Método de Mínimos Cuadrados Ordinarios (MCO-OLS)

¿Cómo se determina la FRM?

- Teníamos que $Y_i = \hat{Y}_i + \hat{\varepsilon}_i$
  - $\hat{\varepsilon}_i = Y_i - \hat{Y}_i$
  - $\hat{\varepsilon}_i = Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i$

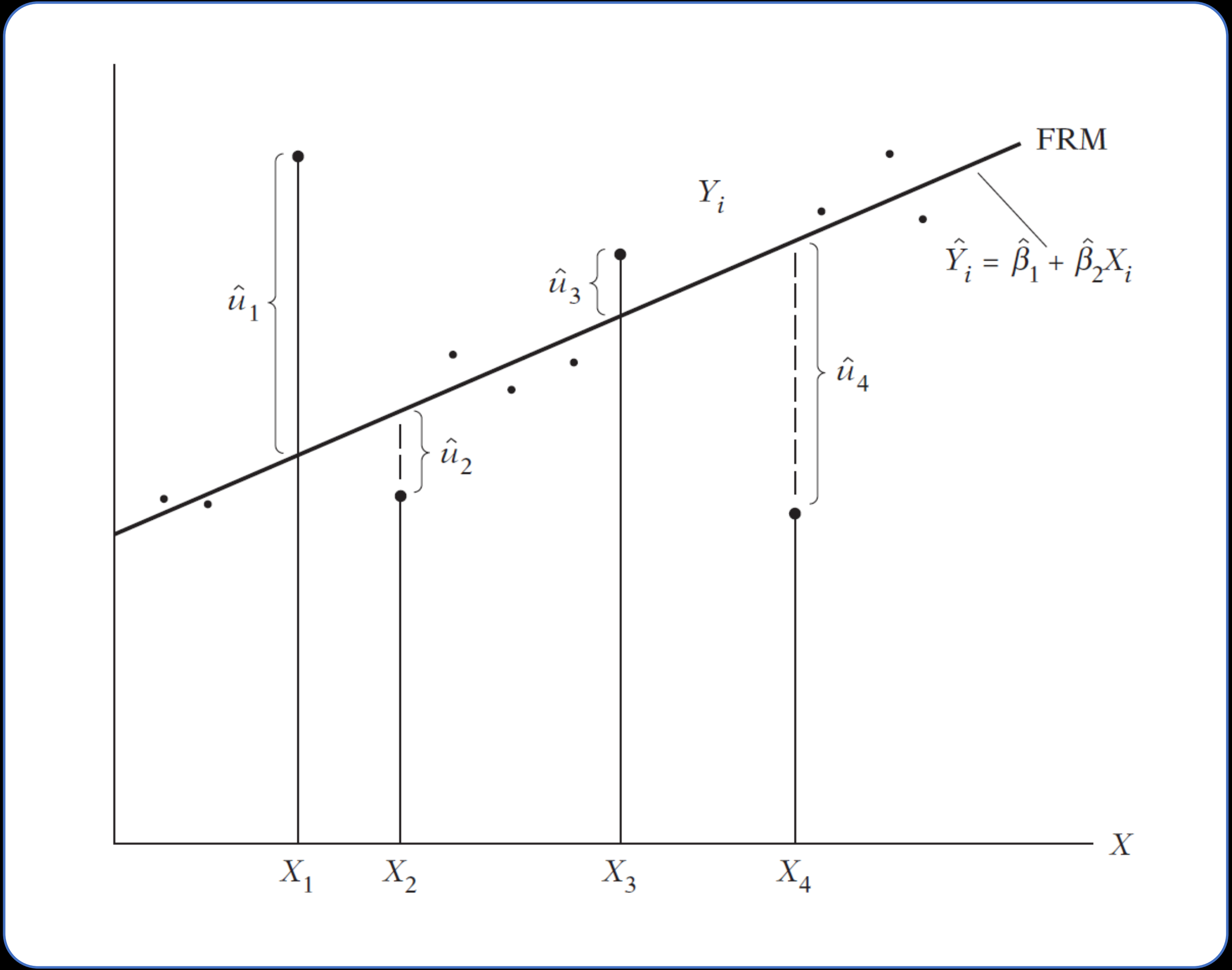
- Los residuos, $\hat{\varepsilon}_i$, son simplemente las diferencias entre los valores observados y los estimados de *Y.*

- Nos interesa que la FRM esté lo más cerca posible de las observaciones (reales).

- ¿Cómo logramos esto?

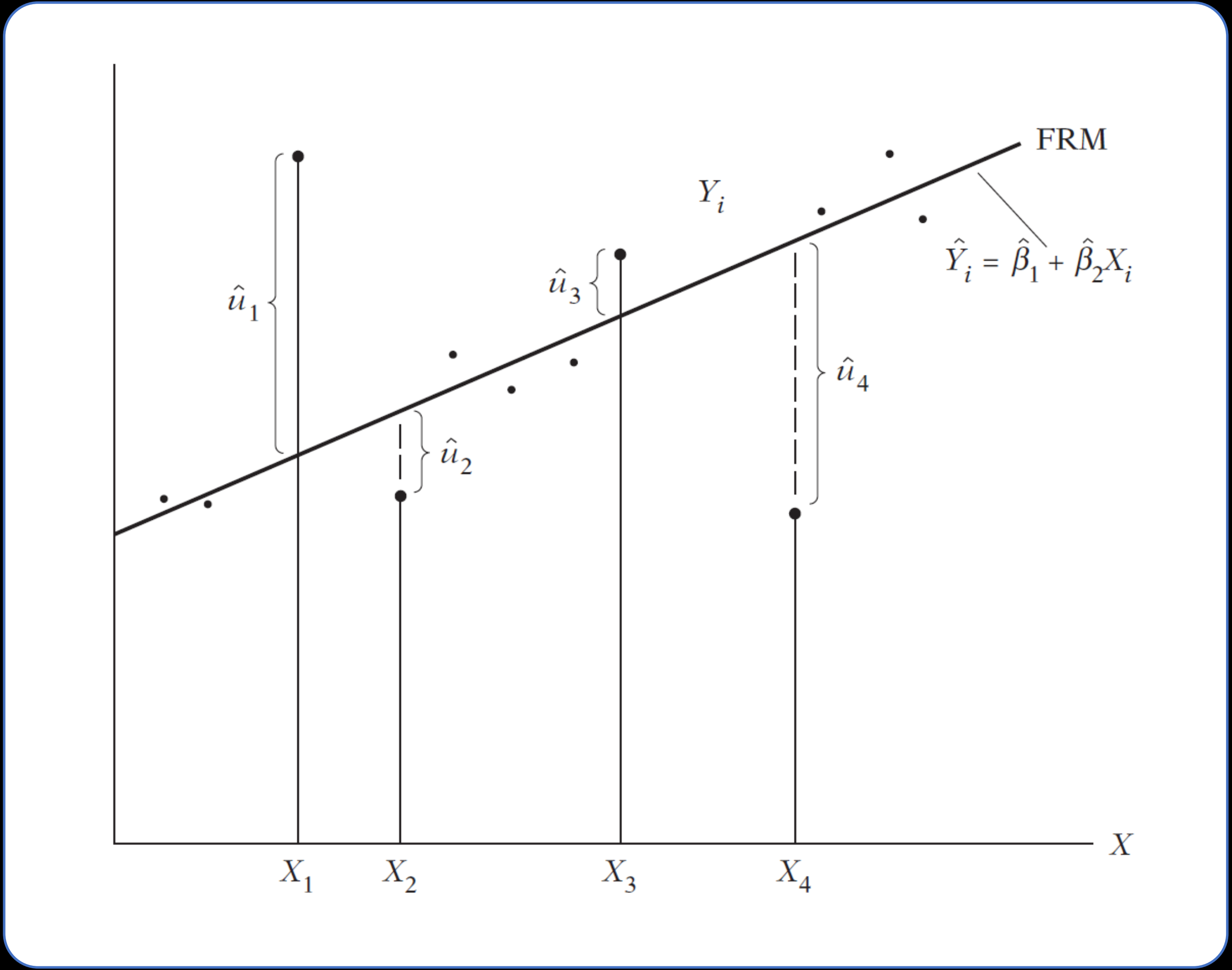
¿Cómo se determina la FRM?

- Podríamos escoger la FRM de manera que la suma de los residuos fuera la menor posible.
- $\sum \hat{\varepsilon}_i = \sum\left(Y_i - \widehat{Y}_i\right)$
- Este criterio no es muy adecuado: todos los residuales tienen el mismo peso.
  - $\hat{\varepsilon}_2$^$\hat{\varepsilon}_3$ están mucho más cerca de la FRM que $\hat{\varepsilon}_1$^$\hat{\varepsilon}_4$.
  - $(\hat{\varepsilon}_2 + \hat{\varepsilon}_3 + \hat{\varepsilon}_1 + \hat{\varepsilon}_4)$
- Es posible que la suma≈ 0, a pesar de que los $\hat{\varepsilon}_i$ estén muy dispersos.



- Entonces, para darle mayor peso a los residuales más alejados, podemos tomar los mínimos cuadrados:
- $\sum \hat{\varepsilon}_i{}^2 = \sum \left(Y_i - \hat{Y}_i\right)^2$
- $\sum \hat{\varepsilon}_i{}^2 = \sum \left(Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i\right)^2$

- A partir de esto, sabemos que
- $\sum \hat{\varepsilon}_i{}^2 = f(\hat{\beta}_1, \hat{\beta}_2)$
- (la suma de los residuales al cuadrado es una función de los estimadores).



¿Cuál experimento es mejor?

		Experimento 1 $\hat{Y}_{1i} = 1.572 + 1.357X_i$			Experimento 2 $\hat{Y}_{2i} = 3 + X_i$		
Y_i	$\boldsymbol{X_i}$	$\hat{Y}_{\mathbf{1}i}$	$\hat{\varepsilon}_{\mathbf{1}i}$	${\hat{\varepsilon}_{\mathbf{1}i}}^2$	$\hat{Y}_{\mathbf{2}i}$	$\hat{\varepsilon}_{\mathbf{2}i}$	${\hat{\varepsilon}_{\mathbf{2}i}}^2$
4	1						
5	4						
7	5						
12	6						
Suma=28	**16**						

¿Cuál experimento es mejor?

		Experimento 1 $\hat{Y}_{1i} = 1.572 + 1.357X_i$			Experimento 2 $\hat{Y}_{2i} = 3 + X_i$		
Y_i	$\boldsymbol{X_i}$	$\hat{Y}_{1i}$	$\hat{\varepsilon}_{1i}$	${\hat{\varepsilon}_{1i}}^2$	$\hat{Y}_{2i}$	$\hat{\varepsilon}_{2i}$	${\hat{\varepsilon}_{2i}}^2$
4	1	2.929	1.071	1.147	4	0	0
5	4	7.000	-2.000	4.000	7	-2	4
7	5	8.357	-1.357	1.841	8	-1	1
12	6	9.714	2.286	5.226	9	3	9
Suma=28	**16**		**0**	**12.214**		**0**	**14**

- Las $\hat{\beta}'s$ del experimento 1 son mejores porque $\sum {\hat{\varepsilon}_{1i}}^2 < \sum {\hat{\varepsilon}_{2i}}^2$

- El método de MCO-OLS elige las $\hat{\beta}_1$^$\hat{\beta}_2$ que minimizan $\sum \hat{\varepsilon}_i{}^2$ para una muestra determinada.

- Derivando parcialmente $\sum \hat{\varepsilon}_i{}^2 = \sum\left(Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i\right)^2$ con respecto a $\hat{\beta}_1$^$\hat{\beta}_2$ llegamos a las **ecuaciones normales:**

$$\sum Y_i = n\hat{\beta}_1 + \hat{\beta}_2 \sum X_i$$

$$\sum Y_i X_i = \hat{\beta}_1 \sum X_i + \hat{\beta}_2 \sum X_i{}^2$$

- con n: tamaño de muestra.

- Resolviendo las ecuaciones simultáneas, obtenemos:

$$\hat{\beta}_2 = \frac{\sum x_i y_i}{\sum x_i^2}$$

$$\hat{\beta}_1 = \bar{Y} - \hat{\beta}_2 \bar{X}$$

- Que son los **estimadores de mínimos cuadrados**. Donde:
  - $\bar{X}, \bar{Y}$ son las medias muestrales
  - $x_i, y_i = (X_i - \bar{X}), (Y_i - \bar{Y})$; respectivamente (desviaciones respecto a los valores medios).

Propiedades numéricas de los estimadores de MCO-OLS

- Son las propiedades que se mantienen como consecuencia del uso de MCO, sin considerar la forma como se generan los datos.

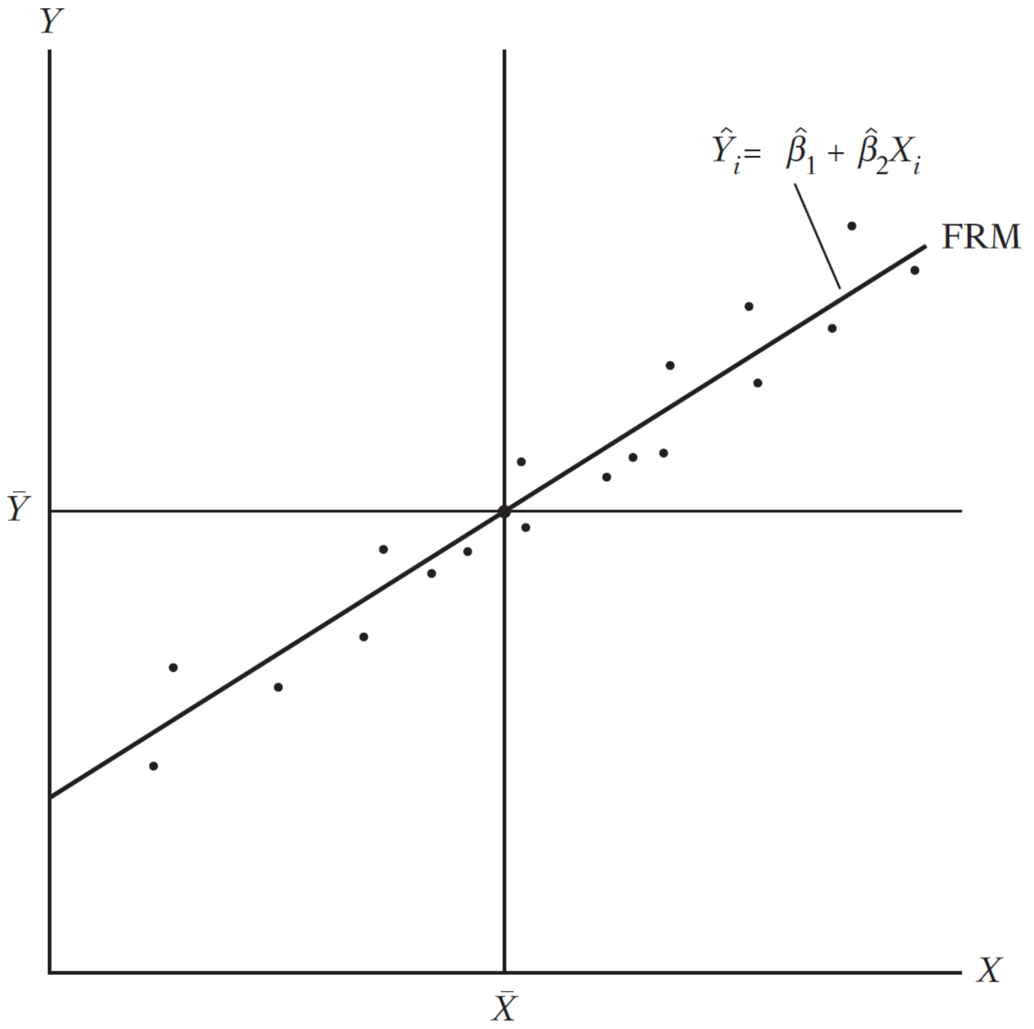
I. Los estimadores de MCO-OLS se expresan únicamente en términos de cantidades (*X,Y*) observables (muestras).

II. Son **estimadores puntuales**: cada estimador otorga un solo valor del parámetro poblacional.

III. Ya obtenidos los estimadores, la **línea de regresión muestral** se consigue sin problemas.

Propiedades de la línea de regresión muestral

1. Pasa a través de las medias muestrales de Y^X.
2. $\overline{\hat{Y}_i} = \bar{Y}$
3. $\overline{\hat{\varepsilon}_i} = 0$
4. Los residuos no están correlacionados con $\hat{Y}_i$; $\sum \hat{Y}_i \hat{\varepsilon}_i = 0$



Modelo clásico de regresión lineal

Fundamentos del método de mínimos cuadrados

- Aparte de obtener los estimadores $\hat{\beta}_1$ ^ $\hat{\beta}_2$, en el análisis de regresión se requiere inferir los verdaderos β_1 ^ β_2.

- Recordando, Y_i depende de X_i y de ε_i (por la FRP:$Y_i = \beta_1 + \beta_2 X_i + \varepsilon_i$).
  - Por lo tanto, necesitamos especificar la forma en cómo se generan X_i y ε_i, para poder hacer inferencias sobre Y_i y sobre β_1 ^ β_2.

  - Necesitamos analizar los supuestos y verificar su validez para poder tener seguridad en los resultados de nuestra regresión.
- El modelo estándar de regresión lineal (MCRL) *{modelo de Gauss o modelo clásico}* es la base de la teoría econométrica y plantea siete supuestos.

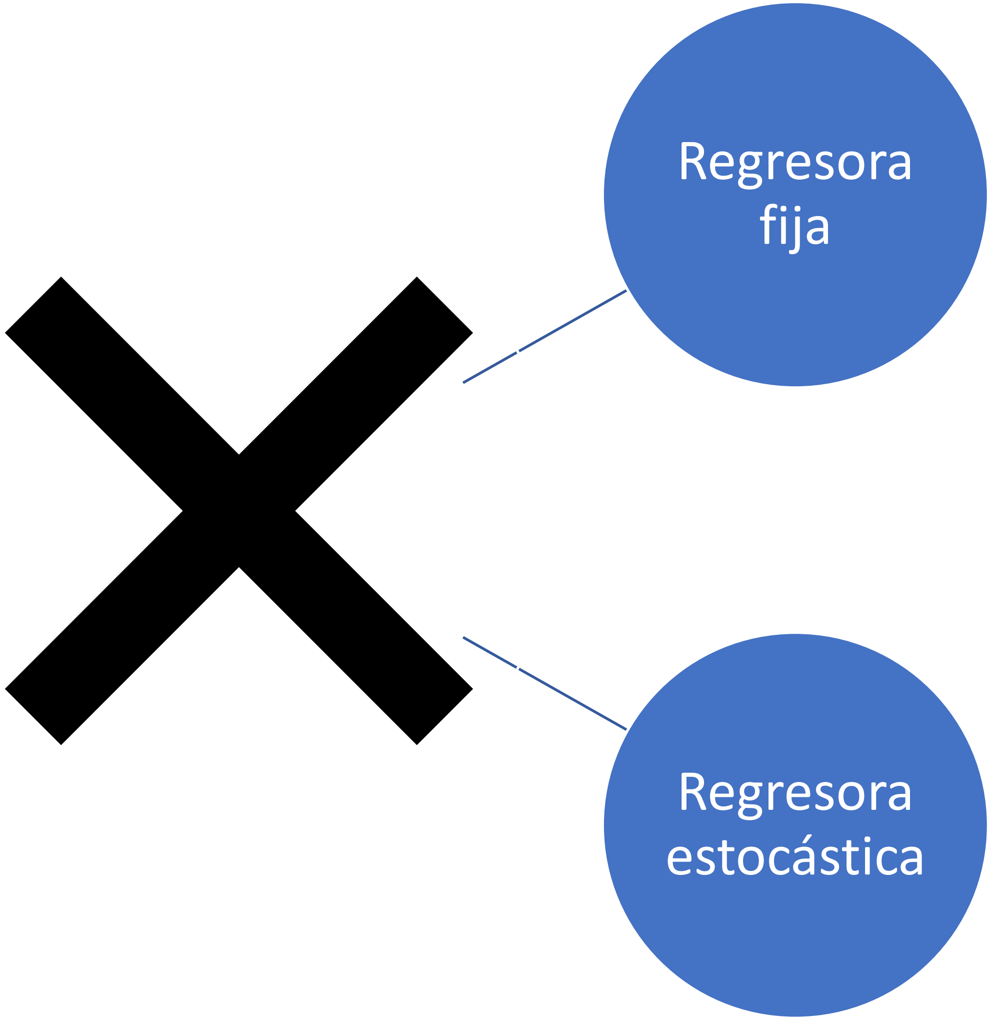
Supuestos del MCRL

Modelo clásico de regresión lineal

1. Modelo de regresión lineal

- El modelo es **lineal en los parámetros (en las $\boldsymbol{\beta's}$)**, aunque puede no ser lineal en las variables.

2. Valores fijos de *X*, o valores de *X* independientes del término de error



Regresora fija:

- Fijos en muestras repetidas.
- ¿Por qué asumirlas fijas?
  - Por simplicidad.
  - Puede ser conveniente.

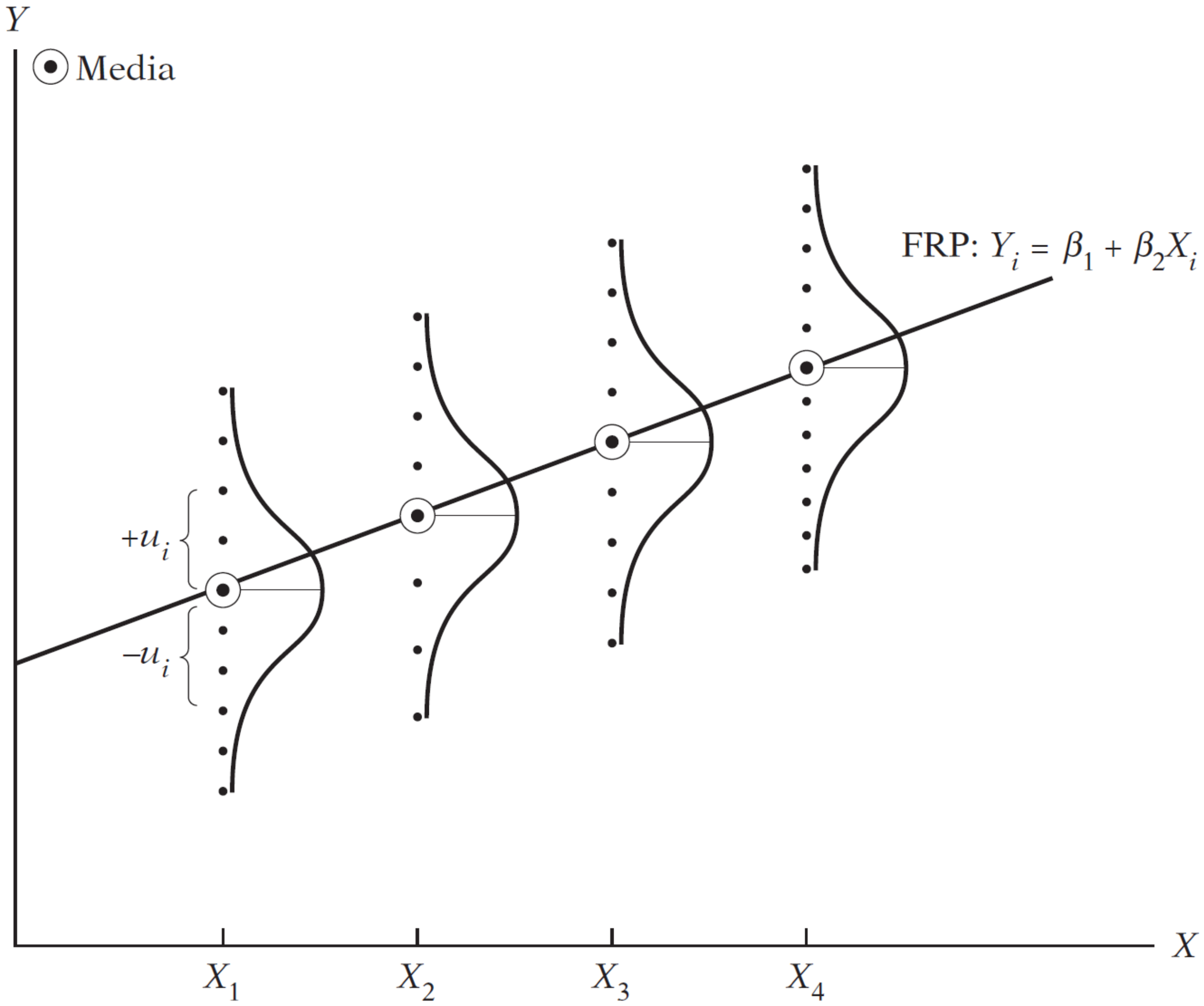
Regresora estocástica:

- Muestreados junto con Y.
- Supone $cov(X_i, \varepsilon_i) = 0$

3. El valor medio de ε_i es igual a cero

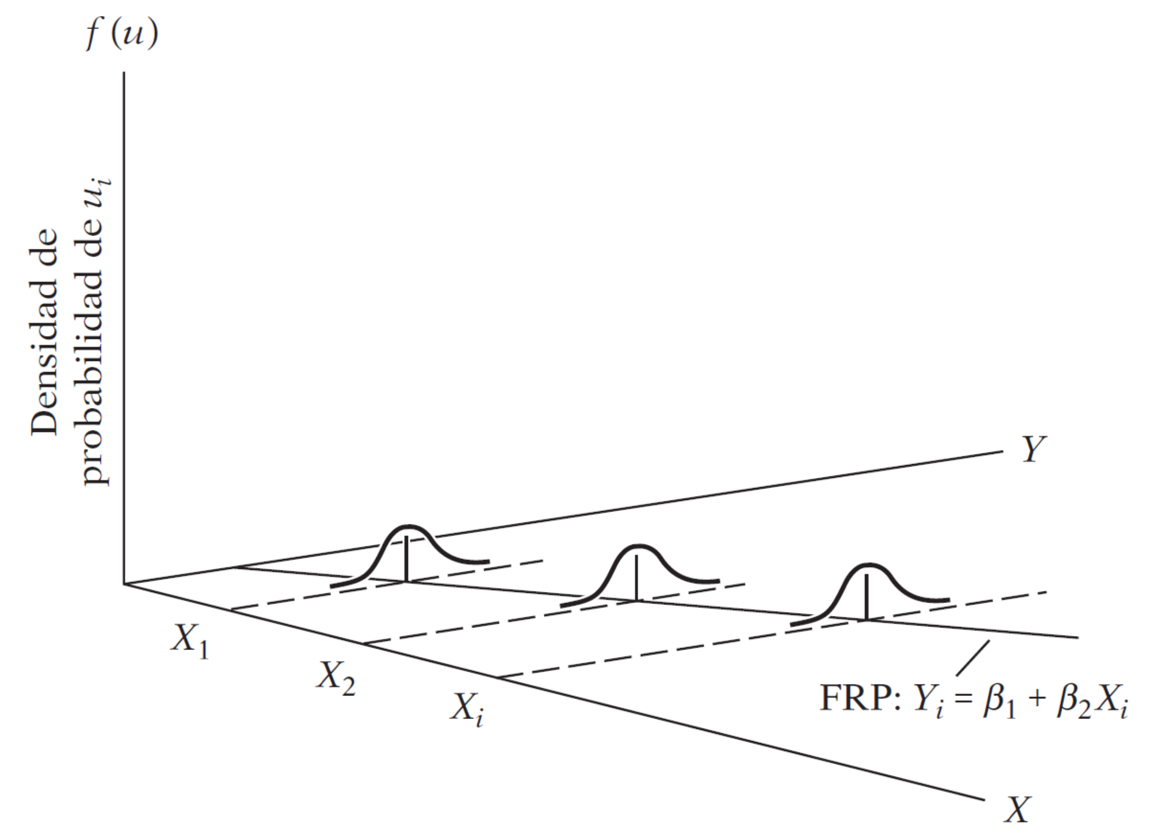
$$E(\varepsilon_i|X_i) = 0$$

- Los factores no incluidos explícitamente no afectan sistemáticamente el valor de la media de Y (los ε_i se cancelan unos con otros).
- Implica que ∄ **sesgo de especificación**.
    - Errores de especificación:
        - Variables importantes omitidas
        - Forma funcional incorrecta
- X_i no están correlacionadas con ε_i.



4. Homoscedasticidad (varianza constante de ε_i)

- La varianza de ε_i es la misma, sin importar el valor de X.
$$var(\varepsilon_i) = \sigma^2$$
- *Homo*: igual, *cedasticidad* (*skedanime*): dispersión= “Igual varianza”.
- La varianza no aumenta ni disminuye conforme cambia X.
- Las varianzas condicionales de Y_i también son homoscedásticas.

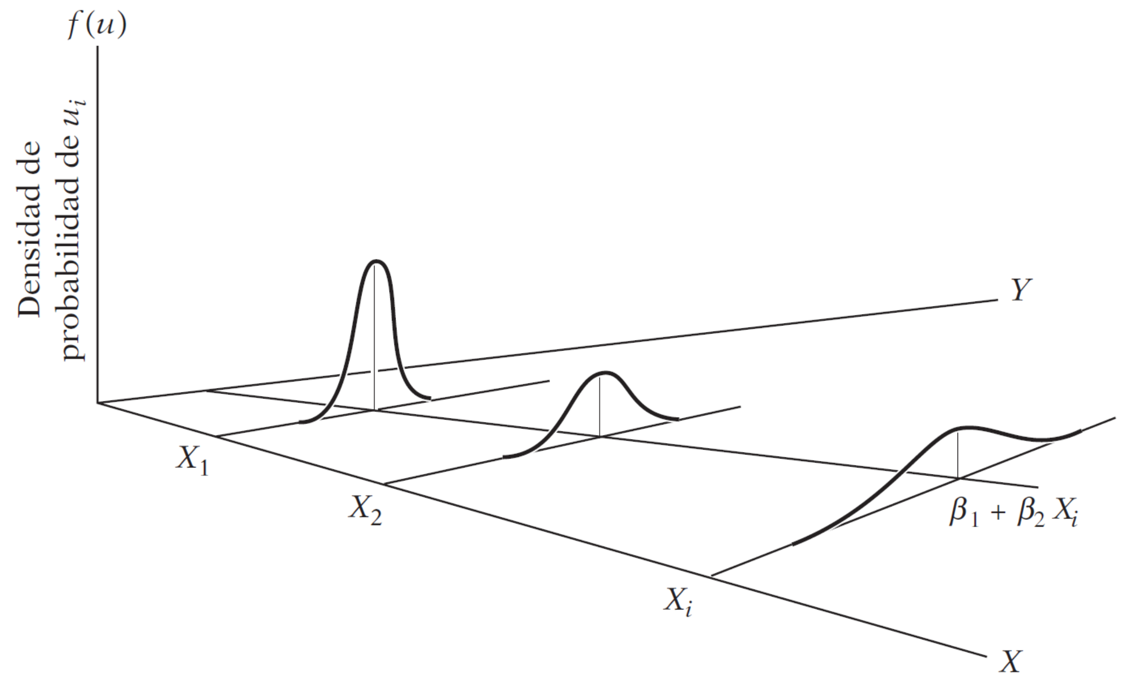


Heteroscedasticidad

$$var(\varepsilon_i|X_i) = \sigma_i^2$$

- Varianza desigual.

- La varianza ya no es constante.



5. No Autocorrelación en las perturbaciones

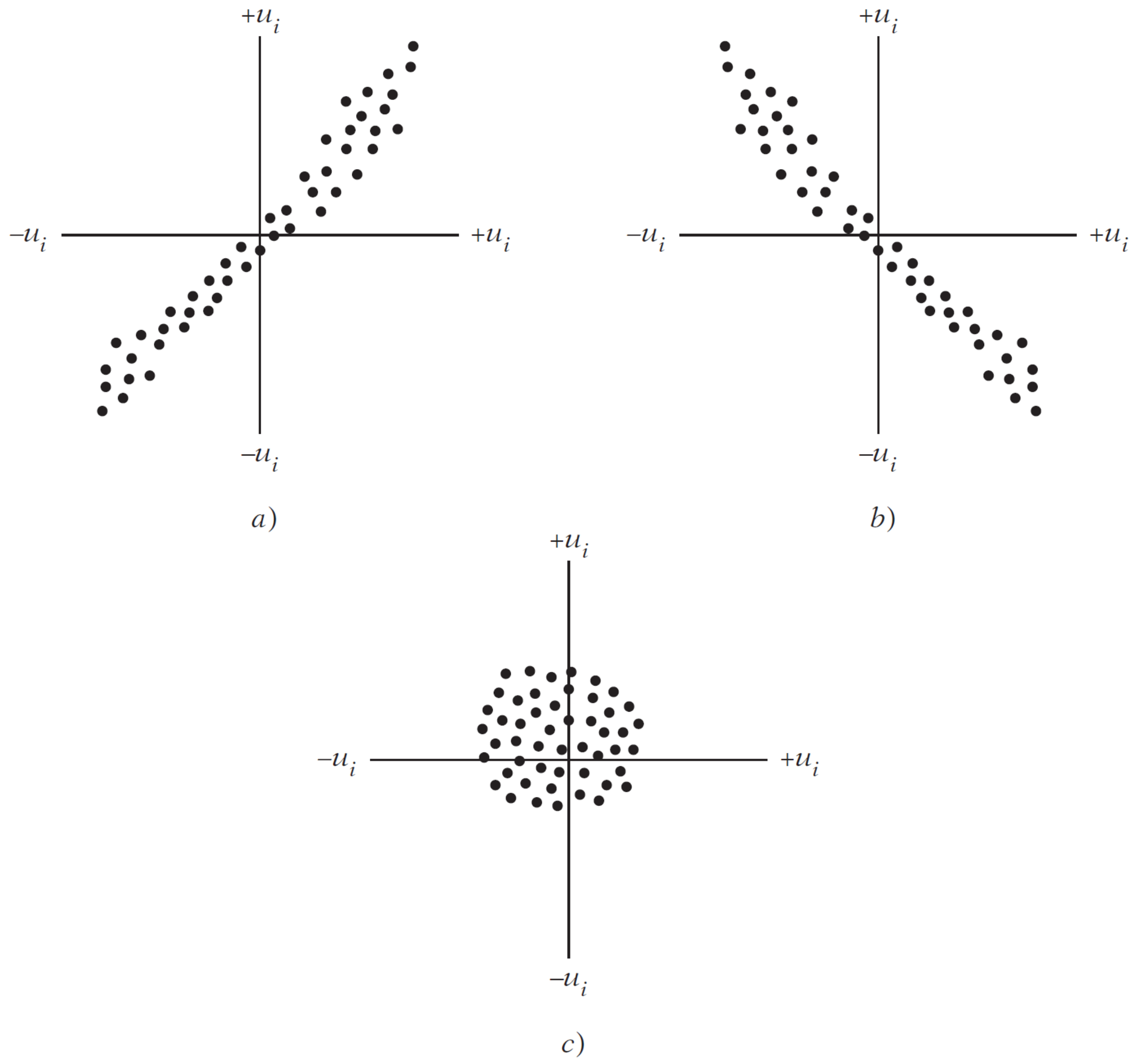
- También llamado supuesto de **no correlación serial** .

- Los términos de error no están correlacionados entre sí.

- $cov(u_i, u_j) = 0.$

$\varepsilon_i, \varepsilon_j, \forall i \neq j$

- Si estuvieran correlacionados, Y_t dependería de X_t y de U_{t-1},…



- En datos de corte transversal
  - Muchas veces es fácil mantener el supuesto de no autocorrelación (al estar los datos muestreados aleatoriamente).

- En series de tiempo
  - Es más complicado mantenerlo.
  - Las observaciones en la serie están correlacionadas (el PIB de un año está correlacionado con los anteriores, etc.)

6. El número de observaciones debe ser mayor que la cantidad de parámetros a estimar

$$n > \#\beta's$$

7. La naturaleza de las variables independientes (X)

- Los valores observados de X no deben ser iguales todos.
  - *Si se quiere medir la desigualdad en Guadalajara, no se debería realizar la encuesta sólo en Andares.*
  - X debe tener varianza positiva, de lo contrario no podremos explicar buena cantidad de los datos.

- No deben haber valores atípicos de X.
  - No se deben incluir valores aislados muy grandes con respecto al resto de las observaciones.
  - Pueden deberse a errores de medición o de mezclar muestras de diferentes poblaciones.

Precisión de los estimadores $\hat{\beta}_1$^$\hat{\beta}_2$

- La precisión (confiabilidad) de un valor estimado se puede medir por su error estándar (ee):

$$var(\hat{\beta}_2) = \frac{\sigma^2}{\sum x_i^2}$$

$$ee(\hat{\beta}_2) = \frac{\sigma}{\sqrt{\sum x_i^2}}$$

$$var(\hat{\beta}_1) = \frac{\sum X_i^2}{n \sum x_i^2} \sigma^2$$

$$ee(\hat{\beta}_1) = \sqrt{\frac{\sum X_i^2}{n \sum x_i^2}} \sigma$$

Donde:

- $var = varianza$
- $ee = error\ estándar$
- $\sigma^2 = varianza\ homoscedástica\ de\ \varepsilon_i$
- $cov(\hat{\beta}_1, \hat{\beta}_2) = -\bar{X} var(\hat{\beta}_2)$

- Para estimar σ^2, utilizamos:

$$\hat{\sigma}^2 = \frac{\sum \hat{\varepsilon}_i^2}{n-2}$$

Donde $\boldsymbol{n-2}$ es conocida como **número de grados de libertad** y $\sum \boldsymbol{\hat{\varepsilon}_i^2}$ como la **suma de cuadrados de los residuos (SCR)**.

$\hat{\sigma} = \sqrt{\frac{\sum \hat{\varepsilon}_i^2}{n-2}}$ es el **error estándar de estimación** (o de regresión).

Número de grados de libertad (gl)

- Es el número total de observaciones en la muestra, n, menos el número de restricciones independientes.

- En el modelo clásico con dos variables (Y ^ X), sólo se necesitan estimar $\hat{\beta}_1$ ^ $\hat{\beta}_2$ (tenemos dos restricciones). Tendríamos $n - 2 = gl$.

- Para el caso general con k variables, los grados de libertad serían $n - k$.

Teorema de Gauss-Markov

- Los estimadores de MCO poseen propiedades que los hacen óptimos. El estimador de MCO-OLS $\hat{\beta}_2$ es el **mejor estimador lineal insesgado (MELI-BLUE)** si se cumple:
    1. Es **lineal**
    2. Es **insesgado** ($E(\hat{\beta}_2) = \beta_2$).
    3. Tiene varianza mínima, o sea que es un **estimador eficiente**.

    "Dados los supuestos del modelo clásico de regresión lineal, los estimadores de MCO-OLS son MELI."