): $X \equiv$ Densidad de la finca A.

$\sum x_i = 537 \rightsquigarrow \bar{x} = \frac{537}{6} = 89.5$

$\sum x_i^2 = 48107 \rightsquigarrow \begin{cases} s_x^2 = \frac{6}{5}\left(\frac{48107}{6} - (89.5)^2\right) = 9.1 \\ s_x = 3.0166 \end{cases}$

$Y \equiv$ Densidad de la finca B.

$\sum y_i = 555 \rightsquigarrow \bar{y} = \frac{555}{6} = 92.5$

$\sum y_i^2 = 51383 \rightsquigarrow \begin{cases} s_y^2 = \frac{6}{5}\left(\frac{51383}{6} - (92.5)^2\right) = 9.1 \\ s_y = 3.0166 \end{cases}$

(b): $CV_x = \frac{s_x}{\bar{x}} = \frac{3.0166}{89.5} = 0.0337$

$CV_y = \frac{s_y}{\bar{y}} = \frac{3.0166}{92.5} = 0.0326$ ← en la finca B

la dispersión relativa de la densidad es un poco menor.

(c): Finca A: ~~86~~ 87 89 89 94 92

Dato PEQUEÑO

(i)

La media disminuye y la varianza se mantiene cte ← NO, ya que la media disminuye al disminuir la cuantía de una observación pero la varianza aumentará pues aumenta la dispersión de los datos.

(ii): La mediana no varía y la varianza crece

Sí es verdadera pues la mediana no varía ya que en su cálculo influye la posición de los datos y no su cuantía. La varianza ya hemos comentado que aumenta.

(iii): El rango intercuartílico permanece cte pues el hecho de que cambie el valor mínimo, éste no influye en el cálculo de Q1 y Q3. ⇒ No es verdadera

P.2:

$P1 \rightarrow p(P1) = 0.34$

$P2 \rightarrow p(P2) = \alpha = 0.33$

$P3 \rightarrow p(P3) = \alpha = 0.33$

$$0.34 + 2\alpha = 1 \Rightarrow \boxed{\alpha = \frac{1-0.34}{2} = 0.33}$$

R ≡ Rendimiento medio de la semilla es superior al 60% ⇒

$p(R/P1) = 0.20$

$p(R/P2) = 0.90$

$p(R/P3) = 0.45$

(a): $$p(R) = \sum_{i=1}^{3} p(Pi)\, p(R/Pi) =$$

$$= 0.34 * 0.20 + 0.33 * 0.90 + 0.33 * 0.45 =$$

$$= 0.068 + 0.297 + 0.1485 = \boxed{0.5135}$$

(b): $p(P2 \cup P3 \mid R) = \dfrac{p[(P2 \cup P3 \cap R)]}{p(R)} =$

$$= \frac{p[\overbrace{(P2 \cap R) \cup (P2 \cap R)}^{\text{incompatibles}}]}{p(R)} = \frac{p(P2 \cap R) + p(P3 \cap R)}{p(R)} =$$

$$= \frac{p(P2)\, p(R \mid P2) + p(P3)\, p(R \mid P3)}{p(R)} = \frac{0.33 * 0.90 + 0.33 * 0.45}{0.5135} =$$

$$= \frac{0.297 + 0.1485}{0.5135} = \boxed{0.8676}$$

(c): $p(P3 \mid \bar{R}) = \dfrac{p(\bar{R} \cap P3)}{p(\bar{R})} = \dfrac{p(P3)\; p(\bar{R} \mid P3)}{p(\bar{R})} =$

$$= \frac{p(P3)\,(1 - p(R \mid P3))}{1 - p(R)} = \frac{0.33 * (1 - 0.45)}{1 - 0.5135} =$$

$$= \frac{0.33 * 0.55}{0.4865} = \boxed{0.7669}$$

P.3: $X \equiv$ Nº de accidentes de trabajo POR SEMANA $\sim P_o(\lambda = 0.5)$

$$\Rightarrow f_X(x) = e^{-0.5} \frac{0.5^x}{x!}, \quad x = 0, 1, 2, 3, 4, \ldots$$

(a): $P(X \leq 1) = e^{-0.5}\left(\dfrac{0.5^0}{0!} + \dfrac{0.5^1}{1!} \right) =$

$$= e^{-0.5}(1 + 0.5) = 1.5 * e^{-0.5} = \boxed{0.9098}$$

(b): $Y \equiv$ Nº de accidentes en 1 mes

$\equiv X_1 + X_2 + X_3 + X_4 \sim Po(4 * 0.5 = 2)$

(con $X_i \sim Po(0.5)$, $i = 1,2,3,4$ INDEPENDIENTES)

$\Rightarrow f_Y(y) = e^{-2} \dfrac{2^y}{y!}, \quad y = 0,1,2,3,\underline{4,5,6,7},8,\ldots$

$p(4 \le Y \le 7) = e^{-2}\left(\dfrac{2^4}{4!} + \dfrac{2^5}{5!} + \dfrac{2^6}{6!} + \dfrac{2^7}{7!}\right) =$

$= e^{-2}\left(\dfrac{16}{24} + \dfrac{32}{120} + \dfrac{64}{720} + \dfrac{128}{5040}\right) =$

$= e^{-2} \dfrac{16}{24} \underbrace{\left(1 + \dfrac{2}{5} + \dfrac{4}{30} + \dfrac{8}{210}\right)}_{330/210} = \boxed{0.1418}$

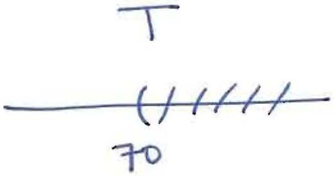
(c): → semana laborable blanca si $X \le 1$

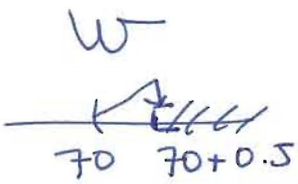
$\Rightarrow p(\text{semana laborable blanca}) = p(X \le 1) = 0.9098 \cong$
$\cong 0.91$

→ Año y medio $= 54 + \frac{54}{2} = 81$ semanas.

Definimos $T \equiv$ Nº de semanas laborables blancas en un total de 81 semanas $\sim B(n = 81, p = 0.91)$

$p(T \geq 70) =$





$$\begin{cases} np = 73.71 \\ np(1-p) = 6.6339 \rightarrow \sigma_W = 2.58 \\ \text{si } W \sim N(73.71, \sigma_W = 2.58) \text{ entonces } T \approx W \end{cases}$$

$$\cong P(T \geq 70.5) = 1 - \phi\left(\frac{70.5 - 73.71}{2.58}\right) =$$

$$= 1 - \phi(-1.24) = \phi(1.24) = \boxed{0.8925}$$

P. 4: $X \equiv$ Tiempo de vida de cada componente eléctrica $\sim Exp(\lambda)$

$E(X) = \frac{1}{\lambda} = 500 \Rightarrow \lambda = \frac{1}{500} = 0.002$

$$\Rightarrow \begin{cases} f_X(x) = \begin{cases} 0, & x \leq 0 \\ 0.002\, e^{-0.002x}, & x > 0 \end{cases} & \equiv \text{f. densidad} \\ F_X(x) = \begin{cases} 0, & x \leq 0 \\ 1 - e^{-0.002x}, & x > 0 \end{cases} & \equiv \text{F. DISTRIBUCIÓN} \end{cases}$$

- Norma de calidad $\Rightarrow X > 200$

(a): $P(\text{C.E. rechazada}) = P(X \leq 200) = F_X(200) =$

$$= 1 - e^{-0.002 * 200} = 1 - e^{-0.4} = \boxed{0.3297}$$

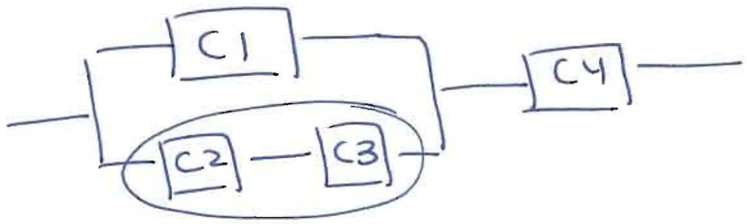
$$\triangleright \quad P(X \geq 300 / X > 200) =$$

$$= \frac{P[(X > 200) \cap (X \geq 300)]}{P(X > 200)} = \frac{P(X \geq 300)}{P(X > 200)} =$$

$$= \frac{e^{-0.002 * 300}}{e^{-0.002*200}} = \frac{e^{-0.6}}{e^{-0.4}} = e^{-0.6+0.4} = e^{-0.2} =$$

$$= \boxed{0.8187}$$

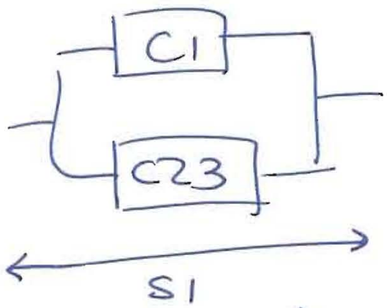
(b): $P(X \geq 200) = e^{-0.002*200} = e^{-0.4} = \boxed{0.6703}$



$\triangleright$ p(C23 funciona al menos 200 h) $= p((C2 \geq 200) \cap (C3 \geq 200)) =$

$$\stackrel{IND}{=} p(C2 \geq 200) * p(C3 \geq 200) = 0.6703^2 = \boxed{0.4493}$$

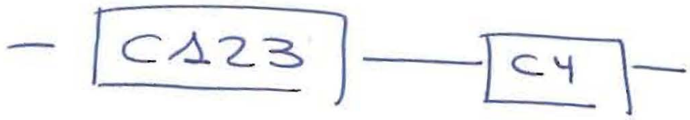
$\triangleright$



p(S1 funcione al menos 200 h) $= p[(C1 \geq 200) \cup (C23 \geq 200)] =$

$$= p(C1 \geq 200) + p(C23 \geq 200) - p(C1 \geq 200) * p(C23 \geq 200) =$$

$= 0.6703 + 0.4493 - 0.6703 * 0.4493 = \boxed{0.8184}$



$p(\text{sistema completo funcione al menos } 200h) =$

$= p\left[\overline{(C123 \geq 200) \cap (C4 \geq 200)}\right] =$

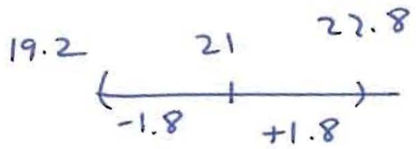
$= p(C123 \geq 200) * p(C4 \geq 200) = 0.8184 * 0.6703 =$

$= \boxed{0.5486}$

P. 5 :- $\left\{ \begin{array}{l} X \equiv \text{longitud de los brotes de espárragos} \\ E(X) = 21 \text{ cm.} \\ \sigma_X = 1.5 \text{ cm.} \end{array} \right.$

- Norma de embalaje que $X \in [19.2, 22.8]$

(a):- $p(\quad \leq X \leq 22.8) = p(|X - \mu| \leq 1.8) \geq$



$\geq p(|X - \mu| < \underbrace{1.8}_{\substack{\parallel \\ k * \sigma_X = k * 1.5 \Rightarrow k = \frac{1.8}{1.5} = 1.2}}) \geq 1 - \frac{1}{(1.2)^2} = \boxed{0.3056}$

Aplicamos la D. de Tchebychev pues no conocemos la distribución de prob. de X sólo la media y la desviación típica (b): sup que $X \sim N(\mu = 21, \sigma = 1.5)$

(b.1): $p(19.2 \leq X \leq 22.8) = \phi\left(\frac{22.8-21}{1.5}\right) -$

$- \phi\left(\frac{19.2-21}{1.5}\right) = \phi(1.2) - \underbrace{\phi(-1.2)}_{1-\phi(1.2)} =$

$= 2\phi(1.2) - 1 = \boxed{0.7698}$ ← prob. de cumplir la norma de embalaje

(b.2): Y = Nº de espárragos que NO cumplen la norma de embalaje en un total de 6 espárragos $\sim B(n=6, p=0.2302)$

$\Rightarrow f_Y(y) = \binom{6}{y} 0.2302^y \, 0.7698^{6-y}$, $y = 0, 1, 2, 3, 4, 5, 6$

$P(Y \leq 1) = \binom{6}{0} 0.2302^0 \, 0.7698^6 + \binom{6}{1} 0.2302^1 \, 0.7698^5 =$

$= 0.7698^6 + 6 * 0.2302 * 0.7698^5 =$

$= 0.7698^5 \underbrace{(0.7698 + 6 * 0.2302)}_{2.1511} = \boxed{0.5815}$

(b.3): sup que $X \sim N(\mu = 21, \sigma)$

¿σ / p(NO cumple la norma) = 0.05?

$\Rightarrow$ p(cumplir la norma de embalaje) = 0.95

$$0.95 = p(19.2 \le X \le 22.8) =$$

$$= \phi\left(\frac{22.8-21}{\sigma}\right) - \phi\left(\frac{19.2-21}{\sigma}\right) = \phi\left(\frac{1.8}{\sigma}\right) - \underbrace{\phi\left(\frac{-1.8}{\sigma}\right)}_{1-\phi\left(\frac{1.8}{\sigma}\right)}$$

$$= 2\phi\left(\frac{1.8}{\sigma}\right) - 1 \Rightarrow \phi\left(\frac{1.8}{\sigma}\right) = \frac{1.95}{2} = 0.975$$

$$\Rightarrow \frac{1.8}{\sigma} = Z_{0.975} = 1.96 \Rightarrow \boxed{\sigma = \frac{1.8}{1.96} \cong 0.92}$$

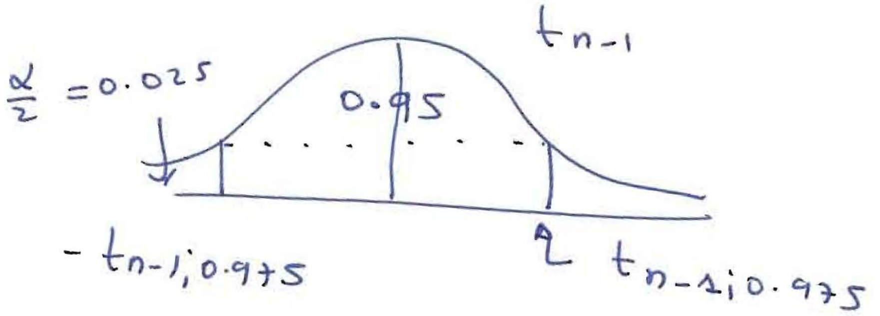
P. 6 :- $X \equiv$ Rendimiento de la nueva variedad de maíz por Ha $\sim N(\mu, \sigma)$

Datos muestrales: $n = 19 \rightarrow \begin{cases} \bar{x} = 270.5 \\ s = \sqrt{56.25} = 7.5 \end{cases}$

(a) :- IC al 95% para μ:

- Tomamos la v.a. $T = \dfrac{\bar{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}$ ya que la varianza es desconocida.

- En la t_{n-1} tomamos los percentiles de valor 0.025 y 0.975:



Entonces $0.95 = p\left(-t_{n-1;0.975} \leq \frac{\bar{x}-\mu}{\hat{s}/\sqrt{n}} \leq t_{n-1;0.975}\right) =$

Despejamos adecuadamente μ

$= p\left(\bar{x} - t_{n-1;0.975} * \frac{\hat{s}}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{n-1;0.975} * \frac{\hat{s}}{\sqrt{n}}\right)$

$\Rightarrow$ El IC al 95% para μ es:

$\bar{x} \pm t_{n-1;0.975} * \frac{\hat{s}}{\sqrt{n}}$

$\Rightarrow$ Para nuestra muestra quedaría:

$\bar{x} \pm t_{18;0.975} * \frac{s}{\sqrt{n}} = 270.5 \pm 2.10 * \frac{7.5}{\sqrt{19}} =$

$= 270.5 \pm 3.61 = (266.89, 274.11)$.

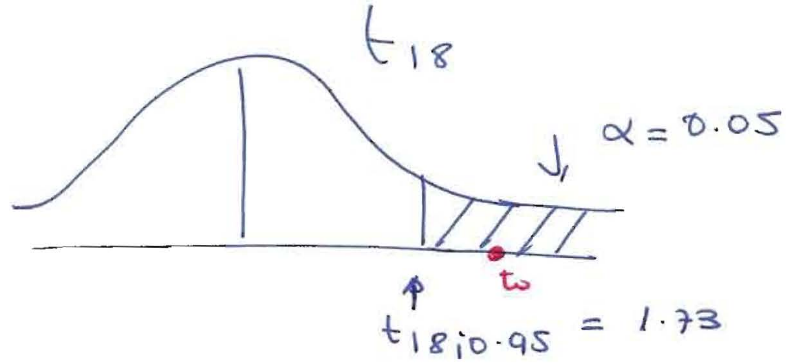
- Si $\uparrow 1-\alpha \Rightarrow$ Aumenta la amplitud del intervalo de confianza.

(b): - $H_0: \mu = 267$
  $H_1: \mu > 267$

- $\alpha = 0.05$
- Est del contraste bajo H_0

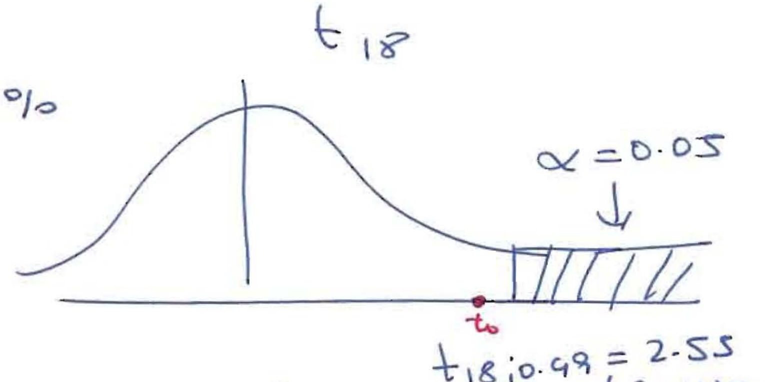
$T_0 \equiv \frac{\bar{x}-\mu_0}{\hat{s}/\sqrt{n}} \sim t_{n-1} \ (\equiv t_{18})$

▸ $R_{95\%} \equiv$



▸ $t_0 = \dfrac{270.5 - 267}{7.5/\sqrt{19}} = 2.03 \in R \Rightarrow$ Rechazamos H_0 al 95% de confianza ⇒ el rendimiento medio por hectárea es significativamente mayor a 267 kg por hectárea ⇒ la semilla puede salir a la venta.

Si $\alpha = 0.01 \Rightarrow R_{99\%}$



$t_0 \notin R_{99\%} \Rightarrow$ Aceptaríamos H_0 al 99% de confianza.

(c): Sup. $X \sim N(\mu, \sigma = 7.1)$

¿n / margen de error < 1.8 al 95%?

$$\Rightarrow Z_{1-\frac{\alpha}{2}} * \frac{\sigma}{\sqrt{n}} < 1.8 \Rightarrow Z_{0.975} * \frac{\sigma}{\sqrt{n}} < 1.8$$

$$\Rightarrow 1.96 * \frac{7.1}{\sqrt{n}} < 1.8 \Rightarrow n > \left(\frac{1.96 * 7.1}{1.8}\right)^2 =$$

$= 59.77$

$\Rightarrow$ Necesitamos plantar al menos 60 parcelas.