
UASLP.png

UASLP

Facultad de Agronomía y Agricultura.

Solución al examen 2 de Estadística.

Programación en R,
Regresión Lineal.

Soluciones

1 COMPARACION DE DOS MODELOS

- **Reporta por escrito en papel** el modelo estimado de la base de datos hombres y el modelo estimado de la base de datos mujeres

Solución:

Modelo hombres: $\text{TiempoLog} = -2.823196 + 1.112214 * \text{DistanciaLog}$

Modelo mujeres: $\text{TiempoLog} = -2.69216 + 1.11167 * \text{DistanciaLog}$

- **Reporta por escrito en papel** cual de los dos modelos indica más rapidez

Solución: Rapidez es la pendiente del modelo, la pendiente del modelo de hombres es ligeramente mayor al del modelo de mujeres, entonces, el modelo de hombres tiene mayor rapidez.

- **Reporta por escrito en papel** las hipótesis nula H_0 y alternativa H_A de la intercept y slope(s) de cada uno de los dos modelos y tu decisión de aceptar o no las hipótesis nula con un nivel de significancia de $\alpha = 0.05$ (usa los p-value). Reporta también si las relaciones son crecientes o no con base en cada pendiente

H_0 : Intercept=0

H_A : Intercept $\neq 0$

Rechazar la hipótesis nula porque los p-values son menores a 0.05

Similarmente para las hipótesis de los coeficientes de cada modelo: rechazar las hipótesis nulas de que las pendientes son cero.

Cada pendiente en cada modelo es positiva, por lo tanto las relaciones son crecientes (a mayor distancia, mayor tiempo)

- **Reporta por escrito en papel** el modelo con los datos concatenados, pero solamente los parámetros (ya sea intercept, β o ambos) significativos al nivel $\alpha = 0.05$

Modelo de datos concatenados, con sexo una dummy variable e interacciones sexo*distancia, solo coeficientes (pendientes) significativas (significativamente distintas de cero):

Solución: $\text{TiempoLog} = -2.6921619 + 1.1116747 * \text{DistanciaLog}$

- **Reporta por escrito en papel** si la interacción Sexo*log(Distancia) es significativa. (Nota: esto se interpreta como los efectos de sexo con log(Distancia) no son significativos y podemos suponer que la rapidez es paralela entre hombres y mujeres y ANCOVA es un buen modelo)

Solución: La prueba de hipótesis de la interacción DistanciaLog*Sexo es de que la pendiente es distinta de cero. Como el p-value es 0.983, que es mayor a 0.05, no se rechaza la hipótesis nula y se concluye que la pendiente es cero (a pesar de que numéricamente la pendiente es 0.0005397, esto es, estadísticamente la pendiente es cero, pero numéricamente no lo es).

-
- **Reporta por escrito en papel** tu predicción del tiempo que los hombres correrían la Distancia de 60 y el tiempo que las mujeres correrían la Distancia 200. Ten cuidado, nota que tu modelo está en la escala log, así que calcularás con calculadora el $\log(\text{Distancia de la pregunta})$ y luego tendrás que usar exponencial porque log es base e.

Solución: a. hombres corren la distancia 60 usando el modelo de hombres en 5.64399 y se resuelve así:

TiempoLog= $-2.823196 + 1.112214 * \log(60)$

$\exp(\text{TiempoLog})$ es 5.64399

(Ve el resto de posibilidades en *rm examen2 solucion knit R.pdf* o en *rm examen2 solucion SAS output.pdf*).

- **Reporta por escrito en papel** la *Multiple R-squared*, esta la encuentras al ejecutar *anova* (o en la tabla de resultados en SAS). Este número es deseable que este cerca de 1.00 porque en porcentaje explica que tan bien la variación de Y está explicada por el modelo

Solución: Para el modelo de hombres, la R-Squared es 0.9995 Localiza el resto de las R-squared en *rm examen2 solucion knit R.pdf* o en *rm examen2 solucion SAS output.pdf*.