a) Suponga que se ajusta un modelo de regresión con una variable categórica, sin interacción, ¿dicho modelo genera rectas secantes?

Supongamos una variable categórica con c categorías

$y_i = \beta_0 + \beta_1 X_i + \beta_2 I_1 + \cdots + \beta_c I_{c-1}$

Si pertenece a la categoría 1

$y_i = (\beta_0 + \beta_2) + \underline{\beta_1} X_i$

Si pertenece a la categoría J

$y_i = (\beta_0 + \beta_{J+1}) + \underline{\beta_1} X_i$

b) En un modelo de regresión lineal simple ajustado solo con factores, las rectas generadas son horizontales.

$y_i = \beta_0 + \beta_1 I_1 + \cdots + \beta_{c-1} I_{c-1}$

Sí, son horizontales

c) El parámetro β_j es la media de Y en la categoría j en el modelo de regresión $Y = \beta_0 + \sum_{k=1}^{c-1} \beta_k I_k$, en caso de que no, ¿cuál es la media?

No es la media porque está el intercepto, pues el modelo donde β_J es la media de la categoría J es

$$Y_i = \beta_1 I_1 + \beta_2 I_2 + \cdots + \beta_c I_c$$

La media es $\beta_J + \beta_0$

$$Y = \beta_0 + \beta_1 I_1 + \beta_2 I_2 + \cdots + \beta_{c-2} I_{c-2} + \beta_{c-1} I_{c-1}$$

$$\beta_0 = \beta_0 \times 1 = \beta_0 (I_1 + I_2 + \cdots + I_c)$$

$$Y = \beta_0 (I_1 + \cdots + I_c) + \beta_1 I_1 + \beta_2 I_2 + \cdots + \beta_{c-2} I_{c-2} + \beta_{c-1} I_{c-1}$$

$$= (\beta_0 + \beta_1) I_1 + (\beta_0 + \beta_2) I_2 + \cdots + (\beta_0 + \beta_{c-1}) I_{c-1} + \beta_0 I_c$$

La media de la categoría J es: $\begin{cases} \beta_0 & \text{si } J = c \\ \beta_0 + \beta_J & \text{si } J \neq c \end{cases}$

¿Por qué $I_1 + I_2 + \cdots + I_c$ siempre es igual a 1?

Tomemos la categoría J, con $1 \leq J \leq c$

$$I_n = \begin{cases} 1 & \text{si } n = J \\ 0 & \text{si } n \neq J \end{cases} \Rightarrow \sum_{n=1}^{c} I_n = \cancel{\sum_{n=1}^{J-1} I_n}^{\,0} + \underbrace{I_J}_{1} + \cancel{\sum_{n=J+1}^{c} I_n}^{\,0} = 1$$

d) La interacción entre variables numéricas y categóricas hace variar la tasa de cambio de la respuesta en cada categoría de la variable categórica.

La interacción hace que la tasa de cambio de la respuesta respecto a alguna covariable varíe.

$$y = \beta_0 + \beta_1 X_i + \beta_2 I_1 + \beta_3 I_1 X_i$$

Si $I_1 = 0 \Rightarrow y = \beta_0 + \beta_1 X_i$

Si $I_1 = 1 \Rightarrow y = (\beta_0 + \beta_2) + (\beta_1 + \beta_3) X_i$