

Bioestadística y uso de software científico



TEMA 11 MÉTODOS MULTIVARIABLES

Algunos ejemplos



- En la regresión lineal hemos visto la relación entre dos variables (x e y).
- En el mundo real, 2 variables no se presentan aisladas
- Queremos conocer cómo influyen varias variables (edad, sexo, tabaquismo,...) en el nivel de tensión arterial
- Para esto usamos modelos multivariables

Índice



- Regresión lineal múltiple
- Regresión logística
- Regresión de Poisson
- Regresión de Cox
- Otros métodos multivariados

Regresión lineal múltiple



- Regresión lineal simple

$$y = \alpha + \beta x$$

$$tensión\ arterial = \alpha + \beta \times colesterol$$

- Regresión lineal múltiple

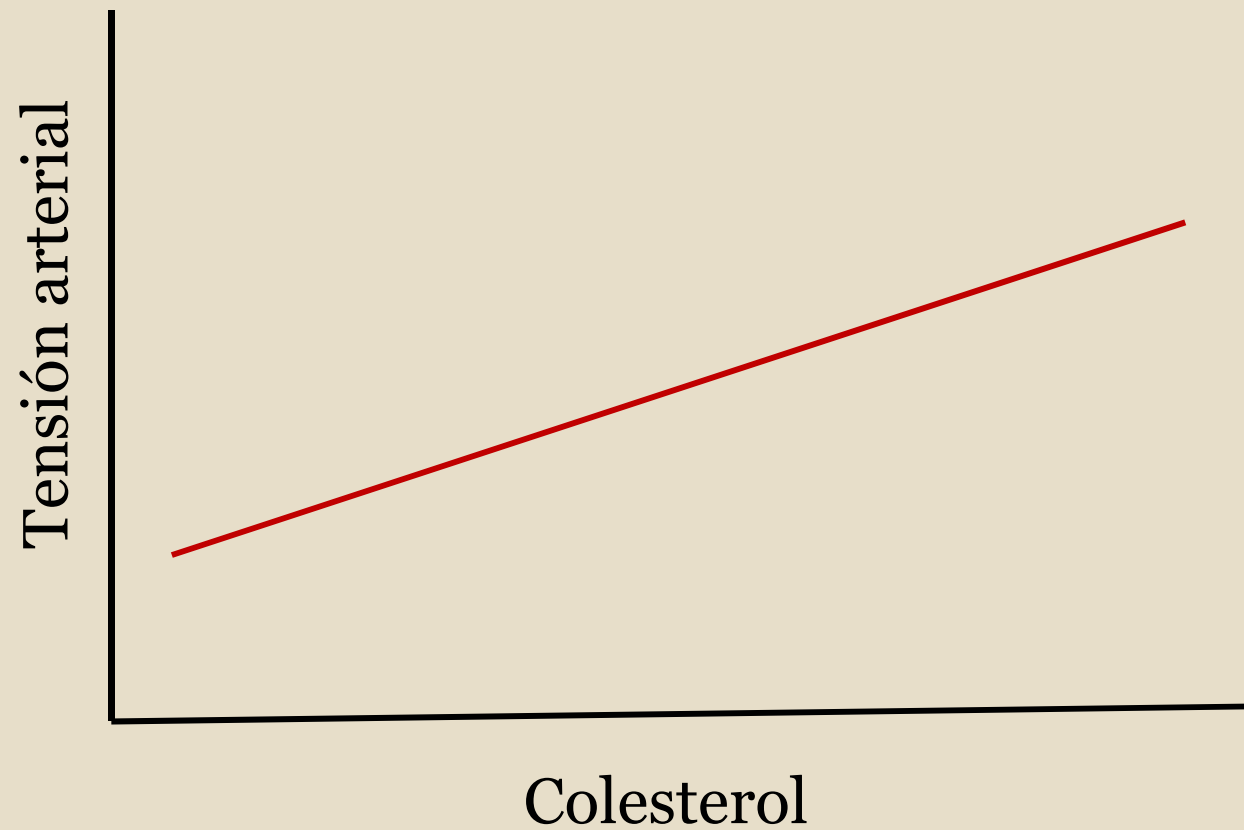
$$y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$$

$$t.a. = \alpha + \beta_1 \times colesterol + \beta_2 \times sexo$$

Regresión lineal múltiple



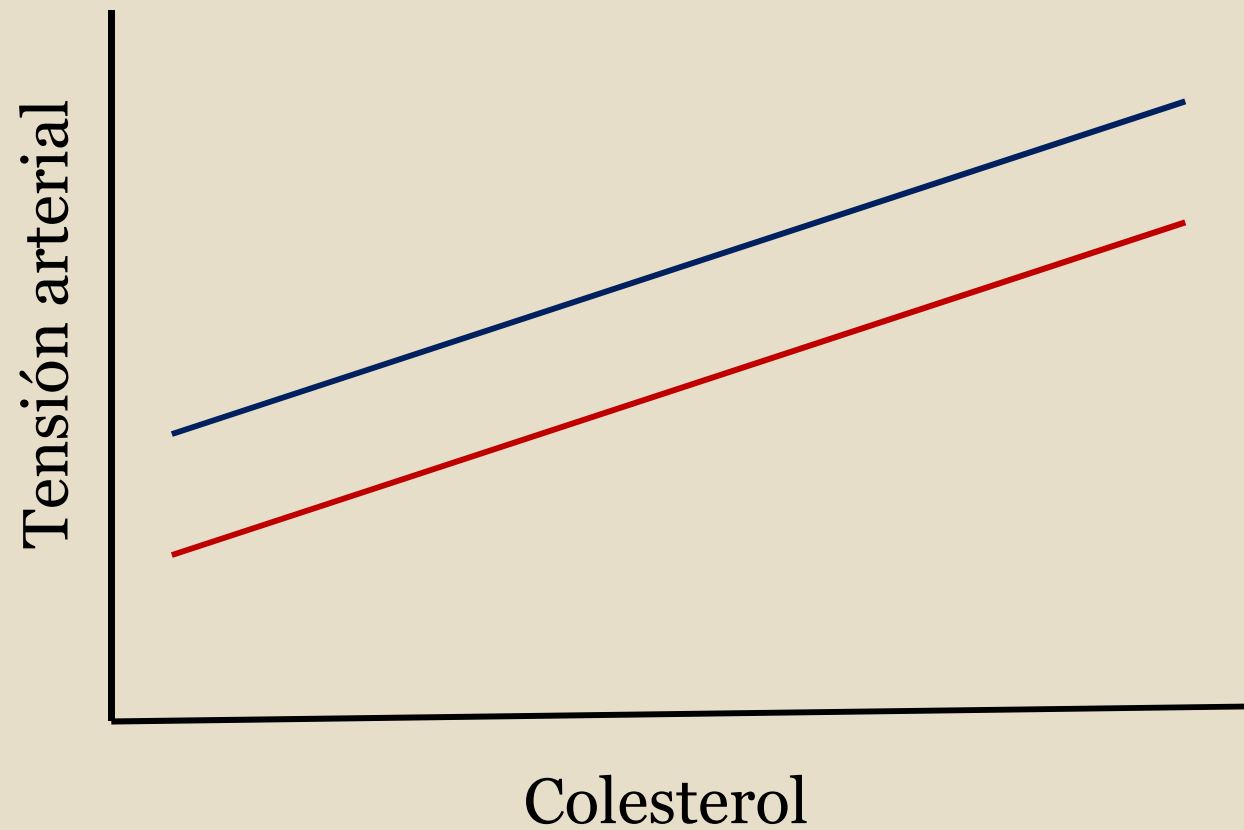
- Regresión lineal simple



Regresión lineal múltiple



- Regresión lineal múltiple



Regresión lineal múltiple



- Regresión lineal múltiple

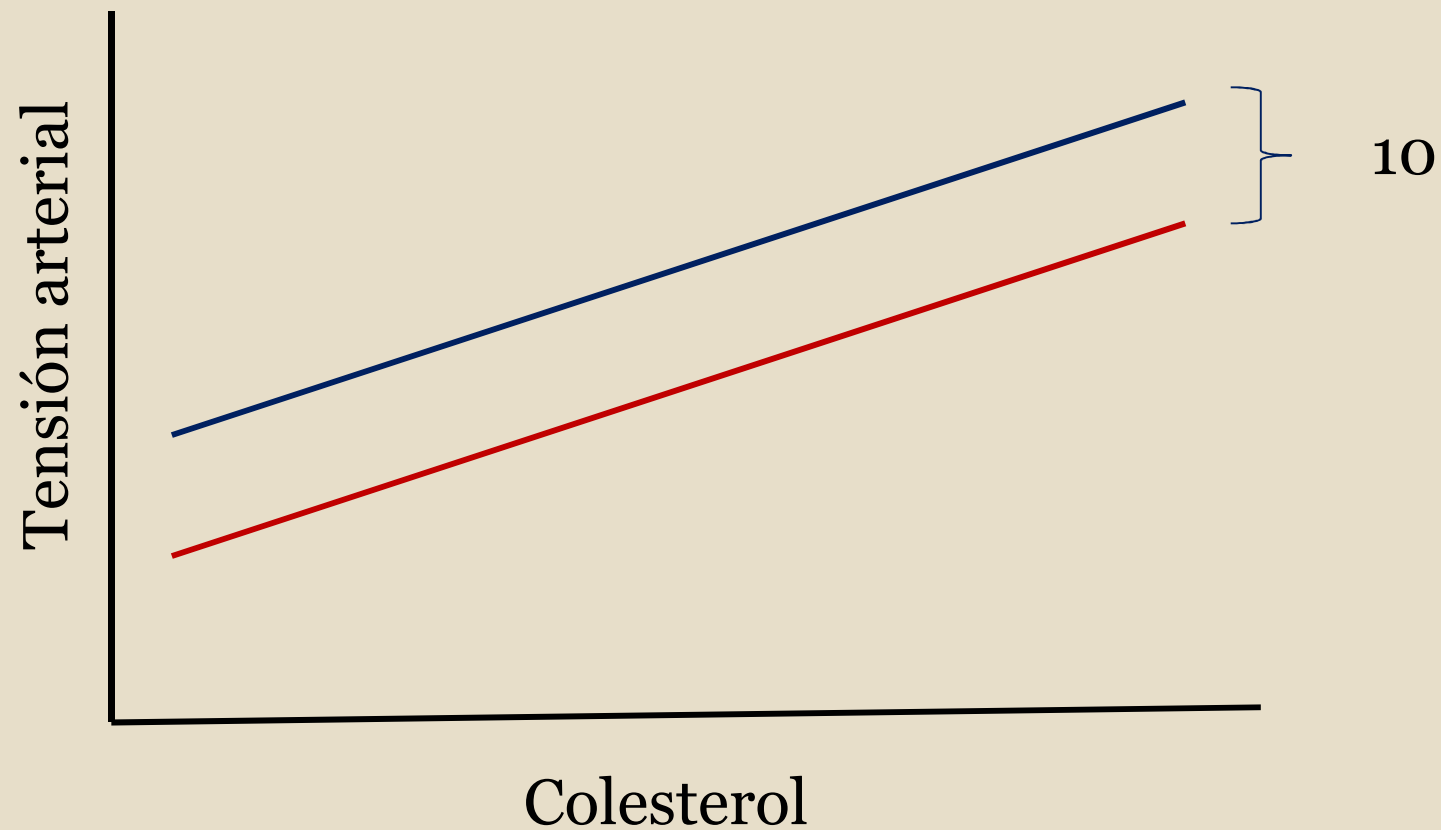
$$t.a. = \alpha + \beta_1 \times colesteroles + \beta_2 \times sexo$$

- ¿Cómo se hacen los cálculos con el sexo?
 - Se crea una variable numérica: varón=1, mujer=0

Regresión lineal múltiple



$$t.a. = \alpha + \beta_1 \times \text{colesterol} + 10 \times \text{varón}$$



Regresión lineal múltiple



- Regresión lineal múltiple

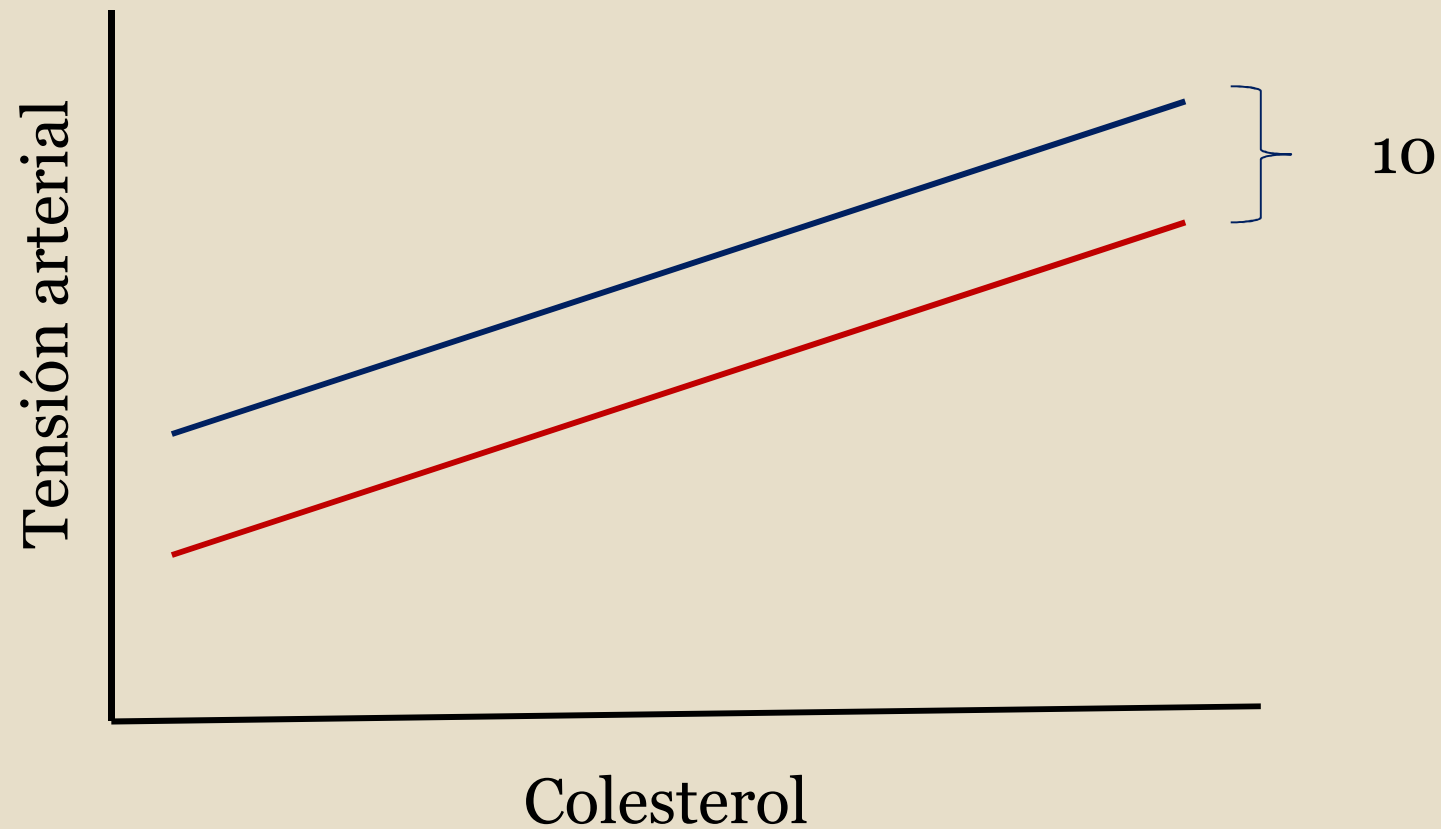
$$t.a. = \alpha + \beta_1 \times colesteroles + \beta_2 \times sexo$$

- ¿Qué pasa si la variable se codifica al revés?
 - mujer=1, varón=0

Regresión lineal múltiple



$$t.a. = \alpha + \beta_1 \times \text{colesterol} - 10 \times \text{mujer}$$



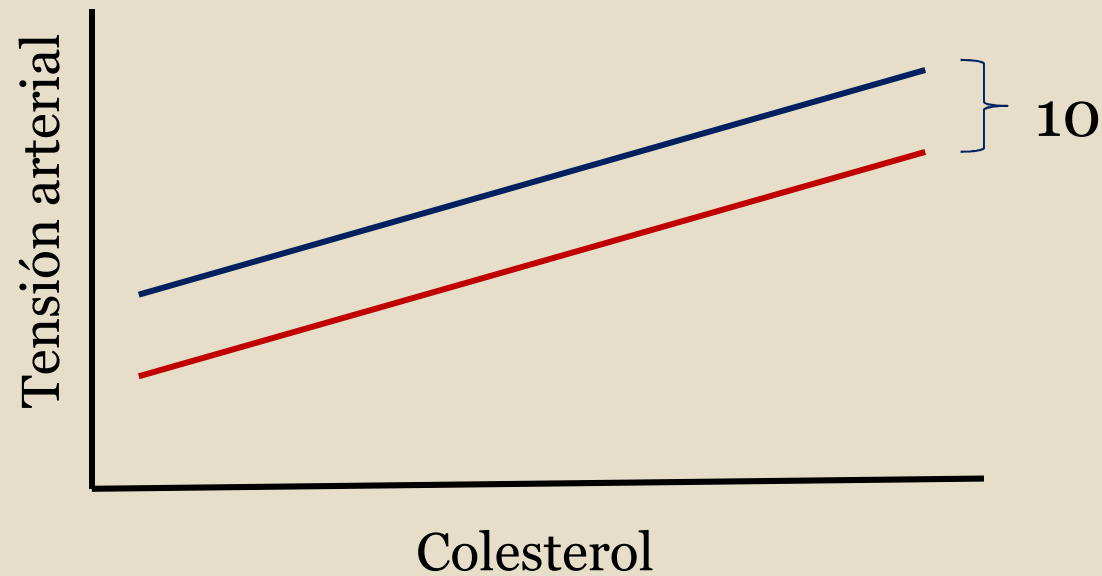
Regresión lineal múltiple



Da igual cómo se codifique: los dos modelos son iguales

$$t.a. = \alpha + \beta_1 \times \text{colesterol} - 10 \times \text{mujer}$$

$$t.a. = \alpha + \beta_1 \times \text{colesterol} + 10 \times \text{varón}$$



Regresión lineal múltiple



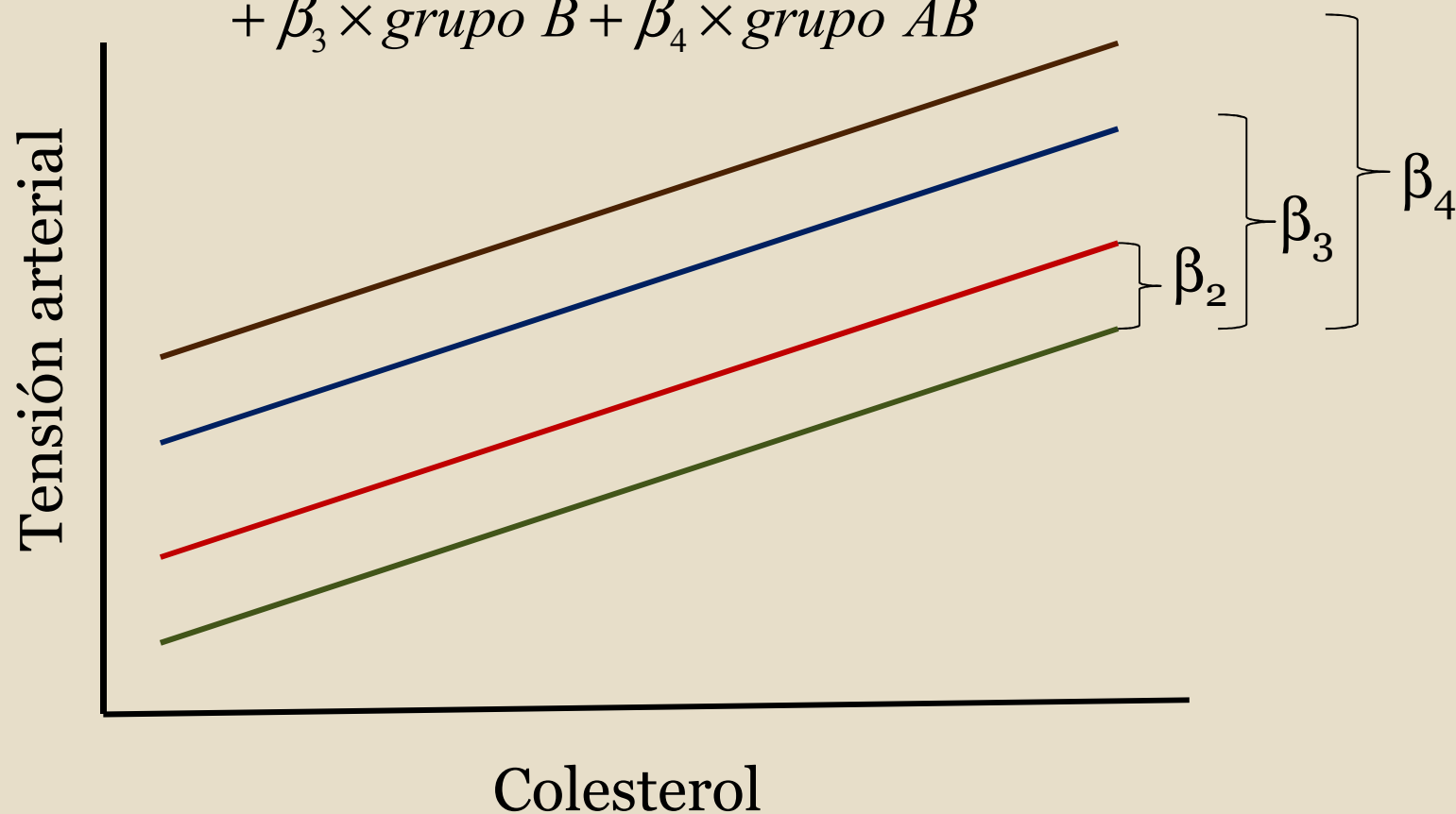
- ¿Cómo se hacen los cálculos si son $k > 2$ categorías?
 - Se crean $k-1$ variables dicotómicas(0/1)
 - ✦ Grupo A: sí=1, no=0
 - ✦ Grupo B: sí=1, no=0
 - ✦ Grupo AB: sí=1, no=0

$$t.a. = \alpha + \beta_1 \times col. + \beta_2 \times grupo A + \\ + \beta_3 \times grupo B + \beta_4 \times grupo AB$$

Regresión lineal múltiple



$$t.a. = \alpha + \beta_1 \times col. + \beta_2 \times grupo A + \\ + \beta_3 \times grupo B + \beta_4 \times grupo AB$$



Regresión lineal múltiple



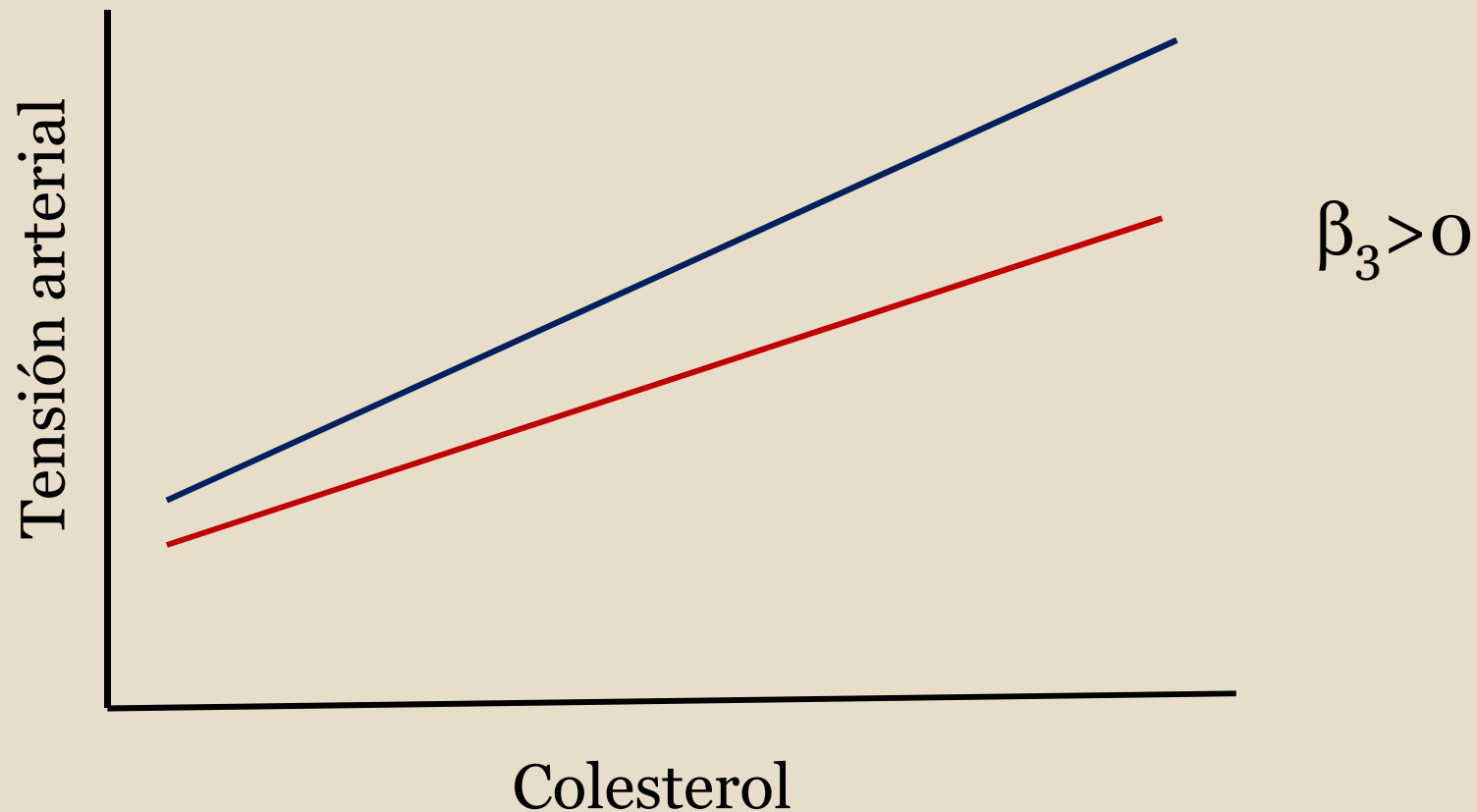
- Medir la interacción entre dos regresores

$$t.a. = \alpha + \beta_1 \times \text{colesterol} + \beta_2 \times \text{varón} + \beta_3 \times \text{colesterol} \times \text{varón}$$

Regresión lineal múltiple



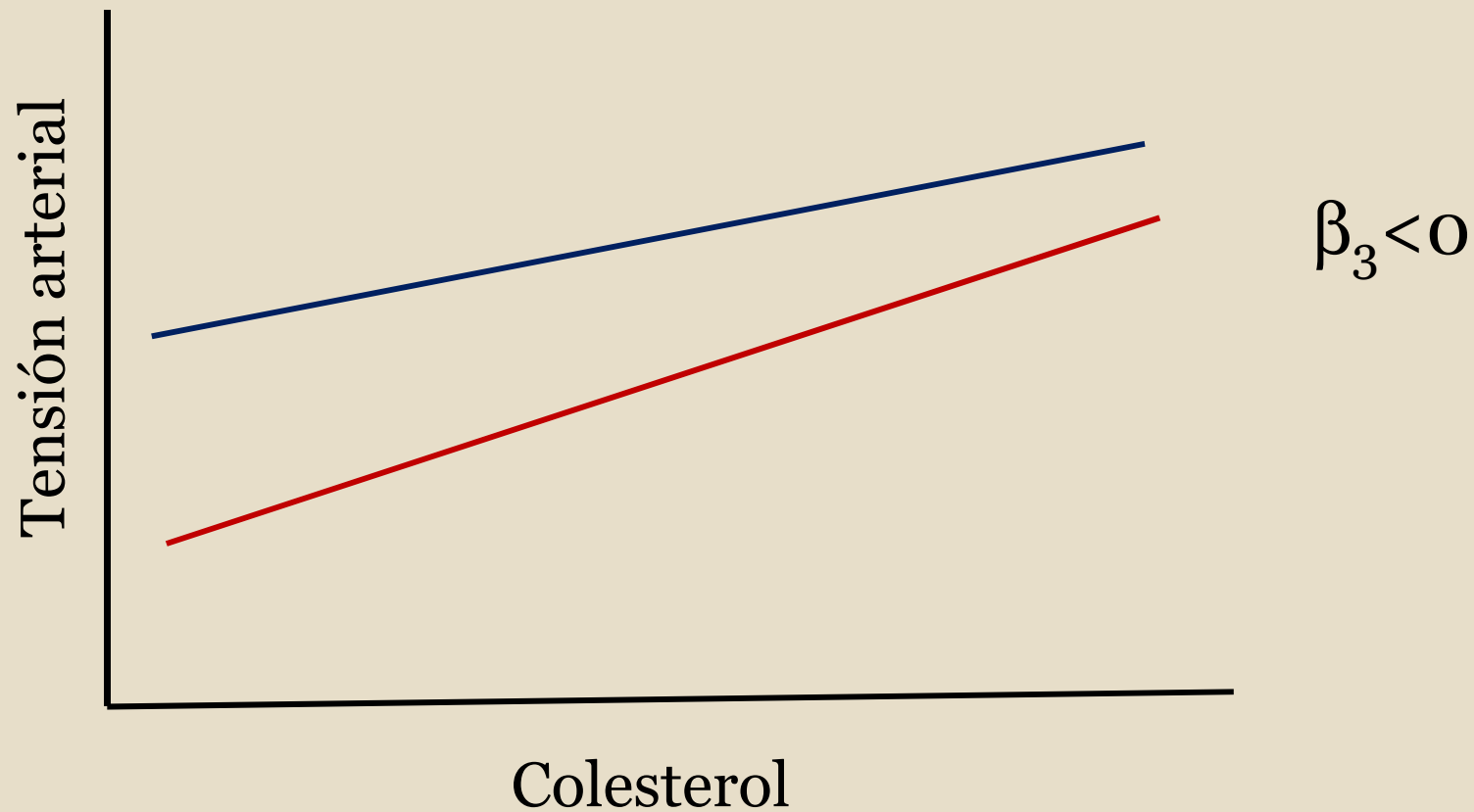
$$t.a. = \alpha + \beta_1 \times \text{colesterol} + \beta_2 \times \text{varón} + \beta_3 \times \text{colesterol} \times \text{varón}$$



Regresión lineal múltiple



$$t.a. = \alpha + \beta_1 \times \text{colesterol} + \beta_2 \times \text{varón} + \beta_3 \times \text{colesterol} \times \text{varón}$$



Otros modelos de regresión



Regresión	Variable dependiente
Lineal	Cuantitativa
Logística	Dicotómica
Poisson	Número de...
Cox	Dicotómica + tiempo

Índice



- Regresión lineal múltiple
- Regresión logística
- Regresión de Poisson
- Regresión de Cox
- Otros métodos multivariados

Regresión logística



- Variable dependiente:
 - Enferma/no enferma
 - Muere/no muere
 - Gana la liga/no gana la liga

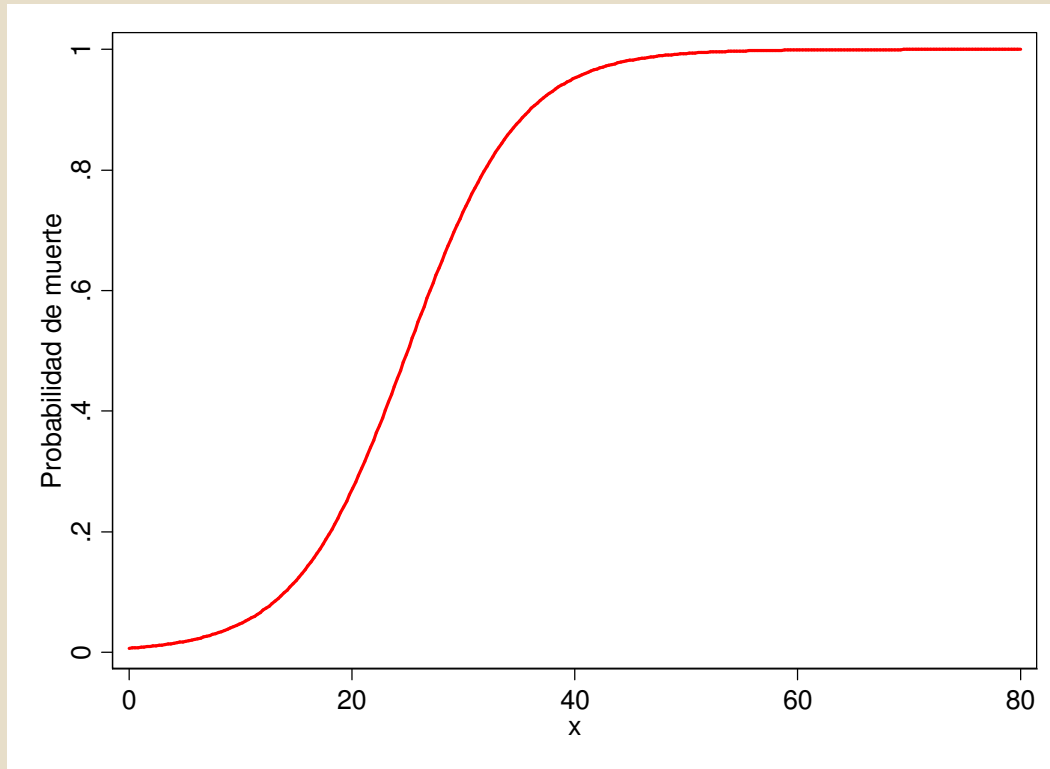
$$\ln \frac{p}{1-p} = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$$

$p = \text{probabilidad de... (enfermar / morir / ...)}$

Regresión logística



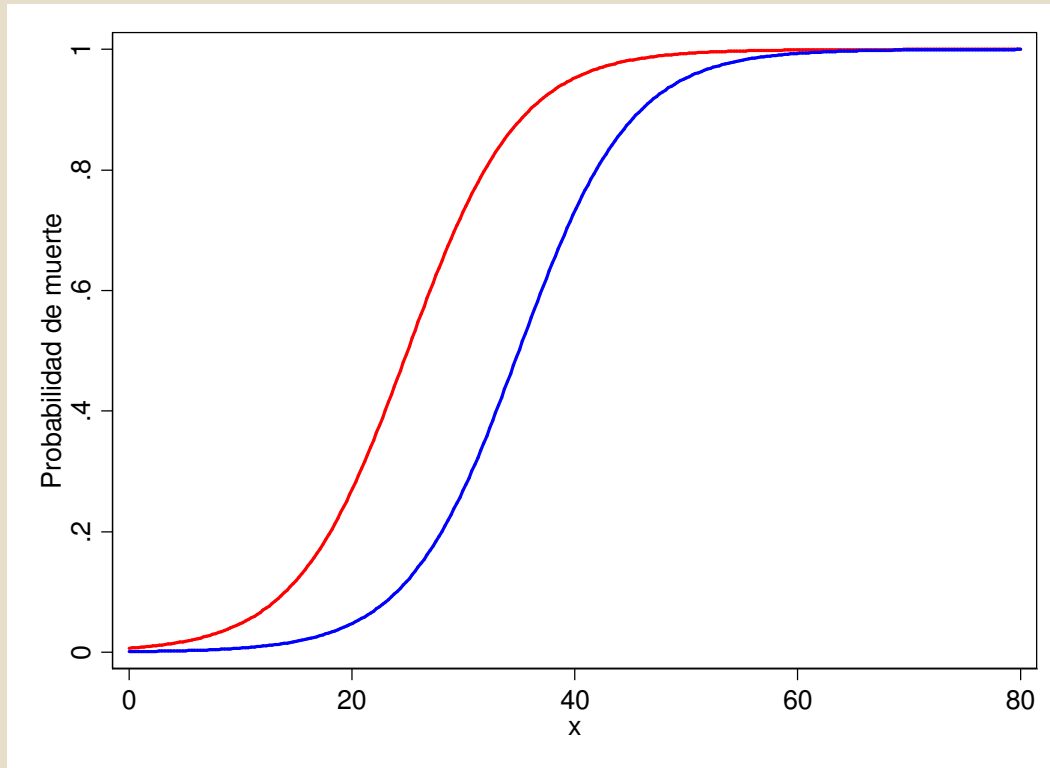
$$\ln \frac{p}{1-p} = \alpha + \beta_1 x_1$$



Regresión logística



$$\ln \frac{p}{1-p} = \alpha + \beta_1 x_1 + \beta_2 \times \text{mujer}$$



Índice



- Regresión lineal múltiple
- Regresión logística
- **Regresión de Poisson**
- Regresión de Cox
- Otros métodos multivariados

Regresión de Poisson



- Variable dependiente:
 - Número de casos de infarto en un año
 - Número de infecciones en cada colegio
 - Número de desintegraciones por hora en una masa radioactiva

$$\ln(y) = \ln(N) + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$$

$N = n^\circ$ de habitantes, n° de alumnos, n° de átomos

Índice



- Regresión lineal múltiple
- Regresión logística
- Regresión de Poisson
- Regresión de Cox
- Otros métodos multivariados

Regresión de Cox



- Variable dependiente:
 - Tiempo hasta la muerte
 - Tiempo hasta el primer infarto
 - Tiempo hasta aprobar bioestadística

$$\ln(\lambda_t) = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$$

λ_t = tasa instantánea de...(enfermar/morir/...) en el tiempo t

Índice



- Regresión lineal múltiple
- Regresión logística
- Regresión de Poisson
- Regresión de Cox
- Otros métodos multivariados

Otros métodos multivariados



- Análisis factorial (componentes principales)
 - Reduce el número de variables
- Análisis de clusters
 - Agrupa los individuos en 2/3/4/... Grupos
- ANCOVA (Análisis de la covarianza)
 - Una variante del ANOVA
- MANOVA
 - ANOVA con varias variables resultado