

# MODELOS LINEALES

## INTERVALOS DE CONFIANZA

Fernando Massa; Bruno Bellagamba

16 de abril 2024



FACULTAD DE  
CIENCIAS ECONÓMICAS  
Y DE ADMINISTRACIÓN

**UESTA** INSTITUTO  
DE ESTADÍSTICA



UNIVERSIDAD  
DE LA REPÚBLICA  
URUGUAY



- 1 INTRODUCCIÓN
- 2 ESTADÍSTICO GENERAL
- 3 RESPUESTA MEDIA
- 4 FUTURA OBSERVACIÓN
- 5 BOOTSTRAP
- 6 PRÓXIMA CLASE



- Veremos los principales usos de los intervalos de confianza en el ámbito de modelos lineales.
- Veremos cómo construir intervalos de confianza para:
  - cada coeficiente del modelo.
  - combinaciones lineales de parámetros.
- Tendremos una *breve* introducción al *Bootstrap* paramétrico aplicado a los modelos lineales.

# INTRODUCCIÓN

Los valores de las estimaciones  $\hat{\beta}$  están sujetos a la incertidumbre propia del muestreo. Por este motivo, por más que nuestra muestra sea tan grande como sea posible, **nunca** conoceremos el verdadero valor de los coeficientes del modelo.

Los valores de las estimaciones  $\hat{\beta}$  están sujetos a la incertidumbre propia del muestreo. Por este motivo, por más que nuestra muestra sea tan grande como sea posible, **nunca** conoceremos el verdadero valor de los coeficientes del modelo.

Sin embargo, podemos construir intervalos de confianza que nos permitan obtener una visión más amplia que la que nos proporcionana las estimaciones puntuales.

Los valores de las estimaciones  $\hat{\beta}$  están sujetos a la incertidumbre propia del muestreo. Por este motivo, por más que nuestra muestra sea tan grande como sea posible, **nunca** conoceremos el verdadero valor de los coeficientes del modelo.

Sin embargo, podemos construir intervalos de confianza que nos permitan obtener una visión más amplia que la que nos proporciona las estimaciones puntuales.

Debido a que estamos en un marco de trabajo paramétrico, los procedimientos que presentamos a continuación requieren de la hipótesis de normalidad de los errores del modelo.

$$\begin{aligned}\varepsilon &\sim N_n(\mathbf{0}, \sigma^2 I_n) \\ Y &\sim N_n(X\beta, \sigma^2 I_n)\end{aligned}$$

Los valores de las estimaciones  $\hat{\beta}$  están sujetos a la incertidumbre propia del muestreo. Por este motivo, por más que nuestra muestra sea tan grande como sea posible, **nunca** conoceremos el verdadero valor de los coeficientes del modelo.

Sin embargo, podemos construir intervalos de confianza que nos permitan obtener una visión más amplia que la que nos proporciona las estimaciones puntuales.

Debido a que estamos en un marco de trabajo paramétrico, los procedimientos que presentamos a continuación requieren de la hipótesis de normalidad de los errores del modelo.

$$\begin{aligned}\varepsilon &\sim N_n(\mathbf{0}, \sigma^2 I_n) \\ Y &\sim N_n(X\beta, \sigma^2 I_n)\end{aligned}$$

Al final de esta clase, tendremos una primera aproximación a la inferencia no paramétrica para modelos lineales empleando *Bootstrap*.

En la clase de pruebas de hipótesis se planteó una equivalencia entre las distribuciones  $t$  y  $F$  al llevar a cabo las pruebas de hipótesis individuales. Esta equivalencia solo es válida cuando en el numerador del estadístico  $F$  solo hay 1 grado de libertad.



En la clase de pruebas de hipótesis se planteó una equivalencia entre las distribuciones  $t$  y  $F$  al llevar a cabo las pruebas de hipótesis individuales. Esta equivalencia solo es válida cuando en el numerador del estadístico  $F$  solo hay 1 grado de libertad.

De esta manera, al emplear una matriz  $C$  distinta a  $(0, 0, \dots, 0, 1, 0, \dots, 0)$  la equivalencia seguiría siendo válida en tanto que la matriz  $C$  siga teniendo una sola fila.

En la clase de pruebas de hipótesis se planteó una equivalencia entre las distribuciones  $t$  y  $F$  al llevar a cabo las pruebas de hipótesis individuales. Esta equivalencia solo es válida cuando en el numerador del estadístico  $F$  solo hay 1 grado de libertad.

De esta manera, al emplear una matriz  $C$  distinta a  $(0, 0, \dots, 0, 1, 0, \dots, 0)$  la equivalencia seguiría siendo válida en tanto que la matriz  $C$  siga teniendo una sola fila.

Esto nos interesa porque plantearemos pruebas (e intervalos) basados en la siguiente cantidad:

$$\lambda' \beta$$

Siendo  $\lambda$  la matriz fila  $C$ .

En la clase de pruebas de hipótesis se planteó una equivalencia entre las distribuciones  $t$  y  $F$  al llevar a cabo las pruebas de hipótesis individuales. Esta equivalencia solo es válida cuando en el numerador del estadístico  $F$  solo hay 1 grado de libertad.

De esta manera, al emplear una matriz  $C$  distinta a  $(0, 0, \dots, 0, 1, 0, \dots, 0)$  la equivalencia seguiría siendo válida en tanto que la matriz  $C$  siga teniendo una sola fila.

Esto nos interesa porque plantearemos pruebas (e intervalos) basados en la siguiente cantidad:

$$\lambda' \beta$$

Siendo  $\lambda$  la matriz fila  $C$ .

Entonces, se necesita conocer la distribución del estadístico  $\lambda' \hat{\beta}$ .

Empleando los mismos desarrollos que la clase anterior, es posible observar que:

$$\lambda' \hat{\beta} \sim N(\lambda' \beta, \sigma^2 \lambda' (X' X)^{-1} \lambda)$$

Empleando los mismos desarrollos que la clase anterior, es posible observar que:

$$\lambda' \hat{\beta} \sim N(\lambda' \beta, \sigma^2 \lambda' (X' X)^{-1} \lambda)$$

Una vez más, el problema con este estadístico es que contiene un parámetro desconocido. Por este motivo, se realiza la misma estandarización realizada en la clase de pruebas de hipótesis.

$$t = \frac{\lambda' \hat{\beta} - \lambda \beta}{\sqrt{\hat{\sigma}^2 \lambda' (X' X)^{-1} \lambda}} \sim t_{n-k-1}$$

A partir del estadístico de la diapositiva anterior es posible plantear

$$\mathbb{P} \left( t_{n-k-1, \frac{\alpha}{2}} \leq \frac{\lambda' \hat{\beta} - \lambda' \beta}{s_{\lambda' \hat{\beta}}} \leq t_{n-k-1, 1-\frac{\alpha}{2}} \right) = 1 - \alpha$$

A partir del estadístico de la diapositiva anterior es posible plantear

$$\mathbb{P} \left( t_{n-k-1, \frac{\alpha}{2}} \leq \frac{\lambda' \hat{\beta} - \lambda' \beta}{s_{\lambda' \hat{\beta}}} \leq t_{n-k-1, 1-\frac{\alpha}{2}} \right) = 1 - \alpha$$

$$\mathbb{P} \left( \lambda' \hat{\beta} - t_{n-k-1, \frac{1-\alpha}{2}} s_{\lambda' \hat{\beta}} \leq \lambda' \beta \leq \lambda' \hat{\beta} + t_{n-k-1, \frac{1-\alpha}{2}} s_{\lambda' \hat{\beta}} \right) = 1 - \alpha$$

Siendo  $s_{\lambda' \hat{\beta}} = \sqrt{\hat{\sigma}^2 \lambda' (X' X)^{-1} \lambda}$ .

A partir del estadístico de la diapositiva anterior es posible plantear

$$\mathbb{P} \left( t_{n-k-1, \frac{\alpha}{2}} \leq \frac{\lambda' \hat{\beta} - \lambda' \beta}{s_{\lambda' \hat{\beta}}} \leq t_{n-k-1, 1-\frac{\alpha}{2}} \right) = 1 - \alpha$$

$$\mathbb{P} \left( \lambda' \hat{\beta} - t_{n-k-1, \frac{1-\alpha}{2}} s_{\lambda' \hat{\beta}} \leq \lambda' \beta \leq \lambda' \hat{\beta} + t_{n-k-1, \frac{1-\alpha}{2}} s_{\lambda' \hat{\beta}} \right) = 1 - \alpha$$

Siendo  $s_{\lambda' \hat{\beta}} = \sqrt{\hat{\sigma}^2 \lambda' (X' X)^{-1} \lambda}$ .

Esto nos indica que para plantear el IC se necesitan 3 ingredientes:

- El valor de  $\lambda' \hat{\beta}$ ,
- El valor de tabla  $t_{n-k-1, \frac{\alpha}{2}}$  y
- El desvío de la estimación de la combinación lineal  $s_{\lambda' \hat{\beta}}$ .



## INTERVALO PARA $\beta_j$

A partir del estadístico planteado en la diapositiva anterior, es posible plantear el IC para cada coeficiente por separado. Para esto, basta con tomar al vector  $\lambda$  como un vector de ceros con un único 1 en la posición  $j + 1$  (recordar que el primer elemento es  $\beta_0$ ).

## INTERVALO PARA $\beta_j$

A partir del estadístico planteado en la diapositiva anterior, es posible plantear el IC para cada coeficiente por separado. Para esto, basta con tomar al vector  $\lambda$  como un vector de ceros con un único 1 en la posición  $j + 1$  (recordar que el primer elemento es  $\beta_0$ ).

El intervalo de confianza al  $(1 - \alpha)\%$  queda definido por:

$$\left( \hat{\beta}_j - t_{n-k-1, \frac{1-\alpha}{2}} s_{\hat{\beta}_j} \quad ; \quad \hat{\beta}_j + t_{n-k-1, \frac{1-\alpha}{2}} s_{\hat{\beta}_j} \right)$$

Siendo  $s_{\hat{\beta}_j} = \hat{\sigma} g_{jj}$ .

## INTERVALO PARA $\beta_j$

A partir del estadístico planteado en la diapositiva anterior, es posible plantear el IC para cada coeficiente por separado. Para esto, basta con tomar al vector  $\lambda$  como un vector de ceros con un único 1 en la posición  $j + 1$  (recordar que el primer elemento es  $\beta_0$ ).

El intervalo de confianza al  $(1 - \alpha)\%$  queda definido por:

$$\left( \hat{\beta}_j - t_{n-k-1, \frac{1-\alpha}{2}} s_{\hat{\beta}_j} \quad ; \quad \hat{\beta}_j + t_{n-k-1, \frac{1-\alpha}{2}} s_{\hat{\beta}_j} \right)$$

Siendo  $s_{\hat{\beta}_j} = \hat{\sigma} g_{jj}$ .


Veamos algún ejemplo.



En la clase pasada utilizamos los datos de los automóviles de 1974 para explicar su rendimiento a partir de ciertas características. Al final, obtuvimos un modelo donde las variables explicativas eran el peso del auto y el tiempo que tarda en recorrer un cuarto de milla.

En la clase pasada utilizamos los datos de los automóviles de 1974 para explicar su rendimiento a partir de ciertas características. Al final, obtuvimos un modelo donde las variables explicativas eran el peso del auto y el tiempo que tarda en recorrer un cuarto de milla.



Volvamos a los datos de los autos en  para construir los  $IC_{95\%}$  de los coeficientes de este modelo.

## IC PARA LA RESPUESTA MEDIA

Hace unas cuantas clases vimos cómo utilizar el modelo lineal para predecir el valor medio de una nueva observación.

Hace unas cuantas clases vimos cómo utilizar el modelo lineal para predecir el valor medio de una nueva observación.

Disponiendo de un cierto modelo lineal

$$y_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_k x_{ik} + \varepsilon_i$$

podemos *predecir* el valor de la variable de respuesta correspondiente a una nueva observación  $y_0$  cuyo vector de variables explicativas  $x_0$  contiene los valores que nos interesan para cada una de las variables predictoras del modelo.

Hace unas cuantas clases vimos cómo utilizar el modelo lineal para predecir el valor medio de una nueva observación.

Disponiendo de un cierto modelo lineal

$$y_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_k x_{ik} + \varepsilon_i$$

podemos *predecir* el valor de la variable de respuesta correspondiente a una nueva observación  $y_0$  cuyo vector de variables explicativas  $x_0$  contiene los valores que nos interesan para cada una de las variables predictoras del modelo.

La forma de obtener la predicción es:

$$\hat{y}_0 = x_0' \hat{\beta} = \hat{\mathbb{E}}(y_0 | x_0)$$

De esta manera, se está estimando la *respuesta media* (RM) que tendría un grupo de individuos con características  $x_0$ .



Retomando el predictor de la respuesta media:

$$\hat{y}_0 = x_0' \hat{\beta} = \hat{\mathbb{E}}(y_0 | x_0)$$

Retomando el predictor de la respuesta media:

$$\hat{y}_0 = x_0' \hat{\beta} = \hat{\mathbb{E}}(y_0 | x_0)$$

Obsérvese cómo, lo que se está haciendo es una combinación lineal de los coeficientes estimados con valores dados por  $x_0$ . Por lo cual, construir un intervalo de confianza es solo un caso especial del estadístico que hemos estado viendo. Lo único que debemos tener en cuenta es que  $\lambda = x_0$ . Entonces:

Retomando el predictor de la respuesta media:

$$\hat{y}_0 = x_0' \hat{\beta} = \hat{\mathbb{E}}(y_0 | x_0)$$

Obsérvese cómo, lo que se está haciendo es una combinación lineal de los coeficientes estimados con valores dados por  $x_0$ . Por lo cual, construir un intervalo de confianza es solo un caso especial del estadístico que hemos estado viendo. Lo único que debemos tener en cuenta es que  $\lambda = x_0$ . Entonces:

$$\left( x_0' \hat{\beta} - t_{n-k-1, \frac{1-\alpha}{2}} s_{RM} \quad ; \quad x_0' \hat{\beta} + t_{n-k-1, \frac{1-\alpha}{2}} s_{RM} \right)$$

Donde  $s_{RM} = \hat{\sigma} \sqrt{x_0' (X'X)^{-1} x_0}$

## IC PARA UNA FUTURA OBSERVACIÓN

Los intervalos para una observación (futura o no) son llamados *intervalos de predicción* y se diferencian de los anteriores en que estos últimos incluyen la variabilidad propia de la observación a ser predicha.

Los intervalos para una observación (futura o no) son llamados *intervalos de predicción* y se diferencian de los anteriores en que estos últimos incluyen la variabilidad propia de la observación a ser predicha.

Veremos que los intervalos para una observación futura suelen ser más anchos que sus análogos para la respuesta media.

Los intervalos para una observación (futura o no) son llamados *intervalos de predicción* y se diferencian de los anteriores en que estos últimos incluyen la variabilidad propia de la observación a ser predicha.

Veremos que los intervalos para una observación futura suelen ser más anchos que sus análogos para la respuesta media.

Al predecir el valor de  $y_0$  mediante  $\hat{y}_0$  queremos cuantificar la variabilidad de  $y_0 - \hat{y}_0$ , esto es:

$$\mathbb{V}ar(y_0 - \hat{y}_0) = \mathbb{V}ar(y_0 - x_0' \hat{\beta}) = \mathbb{V}ar(x_0' \beta + \varepsilon_0 - x_0' \hat{\beta})$$

Retomando la diapositiva anterior:

$$\begin{aligned}\mathbb{V}ar(y_0 - \hat{y}_0) &= \mathbb{V}ar(\varepsilon_0) + \mathbb{V}ar(x_0' \hat{\beta}) \\ &= \sigma^2 + \sigma^2 x_0' (X'X)^{-1} x_0 \\ &= \sigma^2 \left[ 1 + x_0' (X'X)^{-1} x_0 \right]\end{aligned}$$

De esta manera, el intervalo de predicción de una futura observación (FO) es:

Retomando la diapositiva anterior:

$$\begin{aligned}\mathbb{V}ar(y_0 - \hat{y}_0) &= \mathbb{V}ar(\varepsilon_0) + \mathbb{V}ar(x_0' \hat{\beta}) \\ &= \sigma^2 + \sigma^2 x_0' (X' X)^{-1} x_0 \\ &= \sigma^2 \left[ 1 + x_0' (X' X)^{-1} x_0 \right]\end{aligned}$$

De esta manera, el intervalo de predicción de una futura observación (FO) es:

$$\left( x_0' \hat{\beta} - t_{n-k-1, \frac{1-\alpha}{2}} s_{FO} \quad ; \quad x_0' \hat{\beta} + t_{n-k-1, \frac{1-\alpha}{2}} s_{FO} \right)$$

Donde  $s_{FO} = \hat{\sigma} \sqrt{1 + x_0' (X' X)^{-1} x_0}$ .





Retomando el ejemplo de los autos, supongamos que deseamos predecir el rendimiento de un automóvil que pesa 2000kg y al cual le toma 20 segundos recorrer un cuarto de milla.

Retomando el ejemplo de los autos, supongamos que deseamos predecir el rendimiento de un automóvil que pesa 2000kg y al cual le toma 20 segundos recorrer un cuarto de milla.



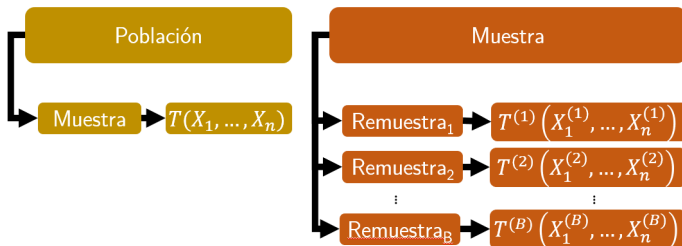
Volvamos a los datos de los autos en  para realizar la predicción y construir los intervalos.

# ¿QUÉ ES BOOTSTRAP?

Bootstrapping es un enfoque general de inferencia estadística basado en la construcción de la distribución en el muestreo de un estadístico **re-muestreando** los datos disponibles. La idea consiste en usar la muestra como si fuese la población y extraer de ella muestras con reposición a partir de las cuales volver a obtener valores del estadístico en cuestión. A estos nuevos valores del estadístico, se le suelen llamar *réplicas*.

# ¿QUÉ ES BOOTSTRAP?

Bootstrapping es un enfoque general de inferencia estadística basado en la construcción de la distribución en el muestreo de un estadístico **re-muestreando** los datos disponibles. La idea consiste en usar la muestra como si fuese la población y extraer de ella muestras con reposición a partir de las cuales volver a obtener valores del estadístico en cuestión. A estos nuevos valores del estadístico, se le suelen llamar *réplicas*.



La población es a la muestra lo que la muestra es a las re-muestras.

La ventaja de este método recae en que nos permite obtener la distribución en el muestreo del estadístico  $T(X)$  de manera empírica sin realizar supuestos sobre la población y sin llevar a cabo procedimientos matemáticos.

La ventaja de este método recae en que nos permite obtener la distribución en el muestreo del estadístico  $T(X)$  de manera empírica sin realizar supuestos sobre la población y sin llevar a cabo procedimientos matemáticos.

El detalle clave a tener en cuenta es que las remuestras son obtenidas **con reposición** y del **mismo tamaño** que la muestra original.

La ventaja de este método recae en que nos permite obtener la distribución en el muestreo del estadístico  $T(X)$  de manera empírica sin realizar supuestos sobre la población y sin llevar a cabo procedimientos matemáticos.

El detalle clave a tener en cuenta es que las remuestras son obtenidas **con reposición** y del **mismo tamaño** que la muestra original.

Este método permite obtener resultados sobre la variabilidad del estadístico  $T(X)$ .

- Sesgo.
- Varianza.
- Intervalos de confianza para el parámetro que estima  $T(X)$ .

Existen diferentes alternativas, para construir IC a partir de remuestras bootstrap, las más comunes son.

- *IC basado en normalidad (asintótica)*

En este caso la idea es similar al método que venimos empleando para construir los IC pero estimando la variabilidad del estadístico a partir de  $B$  las réplicas.

$$IC_{1-\alpha} = T(X) \mp z_{1-\frac{\alpha}{2}} \sqrt{\frac{\sum_k (T^k(X^k) - \bar{T}(X))^2}{B-1}}$$

- *IC basado en percentiles*

Esta otra alternativa se basa en la obtención de los percentiles  $\frac{\alpha}{2}$  y  $1 - \frac{\alpha}{2}$  a partir de las  $B$  réplicas. Tiene la ventaja de que, a diferencia de la anterior, funciona mejor cuando la distribución de las réplicas no es simétrica.

$$IC_{1-\alpha} = \left( T^{[(B+1)\alpha/2]}(X) ; T^{[(B+1)(1-\alpha/2)]}(X) \right)$$



# TIPOS DE REMUESTREO

Si bien cada problema admite distintas estrategias para obtener las remuestras, existen 2 formas que definen 2 maneras de emplear bootstrap.

Si bien cada problema admite distintas estrategias para obtener las remuestras, existen 2 formas que definen 2 maneras de emplear bootstrap.

## Bootstrap paramétrico

Este método se basa en obtener las remuestras a partir de un modelo previamente ajustado. En el caso de utilizarlo en el ámbito de modelos lineales, la idea puede esquematizarse en los siguientes pasos:

- ➊ Asumiendo que el modelo es correcto, se obtienen los valores ajustados  $\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n$ .

Si bien cada problema admite distintas estrategias para obtener las remuestras, existen 2 formas que definen 2 maneras de emplear bootstrap.

## Bootstrap paramétrico

Este método se basa en obtener las remuestras a partir de un modelo previamente ajustado. En el caso de utilizarlo en el ámbito de modelos lineales, la idea puede esquematizarse en los siguientes pasos:

- ➊ Asumiendo que el modelo es correcto, se obtienen los valores ajustados  $\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n$ .
- ➋ A partir de los anteriores, se obtienen los residuos  $\hat{\varepsilon}_1, \hat{\varepsilon}_2, \dots, \hat{\varepsilon}_n$

Si bien cada problema admite distintas estrategias para obtener las remuestras, existen 2 formas que definen 2 maneras de emplear bootstrap.

## Bootstrap paramétrico

Este método se basa en obtener las remuestras a partir de un modelo previamente ajustado. En el caso de utilizarlo en el ámbito de modelos lineales, la idea puede esquematizarse en los siguientes pasos:

- ➊ Asumiendo que el modelo es correcto, se obtienen los valores ajustados  $\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n$ .
- ➋ A partir de los anteriores, se obtienen los residuos  $\hat{\varepsilon}_1, \hat{\varepsilon}_2, \dots, \hat{\varepsilon}_n$
- ➌ Se obtiene una muestra con reposición a partir de los residuos:  $\hat{\varepsilon}_1^*, \hat{\varepsilon}_2^*, \dots, \hat{\varepsilon}_n^*$ .

Si bien cada problema admite distintas estrategias para obtener las remuestras, existen 2 formas que definen 2 maneras de emplear bootstrap.

## Bootstrap paramétrico

Este método se basa en obtener las remuestras a partir de un modelo previamente ajustado. En el caso de utilizarlo en el ámbito de modelos lineales, la idea puede esquematizarse en los siguientes pasos:

- ➊ Asumiendo que el modelo es correcto, se obtienen los valores ajustados  $\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n$ .
- ➋ A partir de los anteriores, se obtienen los residuos  $\hat{\varepsilon}_1, \hat{\varepsilon}_2, \dots, \hat{\varepsilon}_n$ .
- ➌ Se obtiene una muestra con reposición a partir de los residuos:  $\hat{\varepsilon}_1^*, \hat{\varepsilon}_2^*, \dots, \hat{\varepsilon}_n^*$ .
- ➍ Se reconstruye la variable de respuesta  $Y$  mediante  $y_i^* = \hat{y}_i + \hat{\varepsilon}_i^*$ .

Si bien cada problema admite distintas estrategias para obtener las remuestras, existen 2 formas que definen 2 maneras de emplear bootstrap.

## Bootstrap paramétrico

Este método se basa en obtener las remuestras a partir de un modelo previamente ajustado. En el caso de utilizarlo en el ámbito de modelos lineales, la idea puede esquematizarse en los siguientes pasos:

- ➊ Asumiendo que el modelo es correcto, se obtienen los valores ajustados  $\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n$ .
- ➋ A partir de los anteriores, se obtienen los residuos  $\hat{\varepsilon}_1, \hat{\varepsilon}_2, \dots, \hat{\varepsilon}_n$ .
- ➌ Se obtiene una muestra con reposición a partir de los residuos:  $\hat{\varepsilon}_1^*, \hat{\varepsilon}_2^*, \dots, \hat{\varepsilon}_n^*$ .
- ➍ Se reconstruye la variable de respuesta  $Y$  mediante  $y_i^* = \hat{y}_i + \hat{\varepsilon}_i^*$ .
- ➎ De esta forma, manteniendo inalterada la matriz  $X$ , pero con la remuestra  $Y^*$  se vuelve a ajustar el modelo. Obteniéndose  $\hat{\beta}_{MV}^*$ .

## Bootstrap no paramétrico

A diferencia del anterior, este método no requiere asumir nada respecto del modelo. El procedimiento a repetir  $B$  veces es:

- 1 Se obtiene la remuestra muestreando con reposición el conjunto de vectores  $(y_i, x_{i1}, x_{i2}, \dots, x_{ik})$ .

## Bootstrap no paramétrico

A diferencia del anterior, este método no requiere asumir nada respecto del modelo. El procedimiento a repetir  $B$  veces es:

- 1 Se obtiene la remuestra muestreando con reposición el conjunto de vectores  $(y_i, x_{i1}, x_{i2}, \dots, x_{ik})$ .
- 2 Habiendo sorteado  $n$  de estos, se obtiene una nueva matriz  $X^*$  y un nuevo  $Y^*$ .



## Bootstrap no paramétrico

A diferencia del anterior, este método no requiere asumir nada respecto del modelo. El procedimiento a repetir  $B$  veces es:

- 1 Se obtiene la remuestra muestreando con reposición el conjunto de vectores  $(y_i, x_{i1}, x_{i2}, \dots, x_{ik})$ .
- 2 Habiendo sorteado  $n$  de estos, se obtiene una nueva matriz  $X^*$  y un nuevo  $Y^*$ .
- 3 Se vuelve a ajustar el modelo. Obteniéndose  $\hat{\beta}_{MV}^*$

## Bootstrap no paramétrico

A diferencia del anterior, este método no requiere asumir nada respecto del modelo. El procedimiento a repetir  $B$  veces es:

- 1 Se obtiene la remuestra muestreando con reposición el conjunto de vectores  $(y_i, x_{i1}, x_{i2}, \dots, x_{ik})$ .
- 2 Habiendo sorteado  $n$  de estos, se obtiene una nueva matriz  $X^*$  y un nuevo  $Y^*$ .
- 3 Se vuelve a ajustar el modelo. Obteniéndose  $\hat{\beta}_{MV}^*$


En cualquiera de los 2 casos (paramétrico o no paramétrico), una vez obtenidas las réplicas del estadístico, se utilizan los métodos explicados anteriormente para construir los intervalos.



Recalculemos los intervalos de confianza y predicción para el caso de  $x_0 = (1, 2, 20)$  empleando bootstrap paramétrico.

Recalculemos los intervalos de confianza y predicción para el caso de  $x_0 = (1, 2, 20)$  empleando bootstrap paramétrico.



Volvamos a los datos de los autos en  para construir los intervalos usando *bootstrap*.



La próxima hablaremos de:

- Métodos de selección de modelos.
- Haremos una clase/taller sobre los métodos inferenciales que hemos visto hasta hoy.



Carmona, Francesc (2003). *Modelos Lineales (notas de curso)*. Departament d'Estadística.



Faraway, Julian (2014). *Linear Models with R, second edition*. Chapman Hall/CRC.



Rencher, Alvin y Bruce Schaalje (2008). *Linear Models in Statistics, second edition*. John Wiley Sons, Inc.

¿Preguntas?

Muchas Gracias