

Caso Estudio

Regresión con R

CNE/ISCIH

1 Preambulo

1.1 Objetivo

A partir de los controles del estudio de casos y controles EURAMIC se pretende evaluar el efecto de distintos factores sobre el riesgo de hipertensión en hombres adultos. La base incluye las siguientes variables:

- **statcc:** 1 “caso de infarto”, 0 “control”
- **edad:** edad del paciente en años
- **peso:** peso en kg
- **altura:** altura en cm
- **tipfum:** 1 “no fumador”, 2 “ex fumador”, 3 “fumador actual”
- **alcohol:** consumo de alcohol en g/día (0 si no es bebedor actual)
- **hta:** 1 “hipertenso”, 0 “normotenso”
- **diabetes:** 1 “diabético”, 0 “no diabético”
- **hdlcol:** colesterol HDL en mmol/l
- **totcol:** colesterol Total en mmol/l

1.2 Importación de los datos

```
euramic = read.csv("data/euramic.csv")
euramic$tipfum = factor(euramic$tipfum, levels = 1:3, labels = c("nunca", "ex", "activo"))
str(euramic)
```

```
## 'data.frame':    1339 obs. of  10 variables:
## $ statcc : int  1 1 1 1 1 1 1 1 1 1 ...
## $ edad   : int  67 66 51 52 69 43 61 47 66 60 ...
## $ peso   : int  62 70 68 70 72 78 81 90 84 103 ...
## $ altura : int  164 178 158 172 173 172 175 173 165 180 ...
## $ tipfum  : Factor w/ 3 levels "nunca","ex","activo": 1 3 1 2 3 3 3 3 3 2 ...
## $ alcohol : int  0 12 21 80 0 32 4 96 216 73 ...
## $ hta     : int  0 0 0 0 0 1 0 0 0 1 ...
## $ diabetes: int  0 0 0 0 0 0 0 0 0 0 ...
## $ hdlcol  : num  0.89 NA 1.58 0.79 1.29 ...
## $ totcol  : num  6.29 NA 6.96 5.01 4.79 ...
```

2 Correlación y regresión simple

Utilizando la muestra de los controles (representativos de la población general,

1. Representar en un diagrama de dispersión el colesterol HDL y el índice de masa corporal ($\text{imc} = \text{peso (kg)}/\text{altura (m)}^2$).
2. Evaluar la asociación lineal entre estas dos variables (interpretar)
3. Estimar mediante un modelo de regresión lineal el efecto del IMC sobre el nivel de HDL e interpretarlo
4. Chequear las asunciones del modelo
5. Calcular el intervalo de confianza de este efecto (interpretar)
6. Representar la recta de regresión

3 Regresión múltiple

1. Evaluar la no-linealidad del efecto del IMC sobre el nivel de HDL
2. Estimar el efecto del IMC ajustando por el hábito tabáquico (interpretar)
3. Contrastar si el efecto del IMC está modificado por el consumo de tabaco.
4. Representar en un grafico, la relación entre HDL e IMC según el hábito tabáquico

4 Regresión logística

1. Evaluar la asociación (OR) entre ser fumador actual y el riesgo de hipertensión
2. Estimar el efecto crudo del colesterol HDL y su efecto ajustado por IMC, edad y hábito tabaquico
3. Dar la tabla de resultados de esta regresión logística en un formato apto para publicación
4. Representar el riesgo de hipertension en función del IMC para hombres de 30, 50 y 70 años