

UNIVERSIDAD NACIONAL DE COLOMBIA

FACULTAD DE CIENCIAS, SEDE BOGOTÁ  
DEPARTAMENTO DE ESTADÍSTICA  
ANÁLISIS DE REGRESIÓN

Ejercicio en clase 3. Regresión lineal múltiple.

1. En el archivo **atletas.xls** se encuentran los datos de 98 mujeres atletas de élite que fueron entrenadas en el Instituto Australiano del Deporte. El objetivo del análisis de este conjunto de datos es explicar el comportamiento de la masa corporal magra (**lbm**) a partir de la estatura en centímetros (**ht**), el peso en kilogramos (**wt**) y el conteo de glóbulos rojos (**rcc**).

a) Haga un análisis descriptivo del conjunto de datos, es decir, construya diagramas de dispersión de cada variable explicativa con la variable respuesta. Calcule las correlaciones simples y parciales entre las variables. Comente.

b) Formule un modelo lineal que permita explicar la masa corporal magra en función de las otras variables.

c) Estime e INTERPRETE los parámetros de localización y de escala del modelo escogido.

d) Compruebe, haciendo uso de la forma matricial del modelo, que efectivamente las estimaciones de los parámetros corresponden con las proporcionadas por R.

e) Calcule la matriz de varianzas y covarianzas estimada de los parámetros de localización y compare con la que proporciona el modelo.

f) Calcule la matriz de proyección, **H**, y verifique que es simétrica e idempotente.

g) Calcule las estimativas del valor medio estimado para todos los individuos de la muestra.

h) Calcule el vector de residuales del modelo y verifique que al proyectarlos con la matriz **H** se obtiene el vector **0**. ¿Qué significa ese resultado? ¿Cuál es la media de los residuales?

i) ¿Es el modelo de regresión significativo para explicar la variabilidad presentada en **lbm**?

j) ¿La masa corporal magra depende del peso del atleta?

k) ¿La masa corporal magra depende de la estatura del atleta?

l) ¿La masa corporal magra depende del conteo de glóbulos rojos del atleta?

m) ¿Son necesarias las variables de estatura y de conteo de glóbulos rojos para explicar el comportamiento de la masa corporal magra?

n) ¿Se podría decir que el efecto del aumento de 1kg en la masa del atleta es igual al efecto del aumento de 1cm en la altura?

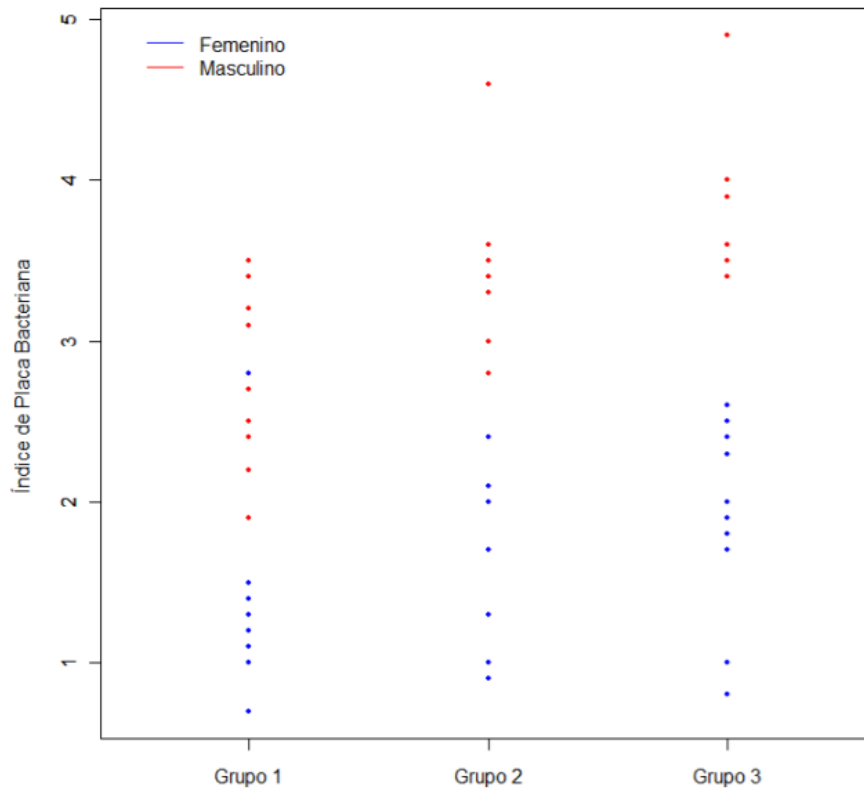
o) Calcule un intervalo del 95% de confianza para el valor medio de individuos con una estatura de 170cm, 70kg y con un conteo de 4.5 glóbulos rojos.

p) Calcule e interprete el coeficiente de determinación y el coeficiente de determinación ajustado.

q) ¿Si se incluye un término cuadrático de la altura se tiene una ganancia significativa?

2. En cierta universidad del Brasil se midió el Índice de Placa Bacteriana antes del cepillado a estudiantes de tres grupos poblacionales, separando por sexo. Los datos del Índice de Placa Bacteriana se encuentran guardados en **ipb**, el grupo en **grupo** y el sexo en **sexo**. (**ipb.xls**)

**UNIVERSIDAD NACIONAL DE COLOMBIA**  
**FACULTAD DE CIENCIAS, SEDE BOGOTÁ**  
**DEPARTAMENTO DE ESTADÍSTICA**  
**ANÁLISIS DE REGRESIÓN**



- Calcule y comente las estadísticas de resumen más importantes (*min*, *Q1*, *mediana*, *media*, *Q3*, *max*, *desviación estándar*) para el Índice de Placa Bacteriana en cada combinación de grupo y sexo.
- Proponga un modelo estadístico para explicar el Índice de Placa Bacteriana de los estudiantes a partir del grupo y el sexo de los mismos. Describa las componentes aleatoria y sistemática e INTERPRETE los parámetros del modelo propuesto.
- Estime e INTERPRETE los parámetros (localización y escala) del modelo propuesto.
- ¿Cree usted que la muestra proporciona evidencia estadísticamente significativa de que el Índice de Placa Bacteriana de los estudiantes depende del grupo poblacional? Justifique su respuesta.
- ¿Cree usted que la muestra proporciona evidencia estadísticamente significativa de que el Índice de Placa Bacteriana de los estudiantes depende del sexo? Justifique su respuesta.
- ¿Cree usted que la muestra proporciona evidencia estadísticamente significativa de que el efecto del grupo poblacional sobre el Índice de Placa Bacteriana de los estudiantes depende del sexo? Justifique su respuesta.
- ¿Cree usted que la muestra proporciona evidencia estadísticamente significativa de que el efecto del grupo poblacional sobre el Índice de Placa Bacteriana es tal que el efecto del grupo 3 es el doble del del grupo 2? Justifique su respuesta.