**Apellido y Nombre: BLET SOFIA**

**Segundo Parcial - Estadística - 14/11/23 - IIB - INTECH**

Guardá este archivo nombrándolo con tu apellido. Pegá las salidas de Infostat que justifiquen tus respuestas. Expresá los resultados con sus unidades y las conclusiones en términos del problema. Recordá grabar periódicamente tus avances. Al finalizar envíalo por mail a agoarzac@agro.uba.ar. ¡Buena suerte!

------------------------------------------------------------------------------------------------------------------------------

**Calificación Final = 6/10**

**Calificación por problema : Problema 1: 2,5/10**

**Problema 2 (parte I): 6,5/10**

**Problema 2 (Parte II): 9/10**

**Problema 1**. La declinación en la diversidad de polinizadores naturales es un problema que puede perjudicar tanto a la flora nativa como a la producción de alimentos. Se realiza un estudio con el objetivo de evaluar si los cultivos con bordes en los que crecen especies vegetales espontáneas, podrían funcionar como refugio de insectos polinizadores favoreciendo su presencia. Para ponerlo a prueba se eligieron al azar 60 parcelas cultivadas con girasol en la zona de Balcarce y Lobería (Buenos Aires). En cada una se registró si tenían o no bordes con vegetación espontánea, y se las clasificó como parcelas con alta, media o baja diversidad de polinizadores, obteniendo los siguientes resultados:

|  |  |  |
| --- | --- | --- |
| Diversidad | Con bordes vegetados | Sin bordes |
| Alta | 15 | 6 |
| Media | 10 | 12 |
| Baja | 5 | 12 |

Total:60 de 30 c/u

1. Identifique:
2. tipo de ensayo (observacional/manipulativo):

Es un tipo de ensayo observacional

1. la unidad muestral: Cada una de las parcelas cultivadas con girasol en la zona de Balcarce y Lobería (Buenos Aires)
2. la/s variable/s aleatoria/s y su clasificación: parcelas con alta, media o baja diversidad de polinizadores. Tenemos 2 variables categorícas ,tipo de diversidad y con borde o sin borde vegetado
3. la población sobre la que aplica la inferencia

muestra :60 parcelas ,30 con bordes vegetados y 30 sin

La población son todas las parcelas cultivadas con girasol de Lobería y balcarce

1. Si se desea saber si la diversidad de polinizadores está asociada a la presencia de bordes vegetados, ¿qué prueba debería implementarse? Escriba sus hipótesis verbales y en términos estadísticos.

**PRUEBA DE INDEPENDENCIA**

H0: la diversidad de polinizadores no está asociada a la presencia de bordes vegetados (es independiente)

Ha: la diversidad de polinizadores está asociada a la presencia de bordes vegetados

H0: µ1= µ2 Las hipótesis estadísticas se plantean sobre el parámetro pi, y se plantean comparando probabilidades condicionales, no se estiman medias poblacionales cuando las variables son categóricas sino que se estiman probabilidades

1. Realice la prueba y concluya en términos del problema. Utilice un nivel de significación de 5% (Puede realizarla con INFOSTAT o a mano).

P(Con bordes vegetados/alta)=15/15+6=0.714

P(Con bordes vegetados/media)=10/10+12=0.45

P(Con bordes vegetados/baja)=5/5+12=0.294

P(sin bordes vegetados/alta)=6/15+6=0.285

P(sin bordes vegetados/media)=12/24=0.5

P(sin bordes vegetados/baja)=12/24=0.5

Grados de libertad: 2

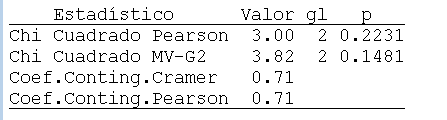
(3-1)\*(2-1) filas y columnas

Los grados de libertad están bien pero las probabilidades condicionales calculaste para qué son? En el punto d) está el resultado de la prueba que debería estar informado acá.

De todas maneras el valor P es incorrecto, esto debe ser porque debe haber habido algún error en el ingreso de la tabla de datos en el infostat ( sólo había que ingresar la tabla con las frecuencias observadas)

1. ¿Qué tipo de error se podría estar cometiendo? Interprételo en el contexto del problema.

Error tipo 1: Diría que la diversidad de polinizadores no está asociada a la presencia de bordes vegetados cuando en realidad si está relacionada.Esta respuesta es incoherente con tu resultado, si nos e rechaza la Ho el posible error es el de tipo II



Debido a que el p-valor (Chi-cuadrado Pearson) es menor a α, entonces se rechaza H0.

En esta tabla el valor P es mayoral alfa

1. ¿Qué aconsejaría para el cultivo de girasol? Justifique con una estimación obtenida del ensayo.

**CONCLUSIÓN:** queda demostrado que la diversidad de polinizadores está asociada a la presencia de bordes vegetados.(Existe una relación lineal). No se está estudiando una relación lienal , las dos variables estudiadas son categóricas

**Aconsejo tener lo** cultivos con bordes en los que crecen especies vegetales espontáneas, para ser refugio de insectos polinizadores favoreciéndose ambos

**Problema 2 (Parte I y II).**

Un cultivo de cobertura es una cobertura vegetal viva, temporal o permanente, que cubre el suelo y que se cultiva en asociación con otras plantas. La inclusión de cultivos de cobertura entre dos cultivos de interés económico busca generar mejoras en las condiciones del suelo que contribuyan al mantenimiento o al aumento de los rendimientos de los cultivos. Estas mejoras pueden tener lugar en las propiedades físicas y químicas del suelo, en el control de la erosión y/o el control de las malezas, o en el uso de agua, entre otras. Las especies más utilizadas como cultivos de cobertura en zonas templadas pertenecen a las familias de las gramíneas y de las leguminosas. Las gramíneas presentan un rápido crecimiento y una mayor producción de biomasa. Por lo tanto, se aconseja su uso como cultivo de cobertura cuando se busca controlar la erosión, absorber nutrientes para evitar su lixiviación, disminuir el plantel de malezas, incrementar el nivel de materia orgánica y mejorar la estabilidad estructural del suelo. Por otro lado, las leguminosas son la alternativa cuando el objetivo es la fijación de N atmosférico debido a la asociación raíz-rizobio. En este contexto, se llevan a cabo las siguientes investigaciones:

1. Se desea encontrar el cultivo de cobertura más eficiente para incorporar nitrógeno al suelo (N-NO3). Para ello, se dispuso de 40 parcelas (INTA Marcos Juárez) en las cuales se cultivó en forma aleatoria y balanceada, una leguminosa (*Vicia villosa, Vv*), una gramínea (Centeno; *Secale cereale, S*), o una mezcla de ambas (V S), mientras que otras fueron dejadas como testigo (t). Después de 6 meses se midió la concentración de nitrógeno (N-NO3 kg ha-1) en suelo (0-60 cm de profundidad). Los resultados se encuentran el archivo (**centeno vicia.IDB2**).
2. Identifique la unidad experimental, enuncie y clasifique las variables involucradas en este estudio.

Unidad experimental: Cada parcela del(INTA Marcos Juárez)

Variable respuesta: Concentración de nitrógeno (N-NO3 kg ha-1) en suelo (0-60 cm de profundidad)

Variable explicativa: Cada tipo de tratamiento; Una leguminosa (*Vicia villosa, Vv*), una gramínea (Centeno; *Secale cereale, S*), una mezcla de ambas (V S), mientras que otras fueron dejadas como testigo (t)

1. Escriba el modelo estadístico de la prueba que usará y enuncie las hipótesis que pondrá a prueba. Falta el modelo estadístico

**PRUEBA DE HIPÓTESIS: ANOVA**

Ho: µVv= µS= µVS=µt (Las especies no difieren, no hay cambios en cuanto a la concentración de nitrógeno) ,es decir que **el contenido medio de nitrógeno en el suelo es similar con los distintos tipos de cobertura**

Ha: algún µ sea distinto. (Al menos una especie difiere en cuanto a la concentración de nitrógeno).

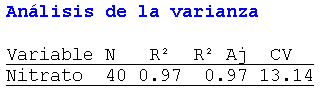
1. ¿Cuáles son los supuestos del modelo? Evalúelos y concluya.

**SUPUESTOS**:

- (las muestras deben ser aleatorias) y observaciones independientes.

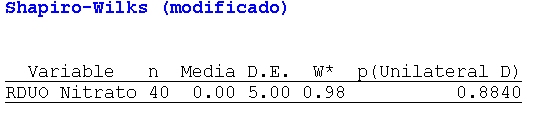
-Distribución normal de la variable en cada subpoblación (NORMALIDAD). Se realizó prueba de normalidad sobre los residuos: H0: errores con buen ajuste ; Ha: errores sin distribución normal.

-Las varianzas deben ser iguales (HOMOCEDASTICIDAD). Se realizó prueba de Levene, utilizando los RABS y la variable explicativa.



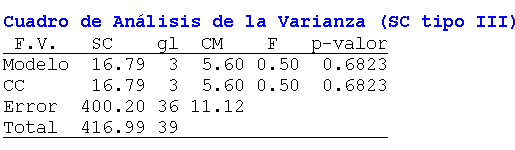


RESULTADOS DE PRUEBA DE SHAPIRO-WILKS

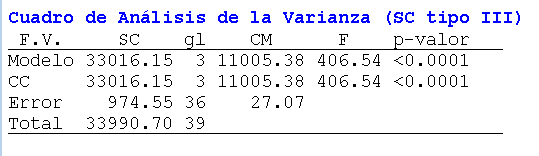


Debido a que el p-valor es mayor a α, no se rechaza la H0, no hay evidencia de falta de normalidad

RESULTADOS PRUEBA DE LEVENE ¿qué se concluye?

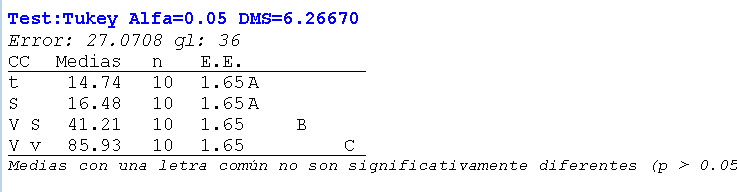


RESULTADOS PRUEBA DE TUKEY (no es la prueba de tukey sino el anova general y debería ser informado en el siguiente punto)



El p-valor es menor que alfa, se rechaza la H0 y, por lo tanto, realiza prueba de Tukey (comparaciones entre todos los tratamientos)





1. Efectúe el análisis y concluya en relación a los objetivos de la investigación. ¿Qué cobertura recomendaría?

Se observan diferencias significativas entre cada especie.

**CONCLUSIÓN:** dentro del grupo de las diferentes especies, la leguminosa (*Vicia villosa, Vv*) fue capaz de tener la mayor concentración (media)de nitrógeno en comparación con el resto de las especies. Por lo tanto, se recomendaría el uso de la leguminosa para poseer el cultivo de cobertura más eficiente para la incorporación de nitrógeno al suelo (N-NO3).

1. Si se realiza un ensayo sólo para estudiar los efectos sobre el suelo de la *Vicia v.* como cultivo de cobertura y este consiste en elegir 10 parcelas al azar, dividirlas a la mitad y sembrar en una mitad de cada una de ellas *Vicia v.* y a la otra mitad dejarla sin cultivo. Luego de seis meses se mide el contenido de nitrógeno en el suelo. ¿De qué tipo de diseño se trata? ¿Qué prueba de hipótesis utilizaría para analizarlo?

Se trata de un diseño experimental y haría un test T (muestras apareadas o independientes)

H0: µ=0 No hubo cambio en cuanto a la concentración de nitrógeno en la parcela con la *Vicia v con la sin cultivo*

Ha: µ≠0 Si hubo cambio en cuanto al el contenido de nitrógeno en el suelo

(las hipótesis no se expresan en tiempo pasado porque la hipotesis no es sobre el ensayo que pasó, sino sobre el proceso. La pregunta de investigación que querés responder es si *Vicia* es efectiva en aumentar el contenido de N en suelo. Entonces la hipótesis es sobre lamedia de las diferencias porque las observaciones están apareadas Ho: µcontro-vicia<= 0, es decir que el contenido medio de N en parcelas sembradas con Vicia no es mayor que en el de parcelas control )

1. Como el aporte de nitrógeno al suelo de *V. villosa* varía con la producción de materia seca, se diseñó un estudio con el objetivo de analizar la relación entre la densidad de siembra y la cantidad de materia seca producida. Para ello se sembró en forma balanceada *V. villosa* en 25 parcelas en densidades que variaron entre 10 y 160 plantas /m2; después de cinco meses se midió la biomasa seca (g m-2) producida en cada parcela. Los resultados se encuentran en el archivo **materia seca.IDB2.**
2. Represente los datos mediante un diagrama de dispersión y describa la relación entre la densidad de siembra y la producción de materia seca.

**PRUEBA DE HIPÓTESIS: REGRESIÓN LINEAL SIMPLE**

**UNIDAD EXPERIMENTAL:** Cada parcela.

**VARIABLE RESPUESTA:** la biomasa seca (g m-2) producida en cada parcela (Cuantitativa continua)

**VARIABLE EXPLICATIVA:** densidades que variaron entre 10 y 160 plantas /m2(Cuantitativa, fija).



Se puede apreciar que si hay relación entre biomasa y densidad

Se observa que la densidad media aumenta con la biomasa de siembra presente, pero la cada parcela individual puede no aumentar con biomasa. Eso se evalúa

1. Enuncie el modelo estadístico que permita responder a la pregunta de investigación e interprete sus términos en el contexto del problema.

Parámetross

**Yi = Beta0 + Beta1.Xi + Error aleatorio i**

i=1-25; Dominio:          10<=X<=160 plantas /m2 ; Ei~NID (0;sigma2)

Yi = Biomasa seca (g m-2) producida en cada parcela.

Beta0= Media de biomasa de todas las parcelas con 0 de densidad (es correcta lainterpretación pero en este caso como el que el valor 0 densidad de plantas no fue incluido en el ensayo, es una extrapolación , el 0 no pertenece al dominio del modelo)

Beta1 = cambio en la media de biomasa por cada cambio unitario de 1 densidad(plantas /m2)

Xi= Densidad de siembra

1. Ponga a prueba la hipótesis nula que afirma que la biomasa seca de *Vicia villosa* no depende linealmente de la densidad de siembra, con un nivel de significación de 0,05 y concluya.

H0: la variabilidad de la biomasa seca de *Vicia villosa* no se explica linealmente por la variación de la densidad de siembra

Ha: la variabilidad de la biomasa seca de *Vicia villosa* si se explica linealmente por la variación de la densidad de siembra.

(H0: β1=0

Ha: β1≠0

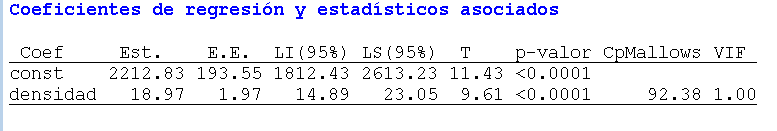
**SUPUESTOS:**

**Tiene que ser una muestra aleatoria, con observaciones independientes, con distribución normal, presentar linealidad e igualdad de varianza.**



El qqplot forma parte de la evaluación del supuesto de normalidad que en este caso no se pide

1. Interprete en el contexto del problema el intervalo de confianza de 95% para la pendiente.



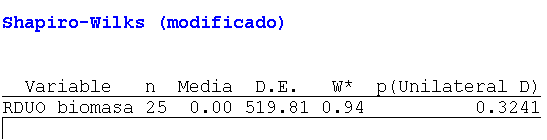
Ordenada

Pendiente

La pendiente (β1)tiene un valor de 18.97 g m-2 / plantas /m2

El intervalo nos da entre 14.89 y 23.05, veo que la pendiente difiere significativamente de cero. Hay evidencia que de la biomasa tiene efectos sobre la densidad de siembra. El p-valor nos da menor que α (α=0.05), se rechaza H0.

Falta completar la interpretación del intervalo de confianza: Con una confianza de 95% el aumento en la media de biomasa seca de *Vicia v.* por cada aumento unitario en la densidad de plantas está contenido entre 14,89 y 23, 05 (g/m2 / plantas/m2)



Este resultado es el de la prueba de normalidad de shapiro wilks y no corresponde informarlo acá.

Modelo estimado: Y= 2212.83 +18.97X

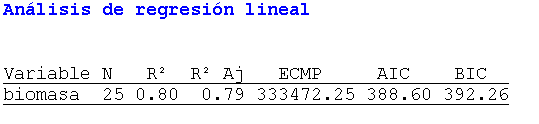
Predicciones solo en el rango considerado de X

Coeficientes de regresión: Indica que ante un aumento de una unidad (g m-2 ) de biomasa, la densidad media aumenta 18.97 g m-2 / plantas /m2

(por dar positivo)

La ordenada al origen indica que cuando no hay densidad de siembra presentes en el suelo, la biomasa es 2212.83 g m-2

1. Informe cuál es el porcentaje de la variabilidad total en la biomasa de materia seca de *Vicia v.* que es explicada por las variaciones en la densidad de siembra.



El 80% de la variabilidad esta explicado por la densidad de siembra y el 20% restante de la variabilidad se debe a otras causas. (error)