IN4402 - Aplicaciones de Probabilidades y Estadística en Gestión

Ronald Leblebici - Angelo Muñoz

ronald.leblebici@gmail.com

gelox97@gmail.com

Facultad de Ciencias Físicas y Matemáticas
Universidad de Chile

16 de octubre del 2018

¿Qué estudiaremos hoy?

- Uso de variables dummies sin interacciones
  - Para estimar promedios y diferencias promedio.
  - Para estimar interceptos distintos.
- Interacciones entre variables dummies para crear sub-categorías
- Interacciones entre variables dummies y variables reales para obtener efectos marginales diferidos entre categorías.
- Interacciones entre variables reales para obtener efectos marginales cambiantes para un regresor a medida que otro regresor varía.

Una amiga médica de usted le pide ayuda para estudiar los efectos del consumo de tabaco –representado por las variables *dummies* de fumar $F_n \in \{0,1\}$ y no fumar $NF_n \in \{0,1\}$ respectivamente–

Una amiga médica de usted le pide ayuda para estudiar los efectos del consumo de tabaco –representado por las variables *dummies* de fumar $F_n \in \{0,1\}$ y no fumar $NF_n \in \{0,1\}$ respectivamente–y las horas de ejercicio $EJ_n \in \mathbb{R}^+$ sobre la capacidad aeróbica $CA_n \in \mathbb{R}^+$ en humanos.

Introducción

Una amiga médica de usted le pide ayuda para estudiar los efectos del consumo de tabaco –representado por las variables *dummies* de fumar $F_n \in \{0,1\}$ y no fumar $NF_n \in \{0,1\}$ respectivamente–y las horas de ejercicio $EJ_n \in \mathbb{R}^+$ sobre la capacidad aeróbica $CA_n \in \mathbb{R}^+$ en humanos.También se tiene información sobre la edad $ED_n \in \mathbb{R}^+$ de las personas estudiadas.

Introducción

Una amiga médica de usted le pide ayuda para estudiar los efectos del consumo de tabaco –representado por las variables *dummies* de fumar $F_n \in \{0{,}1\}$ y no fumar $NF_n \in \{0{,}1\}$ respectivamente–y las horas de ejercicio $EJ_n \in \mathbb{R}^+$ sobre la capacidad aeróbica $CA_n \in \mathbb{R}^+$ en humanos.También se tiene información sobre la edad $ED_n \in \mathbb{R}^+$ de las personas estudiadas. Su amiga se ha tomado el tiempo de crear una variable categórica de rango etario $RE_n \in \{1, 2, 3\}$ que toma los valores 1, 2 y 3 dependiendo de si el individuo n es adolescente ($ED_n \in [14 - 18)$), adulto ($ED_n \in [18 - 60)$) o adulto mayor ($ED_n \in [60, \infty)$) respectivamente.

Una amiga médica de usted le pide ayuda para estudiar los efectos del consumo de tabaco –representado por las variables *dummies* de fumar $F_n \in \{0,1\}$ y no fumar $NF_n \in \{0,1\}$ respectivamente–y las horas de ejercicio $EJ_n \in \mathbb{R}^+$ sobre la capacidad aeróbica $CA_n \in \mathbb{R}^+$ en humanos.También se tiene información sobre la edad $ED_n \in \mathbb{R}^+$ de las personas estudiadas. Su amiga se ha tomado el tiempo de crear una variable categórica de rango etario $RE_n \in \{1, 2, 3\}$ que toma los valores 1, 2 y 3 dependiendo de si el individuo n es adolescente ($ED_n \in [14 - 18)$), adulto ($ED_n \in [18 - 60)$) o adulto mayor ($ED_n \in [60, \infty)$) respectivamente. Como ella sabe que usted tiene un excelente manejo de estadísticas le dice lo siguiente:

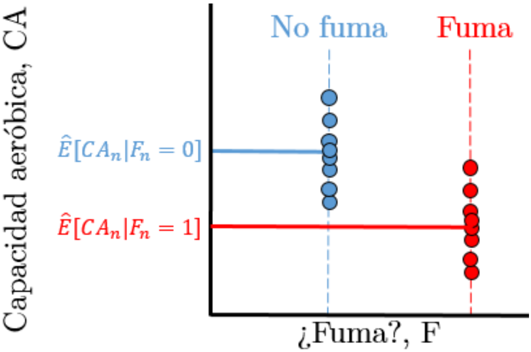
P1. Variables dummies sin interacciones

'En promedio quienes no fuman tienen mejor condición aeróbica que quienes sí". Luego le pide:

a. Hacer un bosquejo de cómo se distribuye la muestra en cuanto a condición aeróbica, separando a fumadores y no fumadores. Mostrar dónde está el promedio muestral.

'En promedio quienes no fuman tienen mejor condición aeróbica que quienes sí". Luego le pide:

a. Hacer un bosquejo de cómo se distribuye la muestra en cuanto a condición aeróbica, separando a fumadores y no fumadores. Mostrar dónde está el promedio muestral.



'En promedio quienes no fuman tienen mejor condición aeróbica que quienes sí". Luego le pide:

b. Plantear dos ejemplos modelos de regresión lineal **sin colinealidad perfecta** y un contraejemplo **con colinealidad perfecta** diseñados para estimar estos promedios. Explique por qué (no) hay colinealidad perfecta en cada caso y por qué este es un problema en el contraejemplo.

'En promedio quienes no fuman tienen mejor condición aeróbica que quienes sí". Luego le pide:

b. Plantear dos ejemplos modelos de regresión lineal **sin colinealidad perfecta** y un contraejemplo **con colinealidad perfecta** diseñados para estimar estos promedios. Explique por qué (no) hay colinealidad perfecta en cada caso y por qué este es un problema en el contraejemplo.

Ejemplo 1: $CA_n = \beta_0 + \beta_F F_n + U_n$.

P1. Variables dummies sin interacciones

'En promedio quienes no fuman tienen mejor condición aeróbica que quienes sí". Luego le pide:

b. Plantear dos ejemplos modelos de regresión lineal **sin colinealidad perfecta** y un contraejemplo **con colinealidad perfecta** diseñados para estimar estos promedios. Explique por qué (no) hay colinealidad perfecta en cada caso y por qué este es un problema en el contraejemplo.

Ejemplo 1: $CA_n = \beta_0 + \beta_F F_n + U_n$.

No hay colinealidad perfecta pues los regresores $\vec{1}$ y F no son perfectamente colineales, ie $\nexists \lambda \in \mathbb{R} | F = \lambda \vec{1}$

'En promedio quienes no fuman tienen mejor condición aeróbica que quienes sí". Luego le pide:

b. Plantear dos ejemplos modelos de regresión lineal **sin colinealidad perfecta** y un contraejemplo **con colinealidad perfecta** diseñados para estimar estos promedios. Explique por qué (no) hay colinealidad perfecta en cada caso y por qué este es un problema en el contraejemplo.

'En promedio quienes no fuman tienen mejor condición aeróbica que quienes sí". Luego le pide:

b. Plantear dos ejemplos modelos de regresión lineal **sin colinealidad perfecta** y un contraejemplo **con colinealidad perfecta** diseñados para estimar estos promedios. Explique por qué (no) hay colinealidad perfecta en cada caso y por qué este es un problema en el contraejemplo.

Ejemplo 2: $CA_n = \beta_{NF} NF_n + \beta_F F_n + U_n$.

'En promedio quienes no fuman tienen mejor condición aeróbica que quienes sí". Luego le pide:

b. Plantear dos ejemplos modelos de regresión lineal **sin colinealidad perfecta** y un contraejemplo **con colinealidad perfecta** diseñados para estimar estos promedios. Explique por qué (no) hay colinealidad perfecta en cada caso y por qué este es un problema en el contraejemplo.

Ejemplo 2: $CA_n = \beta_{NF} NF_n + \beta_F F_n + U_n$.

No hay colinealidad perfecta pues los regresores F y NF no son perfectamente colineales, ie $\nexists \lambda \in \mathbb{R} | F = \lambda NF$.

'En promedio quienes no fuman tienen mejor condición aeróbica que quienes sí". Luego le pide:

b. Plantear dos ejemplos modelos de regresión lineal **sin colinealidad perfecta** y un contraejemplo **con colinealidad perfecta** diseñados para estimar estos promedios. Explique por qué (no) hay colinealidad perfecta en cada caso y por qué este es un problema en el contraejemplo.

P1. Variables dummies sin interacciones

'En promedio quienes no fuman tienen mejor condición aeróbica que quienes sí". Luego le pide:

b. Plantear dos ejemplos modelos de regresión lineal **sin colinealidad perfecta** y un contraejemplo **con colinealidad perfecta** diseñados para estimar estos promedios. Explique por qué (no) hay colinealidad perfecta en cada caso y por qué este es un problema en el contraejemplo.

Contraejemplo: $CA_n = \beta_0 + \beta_{NF} NF_n + \beta_F F_n + U_n$.

'En promedio quienes no fuman tienen mejor condición aeróbica que quienes sí". Luego le pide:

b. Plantear dos ejemplos modelos de regresión lineal **sin colinealidad perfecta** y un contraejemplo **con colinealidad perfecta** diseñados para estimar estos promedios. Explique por qué (no) hay colinealidad perfecta en cada caso y por qué este es un problema en el contraejemplo.

Contraejemplo: $CA_n = \beta_0 + \beta_{NF} NF_n + \beta_F F_n + U_n$.

Hay colinealidad perfecta pues el regresor $\vec{1}$ es colineal con F y NF. Efecto, $\vec{1} = F + NF$.

'En promedio quienes no fuman tienen mejor condición aeróbica que quienes sí". Luego le pide:

b. Plantear dos ejemplos modelos de regresión lineal **sin colinealidad perfecta** y un contraejemplo **con colinealidad perfecta** diseñados para estimar estos promedios. Explique por qué (no) hay colinealidad perfecta en cada caso y por qué este es un problema en el contraejemplo.

Contraejemplo: $CA_n = \beta_0 + \beta_{NF} NF_n + \beta_F F_n + U_n$.

Hay colinealidad perfecta pues el regresor $\vec{1}$ es colineal con F y NF. Efecto, $\vec{1} = F + NF$.

Como consecuencia de esto, la matriz $X^T X$ no será invertible, por lo que será imposible calcular el estimador $\widehat{\beta} = (X^T X)^{-1} X^T$.

'En promedio quienes no fuman tienen mejor condición aeróbica que quienes sí". Luego le pide:

b. Plantear dos ejemplos modelos de regresión lineal **sin colinealidad perfecta** y un contraejemplo **con colinealidad perfecta** diseñados para estimar estos promedios. Explique por qué (no) hay colinealidad perfecta en cada caso y por qué este es un problema en el contraejemplo.

Contraejemplo: $CA_n = \beta_0 + \beta_{NF} NF_n + \beta_F F_n + U_n$.

Hay colinealidad perfecta pues el regresor $\vec{1}$ es colineal con F y NF. Efecto, $\vec{1} = F + NF$.

Como consecuencia de esto, la matriz $X^T X$ no será invertible, por lo que será imposible calcular el estimador $\widehat{\beta} = (X^T X)^{-1} X^T$.

Al haber colinealidad perfecta, es imposible identificar a qué parámetro se deben las variaciones en la variable dependiente (problema de identificación).

"En promedio quienes no fuman tienen mejor condición aeróbica que quienes sí". Luego le pide:

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   i. Calcule $\mathbb{E}[CA_n|F_n = 0]$, $\mathbb{E}[CA_n|F_n = 1]$.

"En promedio quienes no fuman tienen mejor condición aeróbica que quienes sí". Luego le pide:

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   i. Calcule $\mathbb{E}[CA_n|F_n = 0]$, $\mathbb{E}[CA_n|F_n = 1]$.

Contraejemplo: $CA_n = \beta_0 + \beta_F F_n + U_n$.

$$\mathbb{E}[CA_n|F_n = 1] = \beta_0 + \beta_F$$

"En promedio quienes no fuman tienen mejor condición aeróbica que quienes sí". Luego le pide:

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   i. Calcule $\mathbb{E}[CA_n|F_n = 0]$, $\mathbb{E}[CA_n|F_n = 1]$.

Contraejemplo: $CA_n = \beta_0 + \beta_F F_n + U_n$.

$$\mathbb{E}[CA_n|F_n = 1] = \beta_0 + \beta_F$$

$$\mathbb{E}[CA_n|F_n = 0] = \beta_0$$

'En promedio quienes no fuman tienen mejor condición aeróbica que quienes sí". Luego le pide:

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
    ii. Calcule la diferencia esperada en condición aeróbica entre un fumador y un no fumador.

Ejemplo: $CA_n = \beta_0 + \beta_F F_n + U_n$.

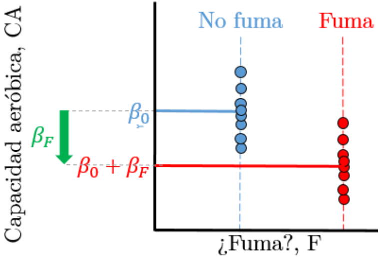
'En promedio quienes no fuman tienen mejor condición aeróbica que quienes sí". Luego le pide:

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   ii. Calcule la diferencia esperada en condición aeróbica entre un fumador y un no fumador.

Ejemplo: $CA_n = \beta_0 + \beta_F F_n + U_n$.

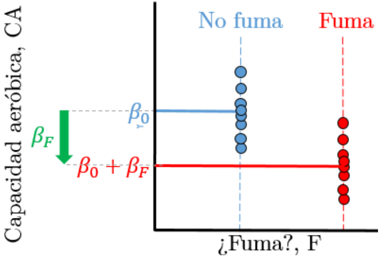
$$\mathbb{E}[CA_n|F_n = 1] - \mathbb{E}[CA_n|F_n = 0] = (\beta_0 + \beta_F) - \beta_0 = \beta_F$$

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   iii. En cuanto a los parámetros poblacionales de su modelo, especule sus signos (+/-), interprete su significado y muéstrelos gráficamente.



c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):

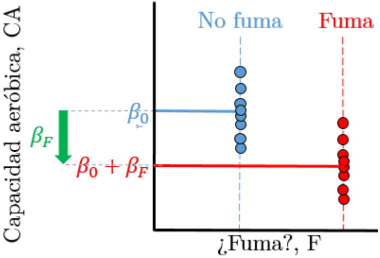
   iii. En cuanto a los parámetros poblacionales de su modelo, especule sus signos (+/-), interprete su significado y muéstrelos gráficamente.



$$\underbrace{\beta_0}_{\text{CA promedio no-fumadores}} > \underbrace{\beta_0 + \beta_F}_{\text{CA promedio fumadores}} > 0.$$

P1. Variables dummies sin interacciones

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   iii. En cuanto a los parámetros poblacionales de su modelo, especule sus signos (+/-), interprete su significado y muéstrelos gráficamente.



$$\underbrace{\beta_0}_{\text{CA promedio no-fumadores}} > \underbrace{\beta_0 + \beta_F}_{\text{CA promedio fumadores}} > 0.$$

$\beta_F < 0$ es la diferencia promedio entre fumadores y no fumadores.

"En promedio quienes no fuman tienen mejor condición aeróbica que quienes sí". Luego le pide:

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   i. Calcule $\mathbb{E}[CA_n|F_n = 0]$, $\mathbb{E}[CA_n|F_n = 1]$.

"En promedio quienes no fuman tienen mejor condición aeróbica que quienes sí". Luego le pide:

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   i. Calcule $\mathbb{E}[CA_n|F_n = 0]$, $\mathbb{E}[CA_n|F_n = 1]$.

Ejemplo 1: $CA_n = \beta_F F_n + \beta_{NF} NF_n + U_n$.

"En promedio quienes no fuman tienen mejor condición aeróbica que quienes sí". Luego le pide:

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   i. Calcule $\mathbb{E}[CA_n|F_n = 0]$, $\mathbb{E}[CA_n|F_n = 1]$.

Ejemplo 1: $CA_n = \beta_F F_n + \beta_{NF} NF_n + U_n$.

$$\mathbb{E}[CA_n|F_n = 1] = \beta_F$$

"En promedio quienes no fuman tienen mejor condición aeróbica que quienes sí". Luego le pide:

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   i. Calcule $\mathbb{E}[CA_n|F_n = 0]$, $\mathbb{E}[CA_n|F_n = 1]$.

Ejemplo 1: $CA_n = \beta_F F_n + \beta_{NF} NF_n + U_n$.

$$\mathbb{E}[CA_n|F_n = 1] = \beta_F$$

$$\mathbb{E}[CA_n|F_n = 0] = \beta_{NF}$$

'En promedio quienes no fuman tienen mejor condición aeróbica que quienes sí". Luego le pide:

- c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
  - ii. Calcule la diferencia esperada en condición aeróbica entre un fumador y un no fumador.

Ejemplo 1: $CA_n = \beta_F F_n + \beta_{NF} NF_n + U_n$.

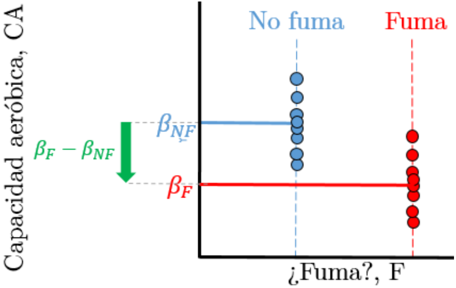
'En promedio quienes no fuman tienen mejor condición aeróbica que quienes sí". Luego le pide:

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   ii. Calcule la diferencia esperada en condición aeróbica entre un fumador y un no fumador.

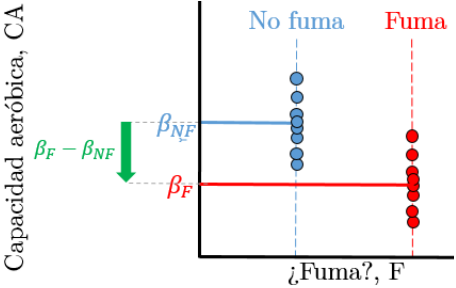
Ejemplo 1: $CA_n = \beta_F F_n + \beta_{NF} NF_n + U_n$.

$$\mathbb{E}[CA_n | F_n = 1] - \mathbb{E}[CA_n | F_n = 0] = \beta_F - \beta_{NF}$$

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   iii. En cuanto a los parámetros poblacionales de su modelo, especule sus signos (+/-), interprete su significado y muéstrelos gráficamente.

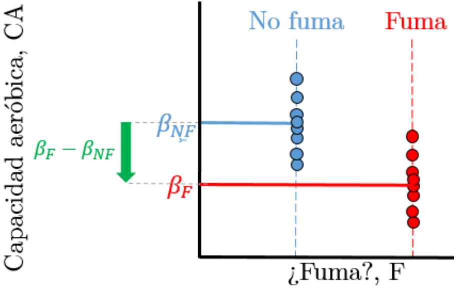


c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   iii. En cuanto a los parámetros poblacionales de su modelo, especule sus signos (+/-), interprete su significado y muéstrelos gráficamente.



$$\underbrace{\beta_{NF}}_{\text{CA promedio no-fumadores}} > \underbrace{\beta_F}_{\text{CA promedio fumadores}} > 0.$$

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   iii. En cuanto a los parámetros poblacionales de su modelo, especule sus signos (+/-), interprete su significado y muéstrelos gráficamente.



$$\underbrace{\beta_{NF}}_{\text{CA promedio no-fumadores}} > \underbrace{\beta_F}_{\text{CA promedio fumadores}} > 0.$$

$\beta_F - \beta_{NF} < 0$ es la dif. promedio entre fumadores y no fumadores.

"En promedio quienes no fuman tienen mejor condición aeróbica que quienes sí". Luego le pide:

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   i. Calcule $\mathbb{E}[CA_n|F_n = 0]$, $\mathbb{E}[CA_n|F_n = 1]$.

"En promedio quienes no fuman tienen mejor condición aeróbica que quienes sí". Luego le pide:

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   i. Calcule $\mathbb{E}[CA_n|F_n = 0]$, $\mathbb{E}[CA_n|F_n = 1]$.

Contraejemplo: $CA_n = \beta_0 + \beta_F F_n + \beta_{NF} NF_n + U_n$.

"En promedio quienes no fuman tienen mejor condición aeróbica que quienes sí". Luego le pide:

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   i. Calcule $\mathbb{E}[CA_n|F_n = 0]$, $\mathbb{E}[CA_n|F_n = 1]$.

Contraejemplo: $CA_n = \beta_0 + \beta_F F_n + \beta_{NF} NF_n + U_n$.

$$\mathbb{E}[CA_n|F_n = 1\beta_0 + \beta_F$$

P1. Variables dummies sin interacciones

"En promedio quienes no fuman tienen mejor condición aeróbica que quienes sí". Luego le pide:

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   i. Calcule $\mathbb{E}[CA_n|F_n = 0]$, $\mathbb{E}[CA_n|F_n = 1]$.

Contraejemplo: $CA_n = \beta_0 + \beta_F F_n + \beta_{NF} NF_n + U_n$.

$$\mathbb{E}[CA_n|F_n = 1\beta_0 + \beta_F$$

$$\mathbb{E}[CA_n|F_n = 0] = \beta_0 + \beta_{NF}$$

'En promedio quienes no fuman tienen mejor condición aeróbica que quienes sí". Luego le pide:

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
    ii. Calcule la diferencia esperada en condición aeróbica entre un fumador y un no fumador.

Contraejemplo: $CA_n = \beta_0 + \beta_F F_n + \beta_{NF} NF_n + U_n$.

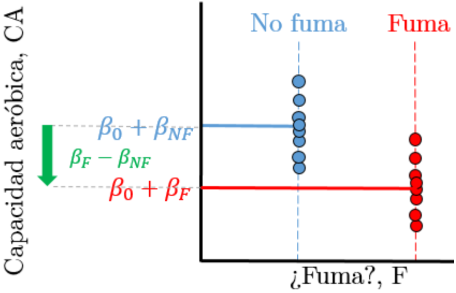
'En promedio quienes no fuman tienen mejor condición aeróbica que quienes sí". Luego le pide:

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
    ii. Calcule la diferencia esperada en condición aeróbica entre un fumador y un no fumador.

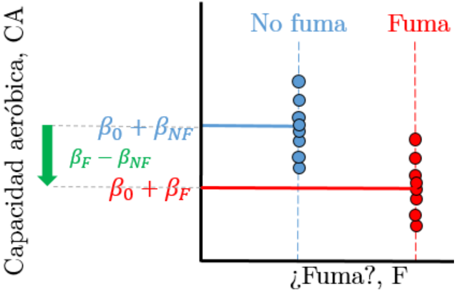
Contraejemplo: $CA_n = \beta_0 + \beta_F F_n + \beta_{NF} NF_n + U_n$.

$$\mathbb{E}[CA_n | F_n = 1] - \mathbb{E}[CA_n | F_n = 0] = (\beta_0 + \beta_F) - (\beta_0 + \beta_{NF}) = \beta_F - \beta_{NF}$$

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   iii. En cuanto a los parámetros poblacionales de su modelo, especule sus signos (+/-), interprete su significado y muéstrelos gráficamente.

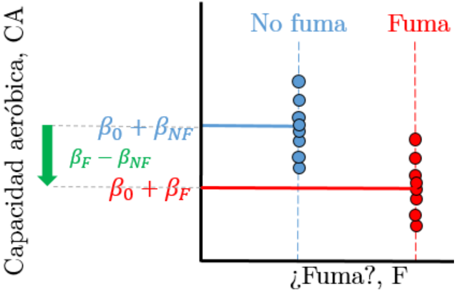


c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   iii. En cuanto a los parámetros poblacionales de su modelo, especule sus signos (+/-), interprete su significado y muéstrelos gráficamente.



$$\underbrace{\beta_0 + \beta_{NF}}_{\text{CA promedio no-fumadores}} > \underbrace{\beta_0 + \beta_F}_{\text{CA promedio fumadores}} > 0.$$

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   iii. En cuanto a los parámetros poblacionales de su modelo, especule sus signos (+/-), interprete su significado y muéstrelos gráficamente.



$$\underbrace{\beta_0 + \beta_{NF}}_{\text{CA promedio no-fumadores}} > \underbrace{\beta_0 + \beta_F}_{\text{CA promedio fumadores}} > 0.$$

$\beta_F - \beta_{NF} < 0$ es la dif. promedio entre fumadores y no fumadores.

Recordemos que en el contraejemplo hay colineadlidad perfecta.

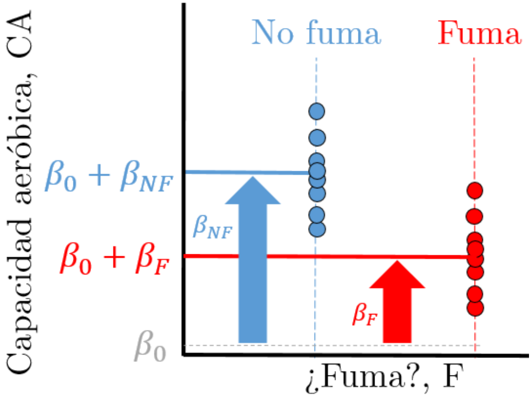
P1. Variables dummies sin interacciones

Recordemos que en el contraejemplo hay colineadlidad perfecta. Esto significa que en realidad no podemos estimar ni β_0, ni β_F, ni β_{NF}.

P1. Variables dummies sin interacciones

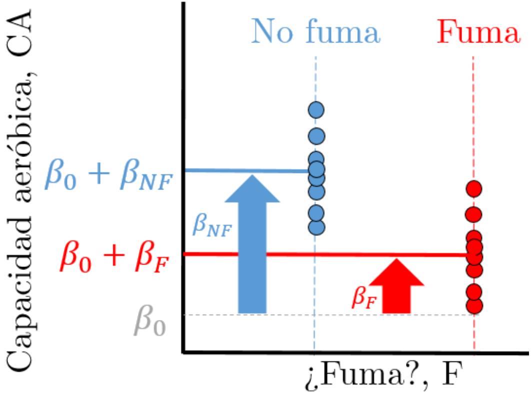
Recordemos que en el contraejemplo hay colineadlidad perfecta. Esto significa que en realidad no podemos estimar ni β_0, ni β_F, ni β_{NF}. El problema de identificación consiste en que existen infinitas combinaciones de estos parámetros que podrían cumplir con las propiedades de la muestra.

Recordemos que en el contraejemplo hay colineadlidad perfecta. Esto significa que en realidad no podemos estimar ni β_0, ni β_F, ni β_{NF}. El problema de identificación consiste en que existen infinitas combinaciones de estos parámetros que podrían cumplir con las propiedades de la muestra.

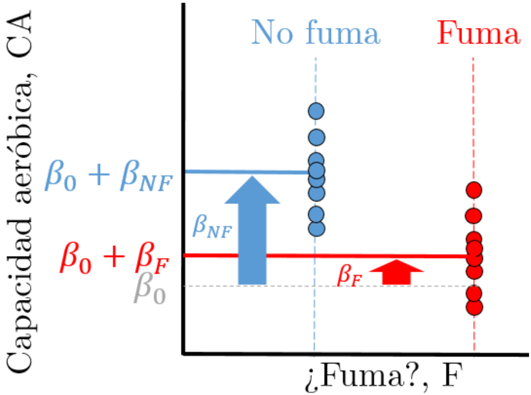


P1. Variables dummies sin interacciones

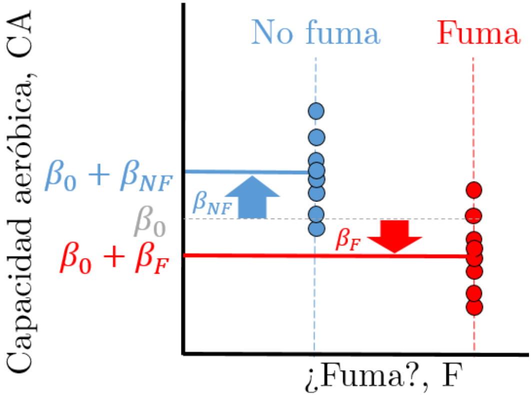
Recordemos que en el contraejemplo hay colineadlidad perfecta. Esto significa que en realidad no podemos estimar ni β_0, ni β_F, ni β_{NF}. El problema de identificación consiste en que existen infinitas combinaciones de estos parámetros que podrían cumplir con las propiedades de la muestra.



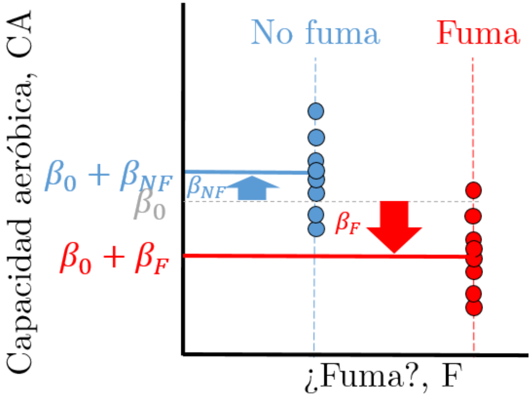
Recordemos que en el contraejemplo hay colineadlidad perfecta. Esto significa que en realidad no podemos estimar ni β_0, ni β_F, ni β_{NF}. El problema de identificación consiste en que existen infinitas combinaciones de estos parámetros que podrían cumplir con las propiedades de la muestra.



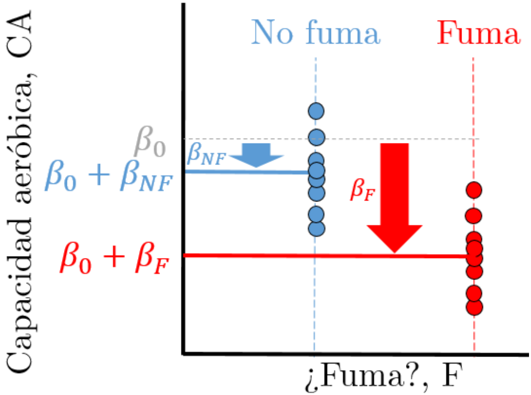
Recordemos que en el contraejemplo hay colineadlidad perfecta. Esto significa que en realidad no podemos estimar ni β_0, ni β_F, ni β_{NF}. El problema de identificación consiste en que existen infinitas combinaciones de estos parámetros que podrían cumplir con las propiedades de la muestra.



Recordemos que en el contraejemplo hay colineadlidad perfecta. Esto significa que en realidad no podemos estimar ni β_0, ni β_F, ni β_{NF}. El problema de identificación consiste en que existen infinitas combinaciones de estos parámetros que podrían cumplir con las propiedades de la muestra.

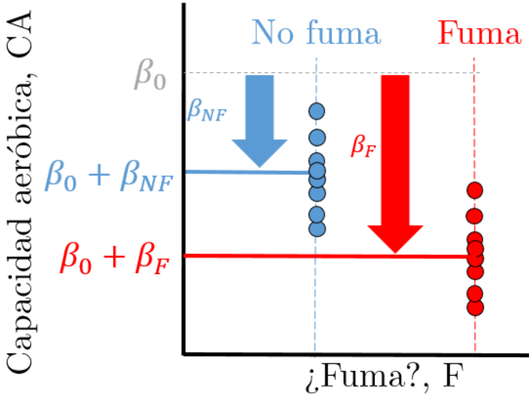


Recordemos que en el contraejemplo hay colineadlidad perfecta. Esto significa que en realidad no podemos estimar ni β_0, ni β_F, ni β_{NF}. El problema de identificación consiste en que existen infinitas combinaciones de estos parámetros que podrían cumplir con las propiedades de la muestra.



P1. Variables dummies sin interacciones

Recordemos que en el contraejemplo hay colineadlidad perfecta. Esto significa que en realidad no podemos estimar ni β_0, ni β_F, ni β_{NF}. El problema de identificación consiste en que existen infinitas combinaciones de estos parámetros que podrían cumplir con las propiedades de la muestra.



P2. Variables dummies sin interacciones

“Me interesa estimar la condición aeróbica promedio de fumadores y no fumadores de todos los rangos etarios.

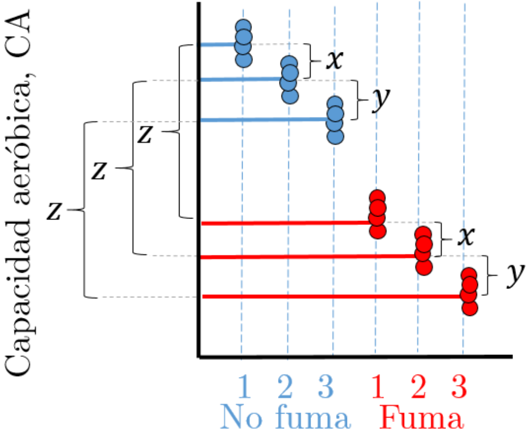
“Me interesa estimar la condición aeróbica promedio de fumadores y no fumadores de todos los rangos etarios. Algunos médicos afirman que la diferencia promedio entre fumadores y no fumadores es aproximadamente la misma para todos los grupos etarios.

P2. Variables dummies sin interacciones

"Me interesa estimar la condición aeróbica promedio de fumadores y no fumadores de todos los rangos etarios. Algunos médicos afirman que la diferencia promedio entre fumadores y no fumadores es aproximadamente la misma para todos los grupos etarios. De la misma forma, las diferencias promedio entre rangos etarios es aproximadamente la misma para fumadores y no fumadores. Podríamos asumir estas diferencias como constantes".

a. Hacer un bosquejo de cómo se distribuye la muestra en cuanto a condición aeróbica, clasificando por rango etáreo (RE_n) y consumo de tabaco (F_n) a la vez. Mostrar dónde está el promedio muestral.

a. Hacer un bosquejo de cómo se distribuye la muestra en cuanto a condición aeróbica, clasificando por rango etáreo (RE_n) y consumo de tabaco (F_n) a la vez. Mostrar dónde está el promedio muestral.



b. Plantear dos ejemplos modelos de regresión lineal **sin colinealidad perfecta**.

Ejemplo 1: $CA_n = \beta_0 + \beta_F F_n + \beta_2 RE_n^2 + \beta_3 RE_n^3 + U_n$

No hay colinealidad perfecta pues no hay ningún regresor que sea colineal con el resto:

$$\nexists(\lambda_1, \lambda_2, \lambda_3) \in \mathbb{R}^3 \mid \vec{1} = \lambda_1 F + \lambda_2 RE^2 + \lambda_3 RE^3$$

b. Plantear dos ejemplos modelos de regresión lineal **sin colinealidad perfecta**.

Ejemplo 2: $CA_n = \beta_{NF} NF_n + \beta_F F_n + \beta_2 RE_n^2 + \beta_3 RE_n^3 + U_n$

No hay colinealidad perfecta pues no hay ningún regresor que sea colineal con el resto:

$$\nexists (\lambda_1, \lambda_2, \lambda_3) \in \mathbb{R}^3 \mid NF = \lambda_1 F + \lambda_2 RE^2 + \lambda_3 RE^3$$

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   i. Calcule $\mathbb{E}[CA_n|F_n, RE_n]$ para todas las combinaciones posibles.

Ejemplo 1: $CA_n = \beta_0 + \beta_F F_n + \beta_2 RE_n^2 + \beta_3 RE_n^3 + U_n$

$$\mathbb{E}[CA_n|F_n = 1, RE_n^1 = 1] = \beta_0 + \beta_F$$
$$\mathbb{E}[CA_n|F_n = 1, RE_n^2 = 1] = \beta_0 + \beta_F + \beta_2$$
$$\mathbb{E}[CA_n|F_n = 1, RE_n^3 = 1] = \beta_0 + \beta_F + \beta_3$$
$$\mathbb{E}[CA_n|F_n = 0, RE_n^1 = 1] = \beta_0$$
$$\mathbb{E}[CA_n|F_n = 0, RE_n^2 = 1] = \beta_0 + \beta_2$$
$$\mathbb{E}[CA_n|F_n = 0, RE_n^3 = 1] = \beta_0 + \beta_3$$

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   ii. Calcule la diferencia esperada en condición aeróbica entre un fumador y un no fumador para todos los rangos etarios.

Ejemplo 1: $CA_n = \beta_0 + \beta_F F_n + \beta_2 RE_n^2 + \beta_3 RE_n^3 + U_n$

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   ii. Calcule la diferencia esperada en condición aeróbica entre un fumador y un no fumador para todos los rangos etarios.

Ejemplo 1: $CA_n = \beta_0 + \beta_F F_n + \beta_2 RE_n^2 + \beta_3 RE_n^3 + U_n$

$$\mathbb{E}[CA_n | F_n = 1, RE_n^1 = 1] - \mathbb{E}[CA_n | F_n = 0, RE_n^1 = 1] = (\cancel{\beta_0} + \beta_F) - \cancel{\beta_0}$$

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   ii. Calcule la diferencia esperada en condición aeróbica entre un fumador y un no fumador para todos los rangos etarios.

Ejemplo 1: $CA_n = \beta_0 + \beta_F F_n + \beta_2 RE_n^2 + \beta_3 RE_n^3 + U_n$

$$\mathbb{E}[CA_n|F_n = 1, RE_n^1 = 1] - \mathbb{E}[CA_n|F_n = 0, RE_n^1 = 1] = (\cancel{\beta_0} + \beta_F) - \cancel{\beta_0}$$

$$\mathbb{E}[CA_n|F_n = 1, RE_n^1 = 2] - \mathbb{E}[CA_n|F_n = 0, RE_n^1 = 2] = (\cancel{\beta_0} + \beta_F + \cancel{\beta_2}) - (\cancel{\beta_0} + \cancel{\beta_2})$$

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   ii. Calcule la diferencia esperada en condición aeróbica entre un fumador y un no fumador para todos los rangos etarios.

Ejemplo 1: $CA_n = \beta_0 + \beta_F F_n + \beta_2 RE_n^2 + \beta_3 RE_n^3 + U_n$

$$\mathbb{E}[CA_n|F_n = 1, RE_n^1 = 1] - \mathbb{E}[CA_n|F_n = 0, RE_n^1 = 1] = (\cancel{\beta_0} + \beta_F) - \cancel{\beta_0}$$

$$\mathbb{E}[CA_n|F_n = 1, RE_n^1 = 2] - \mathbb{E}[CA_n|F_n = 0, RE_n^1 = 2] = (\cancel{\beta_0} + \beta_F + \cancel{\beta_2}) - (\cancel{\beta_0} + \cancel{\beta_2})$$

$$\mathbb{E}[CA_n|F_n = 1, RE_n^1 = 3] - \mathbb{E}[CA_n|F_n = 0, RE_n^1 = 3] = (\cancel{\beta_0} + \beta_F + \cancel{\beta_3}) - (\cancel{\beta_0} + \cancel{\beta_3})$$

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
    iii. Calcule la diferencia esperada en condición aeróbica entre fumadores de distintos rangos etarios y no fumadores de distintos rangos etarios.

Ejemplo 1: $CA_n = \beta_0 + \beta_F F_n + \beta_2 RE_n^2 + \beta_3 RE_n^3 + U_n$

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   iii. Calcule la diferencia esperada en condición aeróbica entre fumadores de distintos rangos etarios y no fumadores de distintos rangos etarios.

Ejemplo 1: $CA_n = \beta_0 + \beta_F F_n + \beta_2 RE_n^2 + \beta_3 RE_n^3 + U_n$

$$\mathbb{E}[CA_n|F_n = 1, RE_n^2 = 1] - \mathbb{E}[CA_n|F_n = 1, RE_n^1 = 1] = (\cancel{\beta_0} + \cancel{\beta_F} + \beta_2) - (\cancel{\beta_0} + \cancel{\beta_F})$$

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):

iii. Calcule la diferencia esperada en condición aeróbica entre fumadores de distintos rangos etarios y no fumadores de distintos rangos etarios.

Ejemplo 1: $CA_n = \beta_0 + \beta_F F_n + \beta_2 RE_n^2 + \beta_3 RE_n^3 + U_n$

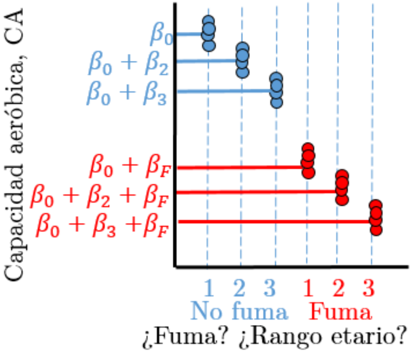
$$\mathbb{E}[CA_n|F_n = 1, RE_n^2 = 1] - \mathbb{E}[CA_n|F_n = 1, RE_n^1 = 1] = (\cancel{\beta_0} + \cancel{\beta_F} + \beta_2) - (\cancel{\beta_0} + \cancel{\beta_F})$$

$$\mathbb{E}[CA_n|F_n = 1, RE_n^3 = 1] - \mathbb{E}[CA_n|F_n = 1, RE_n^1 = 1] = (\cancel{\beta_0} + \cancel{\beta_F} + \beta_3) - (\cancel{\beta_0} + \cancel{\beta_F})$$

Para quienes no fuman pasa lo mismo (propuesto).

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):

   iv. En cuanto a los parámetros poblacionales de su modelo, especule sus signos (+/-), interprete su significado y muéstrelos gráficamente.



$\beta_0 > 0$, $\beta_F, \beta_2, \beta_3 < 0$. Las sumas de coeficientes mostradas en el gráfico muestran los promedios poblacionales de CA de los 6 grupos.

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   i. Calcule $\mathbb{E}[CA_n|F_n, RE_n]$ para todas las combinaciones posibles.

Ejemplo 2: $CA_n = \beta_{NF} NF_n + \beta_F F_n + \beta_2 RE_n^2 + \beta_3 RE_n^3 + U_n$

$$\mathbb{E}[CA_n|F_n = 1, RE_n^1 = 1] = \beta_F$$
$$\mathbb{E}[CA_n|F_n = 1, RE_n^2 = 1] = \beta_F + \beta_2$$
$$\mathbb{E}[CA_n|F_n = 1, RE_n^3 = 1] = \beta_F + \beta_3$$
$$\mathbb{E}[CA_n|F_n = 0, RE_n^1 = 1] = \beta_{NF}$$
$$\mathbb{E}[CA_n|F_n = 0, RE_n^2 = 1] = \beta_{NF} + \beta_2$$
$$\mathbb{E}[CA_n|F_n = 0, RE_n^3 = 1] = \beta_{NF} + \beta_3$$

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   ii. Calcule la diferencia esperada en condición aeróbica entre un fumador y un no fumador para todos los rangos etarios.

Ejemplo 2: $CA_n = \beta_{NF} NF_n + \beta_F F_n + \beta_2 RE_n^2 + \beta_3 RE_n^3 + U_n$

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   ii. Calcule la diferencia esperada en condición aeróbica entre un fumador y un no fumador para todos los rangos etarios.

Ejemplo 2: $CA_n = \beta_{NF}NF_n + \beta_F F_n + \beta_2 RE_n^2 + \beta_3 RE_n^3 + U_n$

$$\mathbb{E}[CA_n|F_n = 1, RE_n^1 = 1] - \mathbb{E}[CA_n|F_n = 0, RE_n^1 = 1] = \beta_F - \beta_{NF}$$

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   ii. Calcule la diferencia esperada en condición aeróbica entre un fumador y un no fumador para todos los rangos etarios.

Ejemplo 2: $CA_n = \beta_{NF} NF_n + \beta_F F_n + \beta_2 RE_n^2 + \beta_3 RE_n^3 + U_n$

$$\mathbb{E}[CA_n | F_n = 1, RE_n^1 = 1] - \mathbb{E}[CA_n | F_n = 0, RE_n^1 = 1] = \beta_F - \beta_{NF}$$

$$\mathbb{E}[CA_n | F_n = 1, RE_n^2 = 1] - \mathbb{E}[CA_n | F_n = 0, RE_n^2 = 1] = (\beta_F + \cancel{\beta_2}) - (\beta_{NF} + \cancel{\beta_2})$$

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   ii. Calcule la diferencia esperada en condición aeróbica entre un fumador y un no fumador para todos los rangos etarios.

Ejemplo 2: $CA_n = \beta_{NF} NF_n + \beta_F F_n + \beta_2 RE_n^2 + \beta_3 RE_n^3 + U_n$

$$\mathbb{E}[CA_n | F_n = 1, RE_n^1 = 1] - \mathbb{E}[CA_n | F_n = 0, RE_n^1 = 1] = \beta_F - \beta_{NF}$$

$$\mathbb{E}[CA_n | F_n = 1, RE_n^2 = 1] - \mathbb{E}[CA_n | F_n = 0, RE_n^2 = 1] = (\beta_F + \cancel{\beta_2}) - (\beta_{NF} + \cancel{\beta_2})$$

$$\mathbb{E}[CA_n | F_n = 1, RE_n^3 = 1] - \mathbb{E}[CA_n | F_n = 0, RE_n^3 = 1] = (\beta_F + \cancel{\beta_3}) - (\beta_{NF} + \cancel{\beta_3})$$

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):

   iii. Calcule la diferencia esperada en condición aeróbica entre fumadores de distintos rangos etarios y no fumadores de distintos rangos etarios.

Ejemplo 2: $CA_n = \beta_{NF} NF_n + \beta_F F_n + \beta_2 RE_n^2 + \beta_3 RE_n^3 + U_n$

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
    iii. Calcule la diferencia esperada en condición aeróbica entre fumadores de distintos rangos etarios y no fumadores de distintos rangos etarios.

Ejemplo 2: $CA_n = \beta_{NF} NF_n + \beta_F F_n + \beta_2 RE_n^2 + \beta_3 RE_n^3 + U_n$

$$\mathbb{E}[CA_n | F_n = 1, RE_n^2 = 1] - \mathbb{E}[CA_n | F_n = 1, RE_n^1 = 1] = (\cancel{\beta_F} + \beta_2) - \cancel{\beta_F}$$

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):

iii. Calcule la diferencia esperada en condición aeróbica entre fumadores de distintos rangos etarios y no fumadores de distintos rangos etarios.

Ejemplo 2: $CA_n = \beta_{NF} NF_n + \beta_F F_n + \beta_2 RE_n^2 + \beta_3 RE_n^3 + U_n$

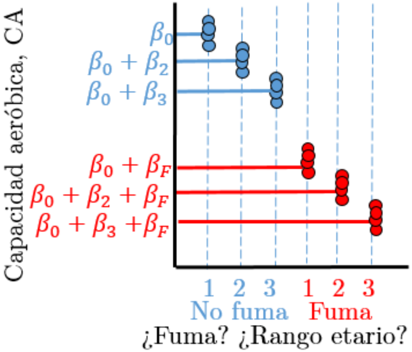
$$\mathbb{E}[CA_n | F_n = 1, RE_n^2 = 1] - \mathbb{E}[CA_n | F_n = 1, RE_n^1 = 1] = (\cancel{\beta_F} + \beta_2) - \cancel{\beta_F}$$

$$\mathbb{E}[CA_n | F_n = 1, RE_n^3 = 1] - \mathbb{E}[CA_n | F_n = 1, RE_n^1 = 1] = (\cancel{\beta_F} + \beta_3) - \cancel{\beta_F}$$

Para quienes no fuman pasa lo mismo (propuesto).

P2. Variables dummies sin interacciones

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   iv. En cuanto a los parámetros poblacionales de su modelo, especule sus signos (+/-), interprete su significado y muéstrelos gráficamente.



$\beta_0 > 0$, $\beta_F, \beta_2, \beta_3 < 0$. Las sumas de coeficientes mostradas en el gráfico muestran los promedios poblacionales de CA de los 6 grupos.

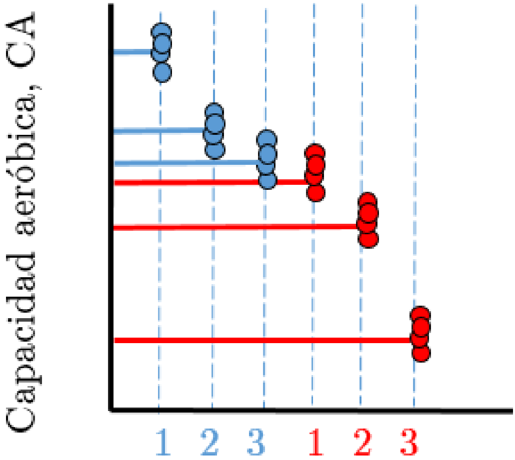
"Creo que lo que te dije la vez pasada es incorrecto. Sería interesante estimar el promedio de los seis grupos (adolescentes fumadores, adolescentes no fumadores, adultos fumadores, etc.) sin asumir esas diferencias constantes". Repita el ejercicio de la parte anterior.

P3. Interacciones dummy-dummy

a. Hacer un bosquejo de cómo se distribuye la muestra en cuanto a condición aeróbica, clasificando por rango etáreo (RE_n) y consumo de tabaco (F_n) a la vez. Mostrar dónde está el promedio muestral.

P3. Interacciones dummy-dummy

a. Hacer un bosquejo de cómo se distribuye la muestra en cuanto a condición aeróbica, clasificando por rango etáreo (RE_n) y consumo de tabaco (F_n) a la vez. Mostrar dónde está el promedio muestral.



b. Plantear dos ejemplos modelos de regresión lineal **sin colinealidad perfecta**.

Ejemplo 2:

$$\begin{aligned} CA_n = \beta_0 &+ \beta_F^1 F_n RE_n^1 + \beta_F^2 F_n RE_n^2 \\ &+ \beta_F^3 F_n RE_n^3 + \beta_{NF}^1 NF_n RE_n^1 + \beta_F^2 NF_n RE_n^2 + U_n \end{aligned}$$

b. Plantear dos ejemplos modelos de regresión lineal **sin colinealidad perfecta**.

Ejemplo 2:

$$CA_n = \beta_0 + \beta_F^1 F_n RE_n^1 + \beta_F^2 F_n RE_n^2 \\ + \beta_F^3 F_n RE_n^3 + \beta_{NF}^1 NF_n RE_n^1 + \beta_F^2 NF_n RE_n^2 + U_n$$

No hay colinealidad perfecta pues no hay ningún regresor que sea colineal con el resto (los 6 grupos son excluyentes):

$$\nexists (\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5) \in \mathbb{R}^5 |$$

$$\vec{1} = \lambda_1 F \cdot RE^1 + \lambda_2 F \cdot RE^2 + \lambda_3 F \cdot RE^3 + \lambda_4 NF \cdot RE^1 + \lambda_5 NRE^2$$

"·" en este caso representa un producto componente a componente.

b. Plantear dos ejemplos modelos de regresión lineal **sin colinealidad perfecta**.

Ejemplo 1:

$$
\begin{aligned}
CA_n = {} & \beta_F^1 F_n RE_n^1 + \beta_F^2 F_n RE_n^2 + \beta_F^3 F_n RE_n^3 \\
& + \beta_{NF}^1 NF_n RE_n^1 + \beta_F^2 NF_n RE_n^2 + \beta_F^3 NF_n RE_n^3 + U_n
\end{aligned}
$$

b. Plantear dos ejemplos modelos de regresión lineal **sin colinealidad perfecta**.

Ejemplo 1:

$$
\begin{aligned}
CA_n = {} & \beta_F^1 F_n RE_n^1 + \beta_F^2 F_n RE_n^2 + \beta_F^3 F_n RE_n^3 \\
& + \beta_{NF}^1 NF_n RE_n^1 + \beta_F^2 NF_n RE_n^2 + \beta_F^3 NF_n RE_n^3 + U_n
\end{aligned}
$$

No hay colinealidad perfecta pues no hay ningún regresor que sea colineal con el resto (los 6 grupos son excluyentes):

$$\nexists(\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5) \in \mathbb{R}^5 |$$

$$F \cdot RE^1 = \lambda_1 F \cdot RE^2 + \lambda_2 F \cdot RE^3 + \lambda_3 NF \cdot RE^1 + \lambda_4 NRE^2 + \lambda_5 NF \cdot RE^3$$

"$\cdot$" en este caso representa un producto componente a componente.

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   i. Calcule $\mathbb{E}[CA_n|F_n, RE_n]$ para todas las combinaciones posibles.

Ejemplo 1:

$$
\begin{aligned}
CA_n &= \beta_F^1 F_n RE_n^1 + \beta_F^2 F_n RE_n^2 + \beta_F^3 F_n RE_n^3 \\
&\quad + \beta_{NF}^1 NF_n RE_n^1 + \beta_F^2 NF_n RE_n^2 + \beta_F^3 NF_n RE_n^3 + U_n
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}[CA_n|F_n = 1, RE_n^1 = 1] &= \beta_F^1 \\
\mathbb{E}[CA_n|F_n = 1, RE_n^2 = 1] &= \beta_F^2 \\
\mathbb{E}[CA_n|F_n = 1, RE_n^3 = 1] &= \beta_F^3 \\
\mathbb{E}[CA_n|F_n = 0, RE_n^1 = 1] &= \beta_{NF}^1 \\
\mathbb{E}[CA_n|F_n = 0, RE_n^2 = 1] &= \beta_{NF}^2 \\
\mathbb{E}[CA_n|F_n = 0, RE_n^3 = 1] &= \beta_{NF}^3
\end{aligned}
$$

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   ii. Calcule la diferencia esperada en condición aeróbica entre un fumador y un no fumador para todos los rangos etarios.

Ejemplo 1:

$$\begin{aligned} CA_n = & \beta_F^1 F_n RE_n^1 + \beta_F^2 F_n RE_n^2 + \beta_F^3 F_n RE_n^3 \\ & + \beta_{NF}^1 NF_n RE_n^1 + \beta_F^2 NF_n RE_n^2 + \beta_F^3 NF_n RE_n^3 + U_n \end{aligned}$$

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   ii. Calcule la diferencia esperada en condición aeróbica entre un fumador y un no fumador para todos los rangos etarios.

Ejemplo 1:

$$\begin{aligned} CA_n = &\beta_F^1 F_n RE_n^1 + \beta_F^2 F_n RE_n^2 + \beta_F^3 F_n RE_n^3 \\ &+ \beta_{NF}^1 NF_n RE_n^1 + \beta_F^2 NF_n RE_n^2 + \beta_F^3 NF_n RE_n^3 + U_n \end{aligned}$$

$$\mathbb{E}[CA_n | F_n = 1, RE_n^1 = 1] - \mathbb{E}[CA_n | F_n = 0, RE_n^1 = 1] = \beta_F^1 - \beta_{NF}^1$$

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
    ii. Calcule la diferencia esperada en condición aeróbica entre un fumador y un no fumador para todos los rangos etarios.

Ejemplo 1:

$$\begin{aligned} CA_n &= \beta_F^1 F_n RE_n^1 + \beta_F^2 F_n RE_n^2 + \beta_F^3 F_n RE_n^3 \\ &\quad + \beta_{NF}^1 NF_n RE_n^1 + \beta_F^2 NF_n RE_n^2 + \beta_F^3 NF_n RE_n^3 + U_n \end{aligned}$$

$$\mathbb{E}[CA_n | F_n = 1, RE_n^1 = 1] - \mathbb{E}[CA_n | F_n = 0, RE_n^1 = 1] = \beta_F^1 - \beta_{NF}^1$$

$$\mathbb{E}[CA_n | F_n = 1, RE_n^2 = 1] - \mathbb{E}[CA_n | F_n = 0, RE_n^2 = 1] = \beta_F^2 - \beta_{NF}^2$$

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):

   ii. Calcule la diferencia esperada en condición aeróbica entre un fumador y un no fumador para todos los rangos etarios.

Ejemplo 1:

$$\begin{aligned} CA_n = {} & \beta_F^1 F_n RE_n^1 + \beta_F^2 F_n RE_n^2 + \beta_F^3 F_n RE_n^3 \\ & + \beta_{NF}^1 NF_n RE_n^1 + \beta_F^2 NF_n RE_n^2 + \beta_F^3 NF_n RE_n^3 + U_n \end{aligned}$$

$$\mathbb{E}[CA_n | F_n = 1, RE_n^1 = 1] - \mathbb{E}[CA_n | F_n = 0, RE_n^1 = 1] = \beta_F^1 - \beta_{NF}^1$$

$$\mathbb{E}[CA_n | F_n = 1, RE_n^2 = 1] - \mathbb{E}[CA_n | F_n = 0, RE_n^2 = 1] = \beta_F^2 - \beta_{NF}^2$$

$$\mathbb{E}[CA_n | F_n = 1, RE_n^3 = 1] - \mathbb{E}[CA_n | F_n = 0, RE_n^3 = 1] = \beta_F^3 - \beta_{NF}^3$$

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   iii. Calcule la diferencia esperada en condición aeróbica entre fumadores de distintos rangos etarios y no fumadores de distintos rangos etarios.

Ejemplo 1:

$$
\begin{aligned}
CA_n = {} & \beta_F^1 F_n RE_n^1 + \beta_F^2 F_n RE_n^2 + \beta_F^3 F_n RE_n^3 \\
& + \beta_{NF}^1 NF_n RE_n^1 + \beta_F^2 NF_n RE_n^2 + \beta_F^3 NF_n RE_n^3 + U_n
\end{aligned}
$$

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   iii. Calcule la diferencia esperada en condición aeróbica entre fumadores de distintos rangos etarios y no fumadores de distintos rangos etarios.

Ejemplo 1:

$$\begin{aligned} CA_n = {} & \beta_F^1 F_n RE_n^1 + \beta_F^2 F_n RE_n^2 + \beta_F^3 F_n RE_n^3 \\ & + \beta_{NF}^1 NF_n RE_n^1 + \beta_F^2 NF_n RE_n^2 + \beta_F^3 NF_n RE_n^3 + U_n \end{aligned}$$

$$\mathbb{E}[CA_n | F_n = 1, RE_n^2 = 1] - \mathbb{E}[CA_n | F_n = 1, RE_n^1 = 1] = \beta_F^2 - \beta_F^1$$

P3. Interacciones dummy-dummy

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   iii. Calcule la diferencia esperada en condición aeróbica entre fumadores de distintos rangos etarios y no fumadores de distintos rangos etarios.

Ejemplo 1:

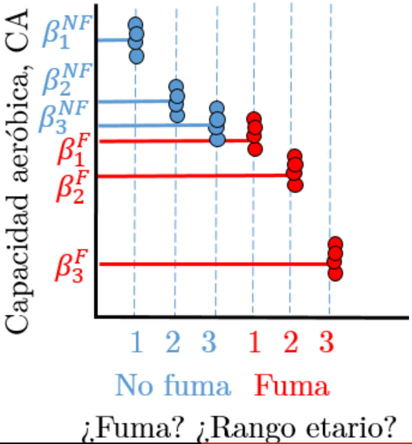
$$
\begin{aligned}
CA_n &= \beta_F^1 F_n RE_n^1 + \beta_F^2 F_n RE_n^2 + \beta_F^3 F_n RE_n^3 \\
&\quad + \beta_{NF}^1 NF_n RE_n^1 + \beta_F^2 NF_n RE_n^2 + \beta_F^3 NF_n RE_n^3 + U_n
\end{aligned}
$$

$$\mathbb{E}[CA_n | F_n = 1, RE_n^2 = 1] - \mathbb{E}[CA_n | F_n = 1, RE_n^1 = 1] = \beta_F^2 - \beta_F^1$$

$$\mathbb{E}[CA_n | F_n = 1, RE_n^3 = 1] - \mathbb{E}[CA_n | F_n = 1, RE_n^1 = 1] = \beta_F^3 - \beta_F^1$$

Para quienes no fuman es análogo (propuesto).

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   iv. En cuanto a los parámetros poblacionales de su modelo, especule sus signos (+/-), interprete su significado y muéstrelos gráficamente.



c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   i. Calcule $\mathbb{E}[CA_n|F_n, RE_n]$ para todas las combinaciones posibles.

Ejemplo 2:

$$CA_n = \beta_0 + \beta_F^1 F_n RE_n^1 + \beta_F^2 F_n RE_n^2 + \beta_F^3 F_n RE_n^3 + \beta_{NF}^1 NF_n RE_n^1 + \beta_F^2 NF_n RE_n^2 + U_n$$

$$\mathbb{E}[CA_n|F_n = 1, RE_n^1 = 1] = \beta_0 + \beta_F^1$$
$$\mathbb{E}[CA_n|F_n = 1, RE_n^2 = 1] = \beta_0 + \beta_F^2$$
$$\mathbb{E}[CA_n|F_n = 1, RE_n^3 = 1] = \beta_0 + \beta_F^3$$
$$\mathbb{E}[CA_n|F_n = 0, RE_n^1 = 1] = \beta_0 + \beta_{NF}^1$$
$$\mathbb{E}[CA_n|F_n = 0, RE_n^2 = 1] = \beta_0 + \beta_{NF}^2$$
$$\mathbb{E}[CA_n|F_n = 0, RE_n^3 = 1] = \beta_0$$

P3. Interacciones dummy-dummy

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   ii. Calcule la diferencia esperada en condición aeróbica entre un fumador y un no fumador para todos los rangos etarios.

Ejemplo 2:

$$CA_n = \beta_0 + \beta_F^1 F_n RE_n^1 + \beta_F^2 F_n RE_n^2 + \beta_F^3 F_n RE_n^3$$
$$+ \beta_{NF}^1 NF_n RE_n^1 + \beta_F^2 NF_n RE_n^2 + U_n$$

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):

   ii. Calcule la diferencia esperada en condición aeróbica entre un fumador y un no fumador para todos los rangos etarios.

Ejemplo 2:

$$
\begin{aligned}
CA_n = \beta_0 + \beta_F^1 F_n RE_n^1 + \beta_F^2 F_n RE_n^2 + \beta_F^3 F_n RE_n^3 \\
+ \beta_{NF}^1 NF_n RE_n^1 + \beta_F^2 NF_n RE_n^2 + U_n
\end{aligned}
$$

$$\mathbb{E}[CA_n|F_n = 1, RE_n^1 = 1] - \mathbb{E}[CA_n|F_n = 0, RE_n^1 = 1] = (\cancel{\beta_0} + \beta_F^1) - (\cancel{\beta_0} + \beta_{NF}^1)$$

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   ii. Calcule la diferencia esperada en condición aeróbica entre un fumador y un no fumador para todos los rangos etarios.

Ejemplo 2:

$$CA_n = \beta_0 + \beta_F^1 F_n RE_n^1 + \beta_F^2 F_n RE_n^2 + \beta_F^3 F_n RE_n^3$$
$$+ \beta_{NF}^1 NF_n RE_n^1 + \beta_F^2 NF_n RE_n^2 + U_n$$

$$\mathbb{E}[CA_n|F_n = 1, RE_n^1 = 1] - \mathbb{E}[CA_n|F_n = 0, RE_n^1 = 1] = (\cancel{\beta_0} + \beta_F^1) - (\cancel{\beta_0} + \beta_{NF}^1)$$

$$\mathbb{E}[CA_n|F_n = 1, RE_n^2 = 1] - \mathbb{E}[CA_n|F_n = 0, RE_n^2 = 1] = (\cancel{\beta_0} + \beta_F^2) - (\cancel{\beta_0} + \beta_{NF}^2)$$

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):

ii. Calcule la diferencia esperada en condición aeróbica entre un fumador y un no fumador para todos los rangos etarios.

Ejemplo 2:

$$CA_n = \beta_0 + \beta_F^1 F_n RE_n^1 + \beta_F^2 F_n RE_n^2 + \beta_F^3 F_n RE_n^3$$
$$+ \beta_{NF}^1 NF_n RE_n^1 + \beta_F^2 NF_n RE_n^2 + U_n$$

$$\mathbb{E}[CA_n|F_n = 1, RE_n^1 = 1] - \mathbb{E}[CA_n|F_n = 0, RE_n^1 = 1] = (\cancel{\beta_0} + \beta_F^1) - (\cancel{\beta_0} + \beta_{NF}^1)$$

$$\mathbb{E}[CA_n|F_n = 1, RE_n^2 = 1] - \mathbb{E}[CA_n|F_n = 0, RE_n^2 = 1] = (\cancel{\beta_0} + \beta_F^2) - (\cancel{\beta_0} + \beta_{NF}^2)$$

$$\mathbb{E}[CA_n|F_n = 1, RE_n^3 = 1] - \mathbb{E}[CA_n|F_n = 0, RE_n^3 = 1] = (\cancel{\beta_0} + \beta_F^3) - \cancel{\beta_0}$$

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
    iii. Calcule la diferencia esperada en condición aeróbica entre fumadores de distintos rangos etarios y no fumadores de distintos rangos etarios.

Ejemplo 2:

$$\begin{aligned} CA_n = \beta_0 + \beta_F^1 F_n RE_n^1 + \beta_F^2 F_n RE_n^2 + \beta_F^3 F_n RE_n^3 \\ + \beta_{NF}^1 NF_n RE_n^1 + \beta_F^2 NF_n RE_n^2 + U_n \end{aligned}$$

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   iii. Calcule la diferencia esperada en condición aeróbica entre fumadores de distintos rangos etarios y no fumadores de distintos rangos etarios.

Ejemplo 2:

$$CA_n = \beta_0 + \beta_F^1 F_n RE_n^1 + \beta_F^2 F_n RE_n^2 + \beta_F^3 F_n RE_n^3 + \beta_{NF}^1 NF_n RE_n^1 + \beta_F^2 NF_n RE_n^2 + U_n$$

$$\mathbb{E}[CA_n|F_n = 1, RE_n^2 = 1] - \mathbb{E}[CA_n|F_n = 1, RE_n^1 = 1] = (\cancel{\beta_0} + \beta_F^2) - (\cancel{\beta_0} + \beta_F^1)$$

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   iii. Calcule la diferencia esperada en condición aeróbica entre fumadores de distintos rangos etarios y no fumadores de distintos rangos etarios.

Ejemplo 2:

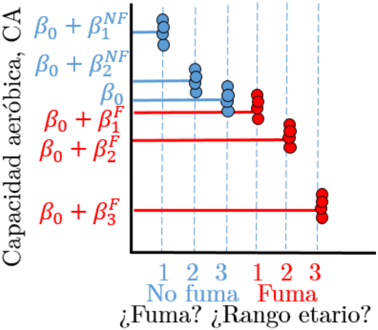
$$CA_n = \beta_0 + \beta_F^1 F_n RE_n^1 + \beta_F^2 F_n RE_n^2 + \beta_F^3 F_n RE_n^3 \\ + \beta_{NF}^1 NF_n RE_n^1 + \beta_F^2 NF_n RE_n^2 + U_n$$

$$\mathbb{E}[CA_n | F_n = 1, RE_n^2 = 1] - \mathbb{E}[CA_n | F_n = 1, RE_n^1 = 1] = (\cancel{\beta_0} + \beta_F^2) - (\cancel{\beta_0} + \beta_F^1)$$

$$\mathbb{E}[CA_n | F_n = 1, RE_n^3 = 1] - \mathbb{E}[CA_n | F_n = 1, RE_n^1 = 1] = (\cancel{\beta_0} + \beta_F^3) - (\cancel{\beta_0} + \beta_F^1)$$

Para quienes no fuman es análogo (propuesto).

c. Para los tres modelos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   iv. En cuanto a los parámetros poblacionales de su modelo, especule sus signos (+/-), interprete su significado y muéstrelos gráficamente.



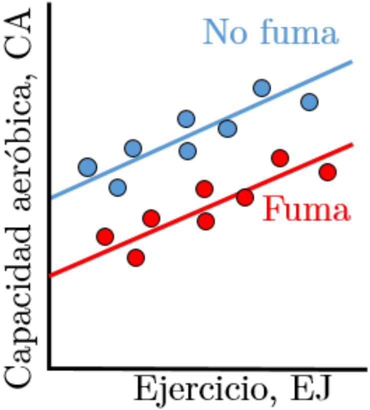
$\beta_0, \beta_1^{NF}, \beta_2^{NF} > 0$, $\beta_1^{F}, \beta_2^{F}, \beta_3^{F} < 0$. Las sumas de coeficientes mostradas en el gráfico muestran los promedios poblacionales de CA de los 6 grupos.

P4. Interceptos distintos con dummies

"Algunos médicos afirman que las horas de ejercicio aeróbico semanal son una variable importante al momento de explicar la capacidad aeróbica de las personas. Además su efecto es el mismo, tanto para fumadores como para no fumadores".

P4. Interceptos distintos con dummies

a. Hacer un bosquejo de cómo se distribuye la muestra en cuanto a condición aeróbica y horas de ejercicio aeróbico semanal, separando a fumadores y no fumadores. En la misma figura grafique un modelo que cumpla con las características descritas por su amiga.



b. Plantear dos ejemplos modelos de regresión lineal **sin colinealidad perfecta** y un contraejemplo **con colinealidad perfecta** diseñados para estimar estos promedios. Explique por qué (no) hay colinealidad perfecta en cada caso y por qué este es un problema en el contraejemplo.

Ejemplo 1: $CA_n = \beta_0 + \beta_F F_n + \beta_{EJ} EJ_n + U_n$

b. Plantear dos ejemplos modelos de regresión lineal **sin colinealidad perfecta** y un contraejemplo **con colinealidad perfecta** diseñados para estimar estos promedios. Explique por qué (no) hay colinealidad perfecta en cada caso y por qué este es un problema en el contraejemplo.

Ejemplo 1: $CA_n = \beta_0 + \beta_F F_n + \beta_{EJ} EJ_n + U_n$

No hay colinealidad perfecta pues los regresores $\vec{1}$, F y EJ no son perfectamente colineales,

b. Plantear dos ejemplos modelos de regresión lineal **sin colinealidad perfecta** y un contraejemplo **con colinealidad perfecta** diseñados para estimar estos promedios. Explique por qué (no) hay colinealidad perfecta en cada caso y por qué este es un problema en el contraejemplo.

Ejemplo 1: $CA_n = \beta_0 + \beta_F F_n + \beta_{EJ} EJ_n + U_n$

No hay colinealidad perfecta pues los regresores $\vec{1}$, F y EJ no son perfectamente colineales, ie

$$\nexists \lambda_1, \lambda_2 \in \mathbb{R}^2 | \vec{1} = \lambda_1 F + \lambda_2 EJ$$

b. Plantear dos ejemplos modelos de regresión lineal **sin colinealidad perfecta** y un contraejemplo **con colinealidad perfecta** diseñados para estimar estos promedios. Explique por qué (no) hay colinealidad perfecta en cada caso y por qué este es un problema en el contraejemplo.

Ejemplo 2:

$$CA_n = \beta_{NF} NF_n + \beta_F F_n + \beta_{EJ} EJ_n + U_n$$

b. Plantear dos ejemplos modelos de regresión lineal **sin colinealidad perfecta** y un contraejemplo **con colinealidad perfecta** diseñados para estimar estos promedios. Explique por qué (no) hay colinealidad perfecta en cada caso y por qué este es un problema en el contraejemplo.

Ejemplo 2:

$$CA_n = \beta_{NF} NF_n + \beta_F F_n + \beta_{EJ} EJ_n + U_n$$

No hay colinealidad perfecta pues los regresores NF, F y EJ no son perfectamente colineales,

b. Plantear dos ejemplos modelos de regresión lineal **sin colinealidad perfecta** y un contraejemplo **con colinealidad perfecta** diseñados para estimar estos promedios. Explique por qué (no) hay colinealidad perfecta en cada caso y por qué este es un problema en el contraejemplo.

Ejemplo 2:

$$CA_n = \beta_{NF} NF_n + \beta_F F_n + \beta_{EJ} EJ_n + U_n$$

No hay colinealidad perfecta pues los regresores NF, F y EJ no son perfectamente colineales, ie

$$\nexists \lambda_1, \lambda_2 \in \mathbb{R}^2 | NF = \lambda_1 F + \lambda_2 EJ$$

b. Plantear dos ejemplos modelos de regresión lineal **sin colinealidad perfecta** y un contraejemplo **con colinealidad perfecta** diseñados para estimar estos promedios. Explique por qué (no) hay colinealidad perfecta en cada caso y por qué este es un problema en el contraejemplo.

Contraejemplo:

$$CA_n = \beta_0 + \beta_{NF} NF_n + \beta_F F_n + \beta_{EJ} EJ_n + U_n$$

Hay colinealidad perfecta pues el regresor $\vec{1}$ es colineal con F y NF. En efecto, $\vec{1} = F + NF$.

c. Para los dos ejemplos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   i. Calcule $\mathbb{E}[CA_n|F_n = 0]$, $\mathbb{E}[CA_n|F_n = 1]$.

Ejemplo 1: $CA_n = \beta_0 + \beta_F F_n + \beta_{EJ} EJ_n + U_n$

$$\mathbb{E}[CA_n|F_n = 0] = \beta_0 + \beta_{EJ} EJ_n$$

$$\mathbb{E}[CA_n|F_n = 1] = \beta_0 + \beta_F + \beta_{EJ} EJ_n$$

P4. Interceptos distintos con dummies

c. Para los dos ejemplos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   ii. Calcule la diferencia esperada en condición aeróbica entre un fumador y un no fumador que hacen la misma cantidad de ejercicio. Obtenga el efecto marginal del ejercicio aeróbico sobre la capacidad aeróbica tanto para fumadores como no fumadores. ¿Son iguales?

c. Para los dos ejemplos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   ii. Calcule la diferencia esperada en condición aeróbica entre un fumador y un no fumador que hacen la misma cantidad de ejercicio. Obtenga el efecto marginal del ejercicio aeróbico sobre la capacidad aeróbica tanto para fumadores como no fumadores. ¿Son iguales?

Diferencia esperada:

$$\begin{aligned}\mathbb{E}[CA_n|F_n = 1] - \mathbb{E}[CA_n|F_n = 1] &= (\cancel{\beta_0} + \beta_F + \cancel{\beta_{EJ}EJ_n}) - (\cancel{\beta_0} + \cancel{\beta_{EJ}EJ_n}) \\ &= \beta_F\end{aligned}$$

c. Para los dos ejemplos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   ii. Calcule la diferencia esperada en condición aeróbica entre un fumador y un no fumador que hacen la misma cantidad de ejercicio. Obtenga el efecto marginal del ejercicio aeróbico sobre la capacidad aeróbica tanto para fumadores como no fumadores. ¿Son iguales?

Diferencia esperada:

$$\begin{aligned}\mathbb{E}[CA_n|F_n=1]-\mathbb{E}[CA_n|F_n=1] &= (\cancel{\beta_0}+\beta_F+\cancel{\beta_{EJ}EJ_n})-(\cancel{\beta_0}+\cancel{\beta_{EJ}EJ_n}) \\ &= \beta_F\end{aligned}$$

Efecto marginal:

$$\frac{\partial\mathbb{E}[CA_n|F_n=0]}{\partial EJ_n}=\frac{\partial}{\partial EJ_n}(\beta_0+\beta_{EJ}EJ_n)=\beta_{EJ}$$

c. Para los dos ejemplos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   ii. Calcule la diferencia esperada en condición aeróbica entre un fumador y un no fumador que hacen la misma cantidad de ejercicio. Obtenga el efecto marginal del ejercicio aeróbico sobre la capacidad aeróbica tanto para fumadores como no fumadores. ¿Son iguales?

Diferencia esperada:

$$\mathbb{E}[CA_n|F_n = 1] - \mathbb{E}[CA_n|F_n = 1] = (\cancel{\beta_0} + \beta_F + \cancel{\beta_{EJ} EJ_n}) - (\cancel{\beta_0} + \cancel{\beta_{EJ} EJ_n})$$
$$= \beta_F$$

Efecto marginal:

$$\frac{\partial \mathbb{E}[CA_n|F_n = 0]}{\partial EJ_n} = \frac{\partial}{\partial EJ_n}(\beta_0 + \beta_{EJ} EJ_n) = \beta_{EJ}$$

$$\frac{\partial \mathbb{E}[CA_n|F_n = 1]}{\partial EJ_n} = \frac{\partial}{\partial EJ_n}(\beta_0 + \beta_F + \beta_{EJ} EJ_n) = \beta_{EJ}$$

c. Para los dos ejemplos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   ii. Calcule la diferencia esperada en condición aeróbica entre un fumador y un no fumador que hacen la misma cantidad de ejercicio. Obtenga el efecto marginal del ejercicio aeróbico sobre la capacidad aeróbica tanto para fumadores como no fumadores. ¿Son iguales?

Diferencia esperada:

$$\begin{aligned}\mathbb{E}[CA_n|F_n=1]-\mathbb{E}[CA_n|F_n=1] &= (\cancel{\beta_0}+\beta_F+\cancel{\beta_{EJ}EJ_n})-(\cancel{\beta_0}+\cancel{\beta_{EJ}EJ_n})\\ &= \beta_F\end{aligned}$$

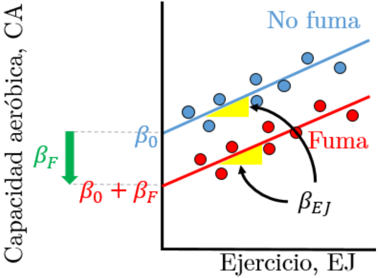
Efecto marginal:

$$\frac{\partial\mathbb{E}[CA_n|F_n=0]}{\partial EJ_n}=\frac{\partial}{\partial EJ_n}(\beta_0+\beta_{EJ}EJ_n)=\beta_{EJ}$$

$$\frac{\partial\mathbb{E}[CA_n|F_n=1]}{\partial EJ_n}=\frac{\partial}{\partial EJ_n}(\beta_0+\beta_F+\beta_{EJ}EJ_n)=\beta_{EJ}$$

Sí, son iguales.

c. Para los dos ejemplos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   iii. En cuanto a los parámetros poblacionales de su modelo, especule sus signos (+/-), interprete su significado y muéstrelos gráficamente.



$\beta_0, \beta_{EJ} > 0$, $\beta_F < 0$. β_0 es el intercepto de los no fumadores (CA cuando $EJ = 0$). β_F es la dif. prom. de CA entre fumadores y no fumadores. β_{EJ} es el retorno marginal del ejercicio sobre CA.

c. Para los dos ejemplos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   i. Calcule $\mathbb{E}[CA_n|F_n = 0]$, $\mathbb{E}[CA_n|F_n = 1]$.

Ejemplo 2: $CA_n = \beta_{NF} NF_n + \beta_F F_n + \beta_{EJ} EJ_n + U_n$

$$\mathbb{E}[CA_n|F_n = 0] = \beta_{NF} + \beta_{EJ} EJ_n$$

$$\mathbb{E}[CA_n|F_n = 1] = \beta_F + \beta_{EJ} EJ_n$$

c. Para los dos ejemplos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):

   ii. Calcule la dif. esperada en condición aeróbica entre un fumador y un no fumador que hacen la misma cantidad de ejercicio. Obtenga el efecto marginal del ej. aeróbico sobre la cap. aeróbica para fumadores y no fumadores. ¿Son iguales?

c. Para los dos ejemplos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
    ii. Calcule la dif. esperada en condición aeróbica entre un fumador y un no fumador que hacen la misma cantidad de ejercicio. Obtenga el efecto marginal del ej. aeróbico sobre la cap. aeróbica para fumadores y no fumadores. ¿Son iguales?

Diferencia esperada:

$$\begin{aligned}\mathbb{E}[CA_n|F_n = 1] - \mathbb{E}[CA_n|F_n = 1] &= (\beta_F + \cancel{\beta_{EJ} EJ_n}) - (\beta_{NF} + \cancel{\beta_{EJ} EJ_n}) \\ &= \beta_F - \beta_{NF}\end{aligned}$$

c. Para los dos ejemplos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):

   ii. Calcule la dif. esperada en condición aeróbica entre un fumador y un no fumador que hacen la misma cantidad de ejercicio. Obtenga el efecto marginal del ej. aeróbico sobre la cap. aeróbica para fumadores y no fumadores. ¿Son iguales?

Diferencia esperada:

$$\begin{aligned}\mathbb{E}[CA_n|F_n=1]-\mathbb{E}[CA_n|F_n=1] &= (\beta_F+\cancel{\beta_{EJ}EJ_n})-(\beta_{NF}+\cancel{\beta_{EJ}EJ_n}) \\ &= \beta_F-\beta_{NF}\end{aligned}$$

Efecto marginal:

$$\frac{\partial\mathbb{E}[CA_n|F_n=0]}{\partial EJ_n}=\frac{\partial}{\partial EJ_n}(\beta_{NF}+\beta_{EJ}EJ_n)=\beta_{EJ}$$

P4. Interceptos distintos con dummies

c. Para los dos ejemplos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):

   ii. Calcule la dif. esperada en condición aeróbica entre un fumador y un no fumador que hacen la misma cantidad de ejercicio. Obtenga el efecto marginal del ej. aeróbico sobre la cap. aeróbica para fumadores y no fumadores. ¿Son iguales?

Diferencia esperada:

$$\begin{aligned}\mathbb{E}[CA_n|F_n=1]-\mathbb{E}[CA_n|F_n=1] &= (\beta_F+\cancel{\beta_{EJ}EJ_n})-(\beta_{NF}+\cancel{\beta_{EJ}EJ_n})\\ &= \beta_F-\beta_{NF}\end{aligned}$$

Efecto marginal:

$$\frac{\partial\mathbb{E}[CA_n|F_n=0]}{\partial EJ_n}=\frac{\partial}{\partial EJ_n}(\beta_{NF}+\beta_{EJ}EJ_n)=\beta_{EJ}$$

$$\frac{\partial\mathbb{E}[CA_n|F_n=1]}{\partial EJ_n}=\frac{\partial}{\partial EJ_n}(\beta_F+\beta_{EJ}EJ_n)=\beta_{EJ}$$

c. Para los dos ejemplos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):
   ii. Calcule la dif. esperada en condición aeróbica entre un fumador y un no fumador que hacen la misma cantidad de ejercicio. Obtenga el efecto marginal del ej. aeróbico sobre la cap. aeróbica para fumadores y no fumadores. ¿Son iguales?

Diferencia esperada:

$$\begin{aligned}\mathbb{E}[CA_n|F_n = 1] - \mathbb{E}[CA_n|F_n = 1] &= (\beta_F + \cancel{\beta_{EJ}EJ_n}) - (\beta_{NF} + \cancel{\beta_{EJ}EJ_n}) \\ &= \beta_F - \beta_{NF}\end{aligned}$$

Efecto marginal:

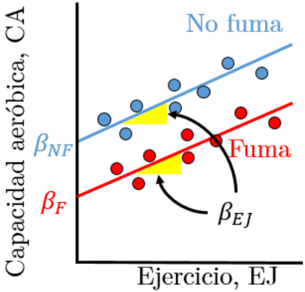
$$\frac{\partial\mathbb{E}[CA_n|F_n = 0]}{\partial EJ_n} = \frac{\partial}{\partial EJ_n}(\beta_{NF} + \beta_{EJ}EJ_n) = \beta_{EJ}$$

$$\frac{\partial\mathbb{E}[CA_n|F_n = 1]}{\partial EJ_n} = \frac{\partial}{\partial EJ_n}(\beta_F + \beta_{EJ}EJ_n) = \beta_{EJ}$$

Sí, son iguales.

c. Para los dos ejemplos y basándose en el gráfico que usted mismo(a) bosquejó, haga lo siguiente (siempre que se pueda):

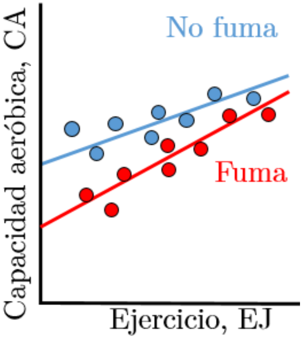
iii. En cuanto a los parámetros poblacionales de su modelo, especule sus signos (+/-), interprete su significado y muéstrelos gráficamente.



$\beta_{NF} > \beta_F > 0$, $\beta_{EJ} < 0$. β_F y β_{NF} son los interceptos (CA cuando $EJ = 0$) para fumadores y no fumadores. β_{EJ} es el retorno marginal del ej. sobre CA.

P5. Interceptos distintos con dummies y pendientes distintas con interacciones

a. Hacer un bosquejo de cómo se distribuye la muestra en cuanto a condición aeróbica y horas de ejercicio aeróbico semanal, separando a fumadores y no fumadores. En la misma figura grafique un modelo que cumpla con las características descritas por su amiga.



P5. Interceptos distintos con dummies y pendientes distintas con interacciones

El resto de la P5 queda propuesta para que ejerciten :)

P6. Interacción de variables continuas

"Leí en un *paper* que el efecto marginal de hacer ejercicio sobre la condición aeróbica disminuye con la edad".

P6. Interacción de variables continuas

"Leí en un *paper* que el efecto marginal de hacer ejercicio sobre la condición aeróbica disminuye con la edad". Proponga un modelo de reg. lineal con int. de variables continuas que permita estimar dicho efecto.

P6. Interacción de variables continuas

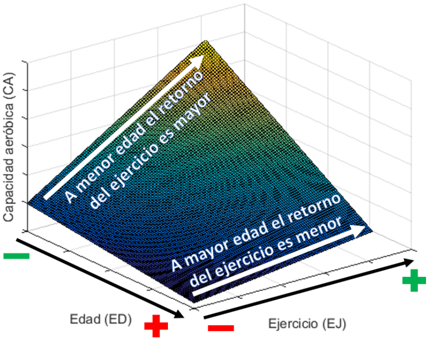
"Leí en un *paper* que el efecto marginal de hacer ejercicio sobre la condición aeróbica disminuye con la edad". Proponga un modelo de reg. lineal con int. de variables continuas que permita estimar dicho efecto. Haga un bosquejo de cómo se comportaría en valor esperado la cap. aeróbica en función de la edad y la cantidad de ejercicio aeróbico semanal suponiendo que lo que dice su amiga es cierto.

$$CA_n = \beta_0 + \beta_1 ED_n EJ_n + \beta_2 EJ_n + U_n$$

P6. Interacción de variables continuas

"Leí en un *paper* que el efecto marginal de hacer ejercicio sobre la condición aeróbica disminuye con la edad". Proponga un modelo de reg. lineal con int. de variables continuas que permita estimar dicho efecto. Haga un bosquejo de cómo se comportaría en valor esperado la cap. aeróbica en función de la edad y la cantidad de ejercicio aeróbico semanal suponiendo que lo que dice su amiga es cierto.

$$CA_n = \beta_0 + \beta_1 ED_n EJ_n + \beta_2 EJ_n + U_n \Rightarrow \frac{\partial \mathbb{E}[CA_n]}{\partial EJ_n} = \beta_1 ED_n + \beta_2$$



¿Qué estudiamos hoy?

- Uso de variables dummies sin interacciones
    - Para estimar promedios y diferencias promedio.
    - Para estimar interceptos distintos.
- Interacciones entre variables dummies para crear sub-categorías
- Interacciones entre variables dummies y variables reales para obtener efectos marginales diferidos entre categorías.
- Interacciones entre variables reales para obtener efectos marginales cambiantes para un regresor a medida que otro regresor varía.