

HETEROCEDASTICIDAD

Se da cuando la varianza de los errores no es constante en las distintas observaciones

Homocedasticidad: $E(U_i^2) = \sigma^2$

Heterocedasticidad: $E(U_i^2) = \sigma_i^2$

Fuentes de Heterocedasticidad

- Características particulares de la regresión
- Factores atípicos
- Errores de especificación del modelo
- Asimetría en la distribución de las variables
- Incorrecta transformación de los datos o forma funcional

Estimación MCO con Heterocedasticidad

- Los estimadores siguen siendo INSESGADOS

$E(U_i) = 0$ Condición de Insesgamiento

$E(U_i) = 0 + \sum K_i E(U_i)$

Si $E(U_i) = 0$ entonces demostramos que el estimador es Insesgado. De esta manera observamos que para determinar la condición de insesgamiento no nos importa si la varianza de U_i es Homocedástica o heterocedástica. Por lo tanto en presencia de Heterocedasticidad los estimadores siguen siendo Insesgados.

- Los estimadores siguen siendo CONSISTENTES

La propiedad de Consistencia es de las muestras grandes y consiste en que la Varianza de β tiende a cero cuando n tiende a ∞ .

Bajo el supuesto de heterocedasticidad se sigue cumpliendo.

$E(U_i X_i) = 0$

- Los estimadores dejan de ser EFICIENTES ya que no son los de mínima varianza.
- Las varianzas y covarianzas de los estimadores de MCO son SESGADAS e INCONSISTENTES. Por este motivo los test de hipótesis ya no son válidos.

Detección de la Heterocedasticidad (métodos informales)

- Naturaleza del problema: Se basa en relaciones ya estudiadas que pueden ser similares a las que estamos analizando
- Casos particulares
- Método gráfico: consiste en hacer una regresión bajo el supuesto de que no existe heterocedasticidad y luego examinar los u con las y para ver si hay algún patrón de comportamiento.

Detección de la Heterocedasticidad (métodos formales)

- **TEST DE GOLDFELD-QUANT:** Se basa en la idea que si la varianza de los errores es igual a través de todas las observaciones, entonces la varianza para una parte de la muestra será la misma que la calculada con otra parte de la misma.

I ETAPA: Se identifica una variable Z relacionada con la varianza de los errores. Si suponemos que la relación es POSITIVA, ordenamos de manera creciente los datos de la muestra.

II ETAPA: Dividimos la muestra en 2 partes omitiendo los valores centrales (c)

III ETAPA: Estimamos las regresiones por separado.

IV ETAPA: Obtenemos SEC de cada una de las regresiones y calculamos las estimaciones de la varianza como $SEC1/n1-k$ y $SEC2/n2-k$

V ETAPA: Calculamos $F_{calc} = \frac{SEC1/n-k}{SEC2/n-k}$

$SEC2/n-k$

VI ETAPA: Comparamos F_{calc} con el valor F tabla con $(n1-K)$ GL numerador y $(n2-K)$ GL denominador. Si $F_{calc} > F_{tabla}$ rechazo H_0 de Homocedasticidad.

Limitación: El éxito depende de (c) y que se seleccione correctamente X.

- **TEST DE BREUSH-PAGAN:** Supone que la varianza de los errores no es una constante sino que está relacionada con un N° de variables Z.

$$Y1 = \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p + U1$$

$$U1 = \beta_{p+1} Z_1 + \dots + \beta_{p+q} Z_q + \epsilon$$

Si $\beta_{p+1} = \dots = \beta_{p+q} = 0$ entonces la varianza de los errores es constante indicando Homocedasticidad. Construimos un testeo con H_0 afirmando dicha igualdad.

I ETAPA: Estimamos $Y1 = \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p + U1$ por MCO y luego calculamos los residuos del modelo.

II ETAPA: $U2i$ (estimado) es una estimación de la varianza de los errores.

Si $\beta_{p+1} = \dots = \beta_{p+q} = 0$ fuese válida podemos esperar que $U2i$ (estimado) esté relacionado con las variables Z lo que sugiere una regresión auxiliar.

$$U2i \text{ (estimado)} = \beta_{p+1} Z_1 + \dots + \beta_{p+q} Z_q + \epsilon$$

(estimado)

III ETAPA: Demostraron que para Muestras Grandes y bajo $H_0: \beta_{p+1} = \dots = \beta_{p+q} = 0$ (Se regresión) $2/2$ tiene distribución Chi-cuadrado con p grados de libertad.

Entonces si $F_{calc} = SEC/2 > F_{tabla} = p$,

Rechazamos H_0 y hay Heterocedasticidad.

Limitación: Es sensible a cualquier tipo de violación del supuesto de normalidad de los errores.

- **TESTE DE WHITE:** También es un test para muestras grandes y es parecido al de Breush-pagan, pero no necesita ningún supuesto previo acerca de las causas de la heterocedasticidad.

I ETAPA: Estimamos el modelo por MCO.

II ETAPA: Calculamos U_i (estimado)

III ETAPA: Estimamos un modelo de regresión utilizando U_i (estimado) como variable dependiente sobre las X originales, las X^2 y los productos cruzados.

IV ETAPA: Calculamos R^2 para la regresión y $n \cdot R^2$

V ETAPA: $H_0: \beta_1 = \beta_2 = \dots = 0$

H_1 : al menos una $\beta \neq 0$

Si $nR^2 > (k-1)$,

Rechazo H_0 y tengo Heterocedasticidad.

–Si al menos una variable explicativa es DUMMY al elevarla al 2 resulta la misma variable y hay problemas de Multicolinealidad Perfecta.

–Si son muchas X podemos eliminar los términos cruzados.

Soluciones a la Heterocedasticidad

- Mínimos Cuadrados Generalizados : Consiste en dividir cada término por σ_i .

Modelo transformado

$Y_i / \sigma_i = X_1 / \sigma_i + X_2 / \sigma_i + \dots$ Este modelo satisface los supuestos de MCO, pero se puede presentar el inconveniente de no conocer σ_i .

- Mínimos Cuadrados Ponderados: es una extensión del MCG.

Definimos $w_i = 1 / \sigma_i$. Y transformando el modelo nos queda

$$Y_i w_i = (X_1 w_i) + (X_2 w_i) + \dots + (U_i w_i)$$

En este modelo transformado cada observación de la variable está ponderada por w_i (inversamente proporcional a σ_i)

- Conocemos la estructura de la Heterocedasticidad

Suponemos $\text{Var}(U_i) = \sigma_i^2$ (se denomina Heterocedasticidad Multiplicativa)

$w_i = 1 / \sigma_i$ Nos queda el modelo transformado.

La $\text{Var}(U_i \text{ transformado}) = \sigma_i^2$ (porque se nos elimina σ_i^2), de esta manera nos queda un modelo Homocedastico.

- Estructura Heterocedasticidad desconocida (MCGE)

I ETAPA: Estimamos el modelo por MCO

II ETAPA: Calculamos los residuos estimados al cuadrado

III ETAPA: Estimamos la regresión particular de los residuos estimados al cuadrado sugerida por White.

IV ETAPA: Usamos las estimaciones y obtenemos la estimación de

$$= + Z^2 + \dots$$

V ETAPA: Hacemos una regresión U_{2i} (estimado) contra cada variable del paso 3

(estimado)

dividida por (estimado).

$$= + X_2 + X_3 + X_2^2 + X_2X_3 \dots \dots \dots \text{etc}$$

VI ETAPA: $W = 1 /$ (estimado por MV)

Puede ocurrir que se obtengan estimaciones NEGATIVAS para las variables. En este caso se eliminan los términos cruzados o se usa el Log de U_2 estimado como término dependiente.