



Pontificia Universidad  
**JAVERIANA**  
Bogotá

MAESTRÍA EN   
**EPIDEMIOLOGÍA**  
CLÍNICA

**BIOESTADÍSTICA AVANZADA**

## MÓDULO I

**Semana 4**

Diagnóstico del modelo

## Práctica semana 4

Carlos Javier Rincón R.

### Escenario de aplicación

Se registra la información de 38 fetos sobre las siguientes variables:

- LCC: longitud cráneo-caudal en milímetros (mm).
- DBP: diámetro biparietal en milímetros (mm).
- DOF: diámetro occipitofrontal en milímetros (mm).
- Distrito: Lugar de residencia de la madre (Distrito1, Distrito2, Distrito3, Distrito4).
- sexo: sexo del feto (H,M).

Los datos se puede crear a partir del siguiente código:

```
id<-c(1:38) # Identificador del sujeto
LCC<-c(66,50.8,61.8,61.6,53,74.1,48.1,57,57,72.6,66.4,64.3,51.1,
      52.3,58.2,57.6,48,77.9,68,69.7,61.8,73,55.3,65.2,72,61.8,74.6,
      67,83,69.3,56.8,51,48.5,72.9,77.1,75.4,71.4,59.6)
DBP<-c(22,16.5,21,26.1,14,22.9,15.8,18,21,26.1,23.1,18.2,16.7,
      18.9,19.8,20,18.1,25.7,21,21.7,21.5,21,17,19.7,23,20.5,26.3,
      19,26,21.6,19.7,15.4,16,23,24.5,21.7,24.6,18.3)
DOF<-c(28,21.9,26.8,27,22,28.3,21.5,23,27,31.4,27.3,23.2,20.7,
      24.6,24,23.5,22.6,31.3,28,25.9,27.8,29,23.2,24.4,30,25.9,33.3,
      28,35,27.3,24.9,18.4,21,32.2,31.5,28.1,31.7,24.4)
Distrito<-c("Distrito4", "Distrito3", "Distrito3", "Distrito3",
            "Distrito3", "Distrito3", "Distrito3", "Distrito2", "Distrito3",
            "Distrito3", "Distrito2", "Distrito2", "Distrito3", "Distrito2",
            "Distrito3", "Distrito3", "Distrito2", "Distrito2", "Distrito2",
            "Distrito1", "Distrito2", "Distrito2", "Distrito2", "Distrito2",
            "Distrito3", "Distrito1", "Distrito1", "Distrito2", "Distrito2",
            "Distrito4", "Distrito1", "Distrito1", "Distrito1", "Distrito1",
            "Distrito1", "Distrito1", "Distrito1", "Distrito1")
sexo<-c("M", "M", "M", "M", "H", "H", "H", "M", "M", "M", "H", "H", "M", "H", "H",
        "M", "H", "H", "H", "H", "M", "H", "M", "M", "H", "M", "H", "M", "M", "M", "H",
        "M", "H", "H", "M", "H", "H", "M")
```

```
bd<-data.frame(id,DBP,LCC,DOF,Distrito,sexo) # Base de datos

# Definición del valor de referencia para las variables distrito y sexo.

bd$Distrito<-factor(bd$Distrito,levels = c("Distrito3","Distrito1",
                                           "Distrito2","Distrito4"))
bd$sexo<-factor(bd$sexo,levels = c("M","H"))
```

Con base en esta muestra de fetos, se ajustó una regresión línea múltiple tomando como desenlace DOF, y como variables independientes el resto de variables incluidas en la base de datos. Se definió como categorías de comparación en las variables categóricas “Distrito3” para la variable Distrito y “M” para la variable sexo.

```
m<-lm(DOF~LCC+DBP+Distrito+sexo,data=bd) # objeto con el modelo ajustado.
summary(m)
```

```
##
## Call:
## lm(formula = DOF ~ LCC + DBP + Distrito + sexo, data = bd)
##
## Residuals:
## Min 1Q Median 3Q Max
## -2.4051 -1.1102 0.2723 0.9105 2.4936
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.57344 1.59770 0.985 0.332337
## LCC 0.18680 0.04706 3.969 0.000398 ***
## DBP 0.61811 0.13174 4.692 5.18e-05 ***
## DistritoDistrito1 0.10338 0.62034 0.167 0.868729
## DistritoDistrito2 0.28130 0.59603 0.472 0.640260
## DistritoDistrito4 -0.03511 1.12405 -0.031 0.975282
## sexoH 0.19552 0.48280 0.405 0.688283
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.411 on 31 degrees of freedom
## Multiple R-squared: 0.8899, Adjusted R-squared: 0.8686
## F-statistic: 41.78 on 6 and 31 DF, p-value: 1.621e-13
```

Con base en los resultados del modelo ajustado, desarrolle las siguientes actividades:

1. Actividad 1. Identifique los datos atípicos utilizando las medidas: leverage, residuales estandarizados, residuales estudentizados, residuales de Jackknife y la distancia de Cook.
2. Actividad 2. Evalúe los supuestos de linealidad y de homocedasticidad del modelo.
3. Actividad 3. Evalúe el supuesto de normalidad sobre los residuales del modelo.
4. Actividad 4. Evalúe posibles problemas de colinealidad entre las variables independientes.