Econometría Avanzada

Regresión Múltiple

Ana MAria Díaz (R de Edward Rubin)

08 August 2023

Prologue

Regresión Múltiple

Regresión Múltiple

De la regresión simple

$$y_i = \beta_0 + \beta_1 x_i + u_i$$

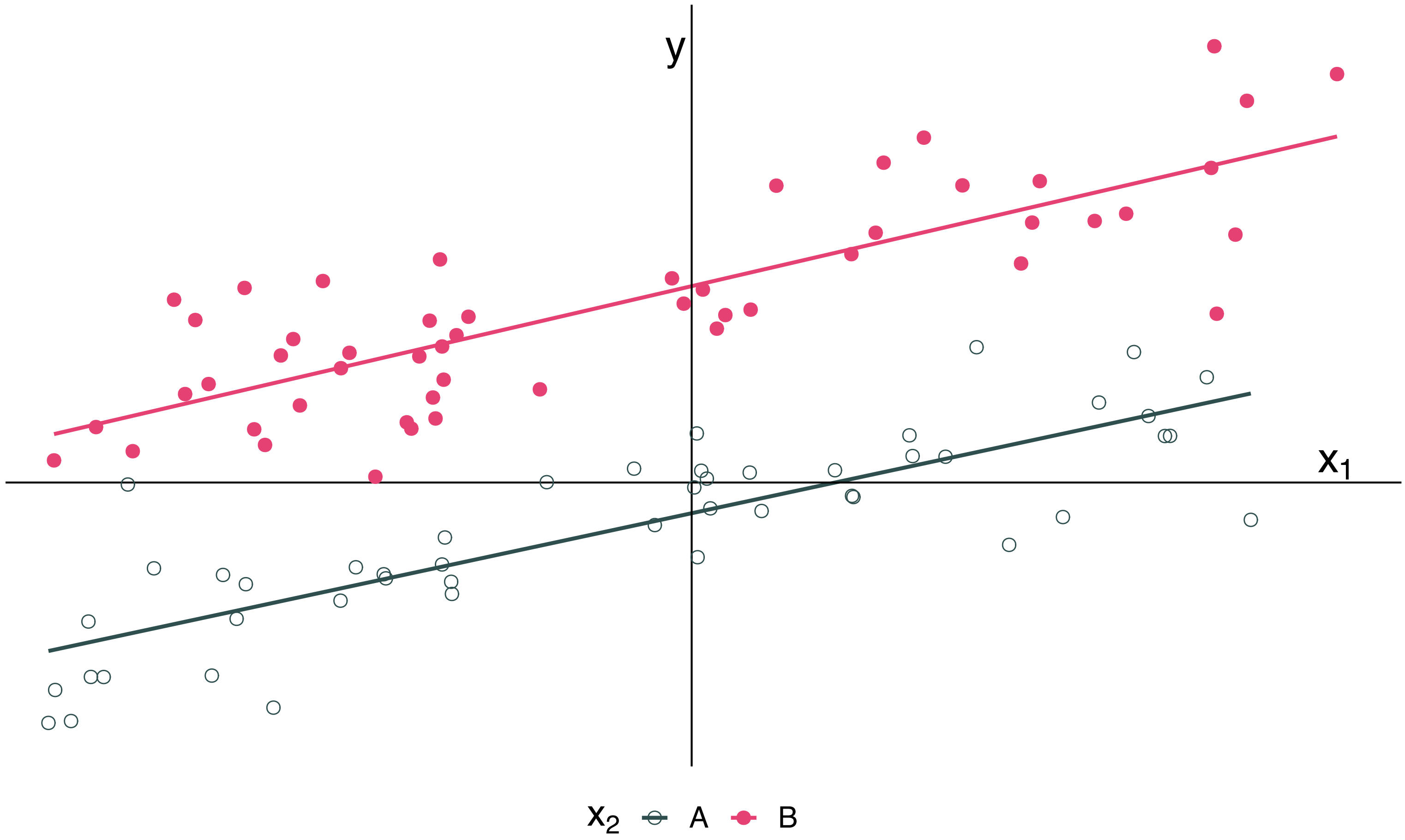
a la regresión múltiple

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_k x_{ki} + u_i$$

Por qué? Explicar mejor la variación de y, mejorar las predicciones, evitar sesgo por variables omitidas, …

Regresión Múltiple

Otra forma de pensarlo:



Regresión Múltiple

Para la regresión lineal simple $y_i = \beta_0 + \beta_1 x_i + u_i$

$$
\begin{aligned}
\hat{\beta}_1 &= \\
&= \frac{\sum_i \left(x_i - \overline{x}\right)\left(y_i - \overline{y}\right)}{\sum_i \left(x_i - \overline{x}\right)} \\
&= \frac{\sum_i \left(x_i - \overline{x}\right)\left(y_i - \overline{y}\right)/(n-1)}{\sum_i \left(x_i - \overline{x}\right)/(n-1)} \\
&= \frac{\hat{\text{Cov}}(x,\, y)}{\hat{\text{Var}}(x)}
\end{aligned}
$$

Regresión Múltiple

Estimador de la regresión lineal simple:

$$\hat{\beta}_1 = \frac{\hat{\text{Cov}}(x, y)}{\hat{\text{Var}}(x)}$$

Al movernos al estimador de la regresión múltiple, el estimador cambia un poco:

$$\hat{\beta}_1 = \frac{\hat{\text{Cov}}(\tilde{x}_1, y)}{\hat{\text{Var}}(\tilde{x}_1)}$$

donde $\tilde{x}_1$ es la variable x_1 limpia *residualized*, que mide la variación restante en x luego de controlar por todas las variablez x.

Regresión Múltiple

$$y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + u_i$$

Nuestra variable x_1 *limpia* (que llamamos $\tilde{x}_1$) proviene de hacer una regresión de x_1 contra el intercepto y las variables de control restantes, *i.e.*,

$$\hat{x}_{1i} = \hat{\gamma}_0 + \hat{\gamma}_2 \, x_{2i} + \hat{\gamma}_3 \, x_{3i}$$
$$\tilde{x}_{1i} = x_{1i} - \hat{x}_{1i}$$

Lo que nos permite entender mejor el estimador de MCO

$$\hat{\beta}_1 = \frac{\hat{\text{Cov}}(\tilde{x}_1, \, y)}{\hat{\text{Var}}(\tilde{x}_1)}$$

Regresión Múltiple