CALCULO ESTADISTICO Y BIOMETRIA

RESOLUCION MODELO SEGUNDO PARCIAL 2020

Ejercicio 1

a) Para construir un intervalo de confianza para la **media poblacional** con varianza poblacional desconocida, usamos la **distribución "t" de Student**.

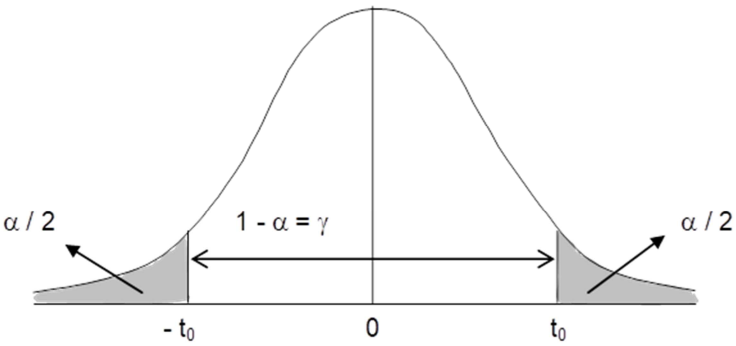
$$t_0 = \frac{\bar{x} - \mu}{S_{\bar{x}}}$$

$$P\left(\bar{X} - t_0 \frac{S}{\sqrt{n}} < \mu < \bar{X} + t_0 \frac{S}{\sqrt{n}} \right) = \gamma = 99\%$$

dónde:
$\bar{X} = 26.65$

S = $\sqrt{17.25}$ = 4.15

$$S_{\bar{x}} = \frac{S}{\sqrt{n}} = \frac{4.15}{\sqrt{30}} = 0.757$$



Entonces, ingresamos a la Tabla de la Distribución de t de Student con los grados de libertad (n - 1 = 29) y una probabilidad (área) a la izquierda del valor que queremos obtener de 0.995 y obtenemos el valor de t_0, en este caso **2,76.**

Finalmente reemplazamos

$$P(26.65 - 2.76 * 0.757 < \mu < 26.65 + 2.76 * 0.757) = 99\%$$

$$P(24.56 < \mu < 28.74) = 99\%$$

entonces el intervalo obtenido es: [24.56 ; 28.74]

b) Para construir el intervalo de confianza para la **varianza poblacional** usamos la **distribución Chi-cuadrado (χ^2)**

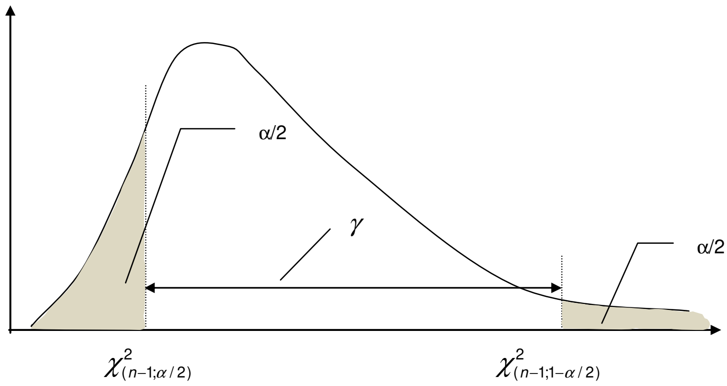
$$\chi^2 = \frac{(n-1)\cdot s^2}{\sigma^2}$$

$$P\left[\frac{(n-1)\cdot S^2}{\chi^2_{(n-1;1-\alpha/2)}} < \quad \sigma^2 \quad < \frac{(n-1)\cdot S^2}{\chi^2_{(n-1;\alpha/2)}}\right] = \gamma = 95\%$$

dónde:

n= 30

$S^2 = 17.25$



Como la tabla de Chi-cuadrado siempre indica el valor de probabilidad acumulada, debemos buscar el valor para el 2,5%, el cual utilizaremos para el límite inferior y el valor para el 97,5% que utilizaremos para el límite superior. En este caso para 29 grados de libertad tenemos:

$$\chi^2_{(29;\,0.025)} = 16.0 \quad y \quad \chi^2_{(29;\,0.975)} = 45.7$$

$$P\left[\frac{(30-1)\cdot 17.25}{45.7} < \sigma^2 < \frac{(30-1)\cdot 17.25}{16.0}\right] = 95\%$$

$$P(10.94 < \sigma^2 < 31.26) = 95\%$$

Por lo tanto el intervalo obtenido es: [10.94 ; 31.26]

__Ejercicio 2__

a) En base a la información del enunciado corresponde realizar un **test bilateral**, ya que no podemos suponer un comportamiento diferencial entre ambas épocas. Como desconocemos las varianza poblacionales y contamos con datos muestrales debemos utilizar la **distribución "t" de Student.**

$$H_0: \mu_{abril} = \mu_{agosto}$$
$$H_1: \mu_{abril} \neq \mu_{agosto}$$

b) Suponiendo varianzas poblacionales desconocidas pero iguales (siempre realizaremos esta suposición en todos los ejercicios de este tipo) realizamos el cálculo de t.

En este caso como los n son iguales podemos utilizar cualquiera de las expresiones siguientes

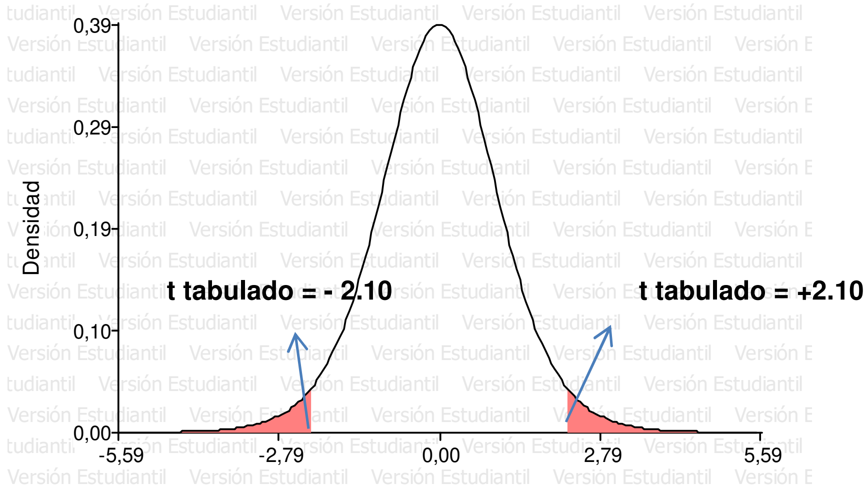
I) $$t_0 = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

II) $$t_0 = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{s_p^2\left[\frac{1}{n_1} + \frac{1}{n_2}\right]}}$$

Por simplicidad utilizaremos la expresión general, o sea la I)

$$t_0 = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} = \frac{40.8 - 26.6}{\sqrt{\frac{12.8}{10} + \frac{17.2}{10}}} = \frac{14.2}{1.732} = 8.19$$

Obtenemos el t tabulado ingresando a la Tabla de la Distribución de t de Student con los grados de libertad (n_1+n_2 - 2 = 18) y una probabilidad (área) a la izquierda del valor que queremos obtener de 0.975 y obtenemos el valor de t_0, en este caso **2,10.**



-----------------+++++++++++++-----------------

Rechazo H_0 No rechazo H_0 Rechazo H_0

Regla de decisión: si "t" calculado > 2,10 ó "t" calculado < -2,10 rechazo H_0,

en caso contrario no rechazo H_0.

Conclusión: Como el "t" calculado > "t" tabla para un nivel de significancia del 5%, rechazo H_0, es decir, que existe evidencia estadísticamente significativa para suponer que la deshidratación de los tubérculos difiere según la época de cosecha.

Ejercicio 3

a) Como necesitamos determinar si la inseminación artificial está relacionada o no con el uso de un suplemento vitamínico y contamos con las unidades experimentales (vacas) clasificadas según estos dos criterios utilizaría una **prueba de Chi-cuadrado de independencia o tabla de contingencia**.

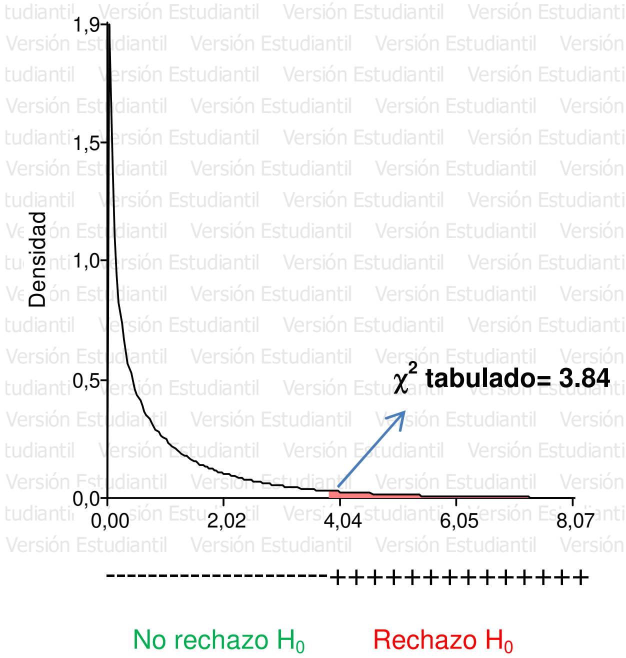
b) Planteo de hipótesis

H_0: Fobs. = F Cal., bajo el supuesto que la vaca este preñada o no es independiente del uso del suplemento vitamínico.

H_1: Fobs. ≠ F Cal., bajo el supuesto que la vaca este preñada o no depende del uso del suplemento vitamínico.

c) χ^2 calculado = 8.07

Obtenemos de la tabla el χ^2 **tabulado o crítico** para un nivel de significancia del 5% y gl = (filas - 1)x(columnas-1)= (2-1)x(2-1) =1



Conclusión: Como el χ^2 **calculado** es mayor que el tabulado para un nivel de significancia del 5%, rechazo H_0, es decir, que existe evidencia estadísticamente significativa para suponer que el éxito de la inseminación artificial (estar preñada) está relacionada con el uso del suplemento vitamínico.

Ejercicio 4

a) Diseño utilizado: Diseño con bloques completamente aleatorizado (DBCA).
La elección de este diseño se debió a la heterogeneidad de las unidades experimentales dado por la pendiente del terreno, a fin de mejorar la precisión o sensibilidad del experimento. Dentro de cada grupo homogéneo de unidades experimentales (bloques) se realizó la asignación aleatoria de cada tratamiento.

b)

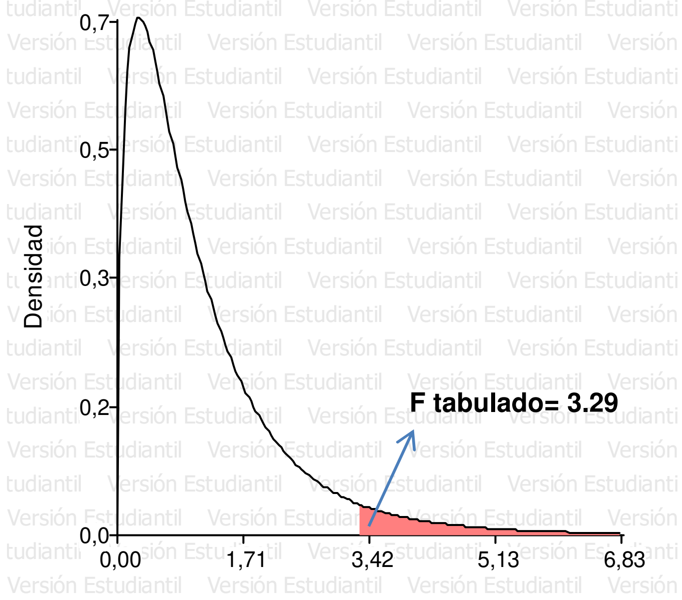
Fuentes de Variación	***Suma de cuadrados***	***gl***	***Cuadrados Medios***	***F** Calculado*
Fechas	357.2	3	119.06	20.18
Bloques	74.6	5	14.92	
Error Experimental	88.5	15	5.9	
Total	520.3	23		

Planteo de hipótesis

H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4$

H_1: Hay al menos una media que difiere de otra

Obtenemos el **F tabulado o crítico** con un nivel de significancia del 5% y grados de libertad del numerador 3 y grados de libertad del denominador 15.



--------------------++++++++++++++

No rechazo H_0 Rechazo H_0

Como el valor de **F calculado** (20.18) es **mayor** que al **F tabulado** (3.29), la **hipótesis nula es rechazada**, por lo tanto podemos concluir para un nivel de significancia que hay al menos una media que difiere de otra. Para detectar cuales difieren significativamente y cuales solo por el azar aplicaremos el Test de Tukey.

c) La diferencia mínima significativa (DMS) calculada por el test de Tukey (4.04) es comparada con la diferencia entre las medias de los tratamientos tomados de a pares. Cuando la diferencia entre las medias no supera la DMS se concluye que ambas medias no difieren significativamente y que solo difieren por el azar, mientras que si la diferencia entre las medias supera la DMS podemos concluir que ambas medias difieren significativamente, es decir que la diferencia entre ambas medias es debido a un efecto del tratamiento.

Tratamiento	*Media*	*Significancia*
Fecha 1	18.7	a
Fecha 3	20.5	a
Fecha 4	25.1	b
Fecha 2	28.3	b

Las medias con la misma letra presentan diferencias no significativas para un α del 5%

Conclusión: Las fechas óptimas de siembra son la fecha 2 y 4, ya que son las de mayor rendimiento, no se diferencian significativamente entre sí y se diferencian significativamente de las otras.

Ejercicio 5

a) Tenemos la relación funcional **Y = 0.76 + 1.59 X** , siendo:

Y: es la variable dependiente aleatoria, que representa la cantidad de nitrógeno en la planta.

X: la variable independiente fija, que representa la cantidad de nitrógeno en el suelo.

b = 1.59 es la pendiente de la recta de regresión, que representa el incremento del contenido de nitrógeno en la planta cuando la cantidad de nitrógeno en el suelo aumenta una unidad (relación positiva)

a = 0.76 es la ordenada al origen que representa el contenido de nitrógeno en la planta cuando el contenido de nitrógeno en el suelo de cero.

b) Prueba de hipótesis para la ordenada en el origen:

i) Las hipótesis $H_0 : \beta_0 = 0$

$H_1 : \beta_0 \neq 0$

ii) el t_a calculado

$t_a = 0.76 / 0.074 = 10.27$

Prueba de hipótesis para la pendiente:

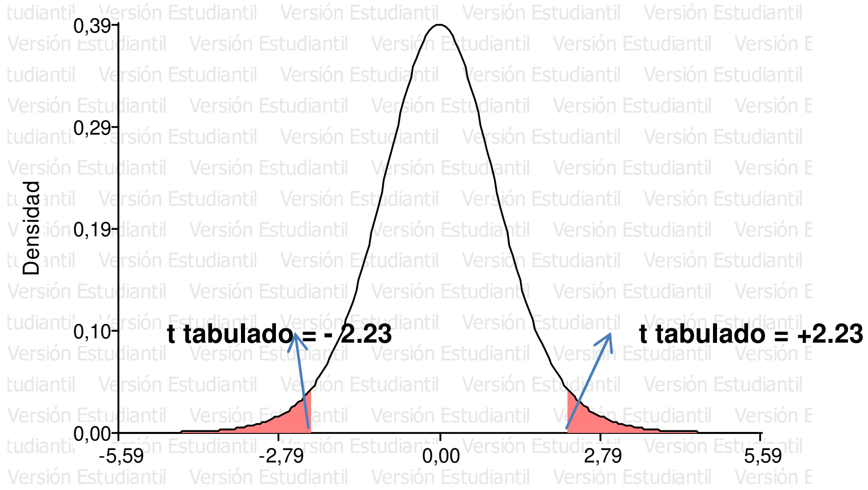
i) Las hipótesis $H_0 : \beta_1 = 0$

$H_1 : \beta_1 \neq 0$

ii) el t_b calculado

$t_b = 1.59 / 0.11 = 14.45$

Obtenemos el **t tabulado**, en función del nivel de significancia 5% y de los grados de libertad (n-k-1) que para una regresión simple se puede calcular como: gl= n – 2 = 12 – 2 = 10



-----------------+++++++++++++-----------------

Rechazo H_0 No rechazo H_0 Rechazo H_0

Conclusión: Las hipótesis nula, para la ordenada en el origen y la pendiente son rechazadas, porque t_a y t_b superan el valor de t_{tabla}. Es decir, que la recta no pasa por el origen y que NS y NP presentan una relación lineal positiva y significativamente diferente de cero para un nivel de significancia del 5%.

c) **R^2 es el coeficiente de determinación**, que mide la proporción de la variación total explicada por el modelo de regresión. En nuestro caso, el 96% de la variación en la cantidad de nitrógeno en la planta puede ser explicada por la disponibilidad de nitrógeno en el suelo.