

Entregar a más tardar hoy a las 21:00

Karla y/o yo estaremos en línea de 17:00 a 21:00 por Skype en letygra y karlyjuarez.

Abriremos un zoom de 5:20 a 6:00 para preguntas,
ID de la reunión 443-676-867 la contraseña es 011590

1. Se registraron todos los tiros libres de las temporadas 1980-81 y 1981-82 del jugador de basquetbol Larry Bird.

tiros libres		
primer tiro	segundo tiro	
	encesta	falla
encesta	251	34
falla	48	5

- a) ¿Qué esquema de muestreo corresponde a este caso? Explica porqué.
- b) Calcula las probabilidades de **encestar el segundo tiro**, p_{e1} cuando encesta el primero y p_{f1} cuando falla el primero.
- c) Calcula el OR y su intervalo de confianza.
- d) Calcula el RR y su intervalo de confianza.
- e) ¿Para Larry Bird la probabilidad de encestar el segundo tiro, cambiaba según el resultado del primer tiro?
2. ¿Da lo mismo al momento de estimar modelos (de regresión por ejemplo logística) sustituir una variable explicativa ordinal de 5 niveles por una de tipo numérico?, (por ejemplo sustituir la variable severidad del DAÑO con etiquetas: pérdida total, muy dañado, algo dañado, muy poco dañado, sin daño, sustituirla por una numérica con valores 1,2,3,4,5)¿Qué modelo ajustaría mejor? Comenta el porqué de tu respuesta.
3. La siguiente tabla contiene información acerca del gusto por cierta película y el grado escolar que cursan los niños encuestados.

Encuesta Niños		
grado escolar	Gusto por la película	
	sí	no
1	10	4
4	6	11

¿El gusto por la película puede decirse que es el mismo para ambos grados? Trabaja con un $\alpha = 0,10$

4. Para la tabla de contingencias de cinco actividades en cuatro lugares distintos, se hizo un análisis de correspondencias. (se anexa salida de R)
- a) ¿Qué lugar y qué actividad son las de mayor masa?, ¿qué nos dice eso?
- b) ¿El biplot qué tan bien representa a la tabla?, ¿todos los puntos están bien representados? Justifica tus respuestas.
- c) Haz una interpretación del Biplot.
- d) ¿Cómo se puede obtener el valor de la prueba de la ji-cuadra a partir de la salida de correspondencias?
5. La base de datos pima del paquete faraway de R, contiene datos de 768 mujeres adultas. (library(faraway);data(pima))

- a) Para la variable respuesta *test* ajusta el modelo logístico que resulte del comando *step*.
(res.diabetes < -glm(test insulin+glucose+diastolic+triceps+insulin+bmi+diabetes , family=binomial,
data= pima)
diabetesfinal < -step(res.diabetes))
- b) Escribe la fórmula del modelo ajustado.
- c) ¿Los datos son agrupados o desagrupados? Justifica tu respuesta
- d) ¿El modelo ajusta bien? Justifica tu respuesta.
- e) Interpreta los coeficientes del modelo
- f) Ajusta un modelo similar con liga probit .
- g) Escribe la fórmula del modelo ajustado.
- h) Ajusta un modelo similar con liga cloglog.
- i) Escribe la fórmula del modelo ajustado.
- j) ¿Cuál liga te parece da mejor ajuste. Justifica tu respuesta con alguna estadística, alguna gráfica o análisis de residuales.