

Examen Segundo Parcial

⚠ Esta es una vista previa de la versión publicada del examen

Comenzado: 7 de sep en 10:26

Instrucciones del examen

El siguiente examen consta de tres temas, dos de desarrollo en R e interpretación y la tercera sólo de interpretación.

Se debe adjuntar al final un Archivo.R, .rmd, archivo generado .html o archivo .txt . El archivo debe contener las respuestas a las preguntas.

Cuentan con 120 minutos. Distribuya bien su tiempo.

Es un examen a libro abierto, pueden utilizar el material de clase, páginas de internet, etc.

Es un examen individual, está prohibida la comunicación con sus compañeros o cualquier tercera persona durante el examen.

Cualquier acto de deshonestidad será informado a las autoridades competentes.

Al comenzar el examen, el estudiante acepta el siguiente compromiso:

COMPROMISO DE HONOR

Reconozco que el presente examen está diseñado para ser resuelto de manera individual, y no se permite la ayuda de fuentes no autorizadas ni copiar.

TEMA 1

La base de datos [GROWTHGV.csv](#) contiene información de 16 niñas a las que se le midió su altura anualmente desde los 6 a los 10 años.

Las niñas fueron clasificadas según la altura de su madre (pequeñas (1): < 155 cm, medianas (2): 155-164 cm, altas(3): > 164 cm).

Se desea modelar el crecimiento de las niñas, por lo que la variable de respuesta es la altura.

1. Análisis exploratorio gráfico de la evolución de la altura de las niñas.

1.a Realice un gráfico de el crecimiento de cada niña desde los 6 a los 10 años. **(5 pts)**

1.b Realice un gráfico de interacción para la altura promedio entre la edad y el grupo de la madre. **(5 pts)**

1.c En base a los gráficos sugiera ideas de modelación para la altura **(5 pts)**

2. Modelo general

2.a Estime un modelo con intercepto y pendiente aleatoria con correlación entre las dos y con una estructura fija de edad + grupo + la interacción de edad y grupo. Nota: El grupo es un factor fijo, no es aleatorio. **(5 pts)**

3. Inferencias. Para esta sección debe usar el modelo 2 como base y debe basar su respuesta en algún método estadístico válido de inferencia o selección y debe especificar claramente en qué está basando su respuesta.

3.a ¿El crecimiento aleatorio de cada niña depende de su altura a los 6 años? **(5 pts)**

3.b ¿El crecimiento de cada niña de los 6 a los 10 años es distinto? o ¿se puede decir que en promedio tienen el mismo patrón de crecimiento? **(5 pts)**

3.c ¿El crecimiento de las niñas en promedio depende de qué tan altas son las madres? **(5 pts)**

3.d ¿Habría una diferencia importante entre la altura promedio de cada grupo de niñas a los 6 años? **(5 pts)**

TEMA 2

La base de datos [aerobics.txt](#) contiene datos sobre un estudio de una muestra de personas de 65 años o más que gozan de una salud razonablemente buena. La pregunta de investigación es si una sesión informativa y/o ejercicios influyen en el número de caídas que pueden sufrir estas personas. En el estudio participaron 100 personas. Fueron asignados aleatoriamente a uno de dos grupos. El grupo 1 asistió únicamente a la sesión informativa ($x_1 = 0$), mientras que el otro grupo asistió a la sesión informativa además de realizar ejercicios aeróbicos ($x_1 = 1$). El estudio también recopiló información sobre el género ($x_2 = 0$ para mujer, y $x_2 = 1$ para hombre), un índice de equilibrio x_3 , cuanto mayor es este valor, más estable es la persona, un índice de fuerza x_4 , cuanto mayor es este valor, más fuerte es la persona. La variable respuesta es el número de caídas que realizó la persona durante los seis meses de estudio.

1. Escoja un subconjunto aleatorio de 25 personas en cada grupo. Para la generación aleatoria de la muestra utilice como semilla su número de matrícula de la ESPOL. Ése será el set de datos con el que usted trabajará en los literales 2 al 7.

```
df=read.table("aerobics.txt", header = T)
set.seed(200407814) # aquí pone su número de matrícula
sp=sample(1:50,25) # esta es la generación aleatoria de la muestra

df1=df[df$x1==1,]

df2=df[df$x1==0,]

data1=rbind(df[sp,],df2[sp,])#este es su nuevo set de datos
```

2. Estime un modelo lineal generalizado adecuado con el enlace canónico. Utilice todos los efectos principales y sólo la interacción entre grupo y género. **(10 pts)**

3. Realice una prueba de hipótesis formal para probar si la variable género y su interacción con el grupo no son importantes para explicar el número de caídas. Debe incluir el contraste de hipótesis con notación estadística correcta, identificando cada coeficiente beta. Concluya a un 5% de significancia **(10 pts)**

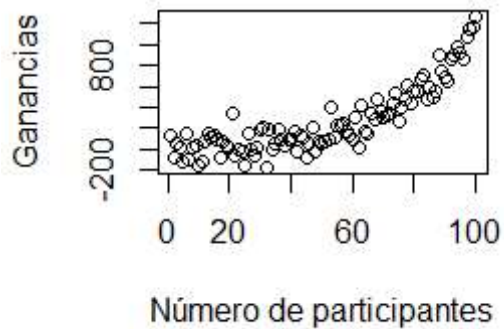
4. Estime el modelo simplificado (sin género ni la interacción con el grupo) en el grupo de datos que no utilizó previamente.

`data2=rbind(df[-sp,],df2[-sp,])` #este es su nuevo set de datos

Obtenga intervalos de confianza para el coeficiente estimado del grupo. Interpretelo. ¿Se puede recomendar que todas las personas de este grupo de edad realicen ejercicios aeróbicos? Indique en qué basa su respuesta **(10 pts)**

TEMA 3

El siguiente gráfico muestra los datos de ganancias de un negocio piramidal mientras más participantes se unen. Para estimar la curva se usó un modelo de crecimiento logístico. Se muestra además el resumen del modelo estimado.



Formula: $y \sim \text{SSlogis}(t, \text{phi1}, \text{phi2}, \text{phi3})$

Parameters:

	Estimate	Std. Error	t value	Pr(> t)
phi1	12981.688	43064.484	0.301	0.764
phi2	149.947	86.163	1.740	0.085 .
phi3	21.774	3.415	6.376	6.17e-09 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 107.1 on 97 degrees of freedom

Number of iterations to convergence: 0

Achieved convergence tolerance: 2.533e-07

a. ¿A partir de cuántos participantes se estima que el crecimiento de las ganancias del negocio piramidal empiece a desacelerarse? (10 pts)

b. ¿Es esta desaceleración significativa? Utilice un 5% de nivel de significancia. Indique en qué específicamente está basando su respuesta. **(10 pts)**

c. ¿El crecimiento de las ganancias de este negocio se explicaría mejor con un modelo de crecimiento exponencial o logístico? Utilice un 5% de nivel de significancia. Indique en qué específicamente está basando su respuesta. **(10 pts)**

Pregunta 1

100 pts

Aquí debe subir el archivo de resolución con las respuestas a los 3 temas

TEMA 1 - 40 PTS

TEMA 2 -30 PTS

TEMA 3 - 30 PTS

Cargar

Elegir un archivo

Examen guardado en 10:27

Entregar examen