

Maestría en Econometría - UTDT

Trabajo Final - Econometría

PRIMERA PARTE: Inferencia en el Modelo de Regresión Lineal.

El objetivo de este ejercicio es estudiar la inferencia asintótica en la práctica cuando no se asume normalidad de los errores.

Considerar el siguiente modelo:

$$y = \alpha + \beta x + u \quad (1)$$

donde x es un escalar, $\alpha = 3$, $\beta = 1$ y $x \sim \mathcal{N}(2, 1)$. Analizar cómo cambia la inferencia estadística si no se asume normalidad en el término de error. El benchmark de comparación será $u \sim \mathcal{N}(0, 12)$, es decir, la varianza de los errores es igual a 12.

Para hacer el ejercicio, se van a analizar los siguientes casos:

- i. La distribución de los errores es uniforme.*
- ii. La distribución de los errores es asimétrica, como una transformación lineal de la distribución Beta o con una transformación de una distribución Bernoulli.*
- iii. La distribución de los errores tiene “colas gruesas” como la distribución t-Student.*
- iv. La distribución benchmark.*
- v. La distribución benchmark con una observación que es un outlier.*

Todas las simulaciones se basan en 2000 replicaciones.

Ejercicio 1.

Si el modelo verdadero cumple que $u \sim_{iid} U[-6, 6]$, se pide:

(a) *Generar 2000 muestras de $n = 10$ observaciones. Para cada muestra, estimar por MCC los parámetros del modelo y contrastar la hipótesis nula $H_0: \beta = 1$ versus una hipótesis alternativa $H_1: \beta \neq 1$ con un test t y un nivel de significación del 5%. Luego, reportar el tamaño del test (la proporción de las 2000 muestras en las que se rechaza la hipótesis nula).*

(b) *Repetir el inciso anterior con $n = 20$.*

(c) *Repetir el inciso anterior con $n = 100$.*

(d) *Repetir el inciso anterior con $n = 200$.*

(e) *Repetir el inciso anterior con $n = 500$.*

(f) *Repetir el inciso anterior con $n = 1000$.*

(g) *Repetir el inciso anterior con $n = 5000$.*

(h) Repetir el inciso anterior con $n = 10000$.

(i) Resumir los pasos anteriores en una tabla para mostrar el tamaño del test en cada caso y reportar qué ocurre a medida que el tamaño de muestra va aumentando.

(j) A partir de los resultados obtenidos, decir a qué conclusiones se puede arribar respecto a la inferencia en el modelo en la práctica.

Ejercicio 2.

Repetir los incisos anteriores para $u \sim_{iid} \sqrt{\frac{2352}{5}} (\text{Beta}(2, 5) - \frac{2}{7})$.

Ejercicio 3.

Repetir los incisos anteriores para $u \sim_{iid} t_{2,1818}$.

Ejercicio 4.

Repetir los incisos anteriores para $u \sim_{iid} \sqrt{75} (\text{Be}(0,8) - 0,8)$.

Ejercicio 5.

Repetir los incisos anteriores para $u \sim_{iid} \mathcal{N}(0, 12)$.

Ejercicio 6.

Repetir los incisos anteriores para $u \sim_{iid} \mathcal{N}(0, 12)$ cuando la primera observación del error toma el valor $u_1 = 500$.

Ejercicio 7.

Comparar cómo difiere la velocidad de convergencia en cada caso y con respecto a tener errores normales.

En la tabla 1, para cada combinación de tamaño de muestra (10, 20, 100, 200, 500, 1000, 5000, 10000) y modelo (según las diferentes distribuciones propuestas de los errores), se presenta el tamaño del test con un nivel de significación del 5% (considerando la hipótesis nula $H_0: \beta = 1$) para la estimación por MCC del modelo (1), considerando 2000 replicaciones.

Por un lado, se puede observar que, a medida que el tamaño de muestra aumenta, el tamaño del test parece converger al nivel de significación del 5%. A partir de estos resultados obtenidos, respecto a la inferencia en el modelo en la práctica, se puede decir que es esperable que la proporción de veces que se rechaza la hipótesis nula cuando ésta es cierta tienda al nivel de significación, ya que se esperaría que la cantidad de veces que, en $R = 2000$ replicaciones independientes, se rechaza la hipótesis nula incorrectamente

tienda (o sea igual) al nivel de significación ($\alpha = 5\%$) por la cantidad de replicaciones independientes ($R = 2000$).

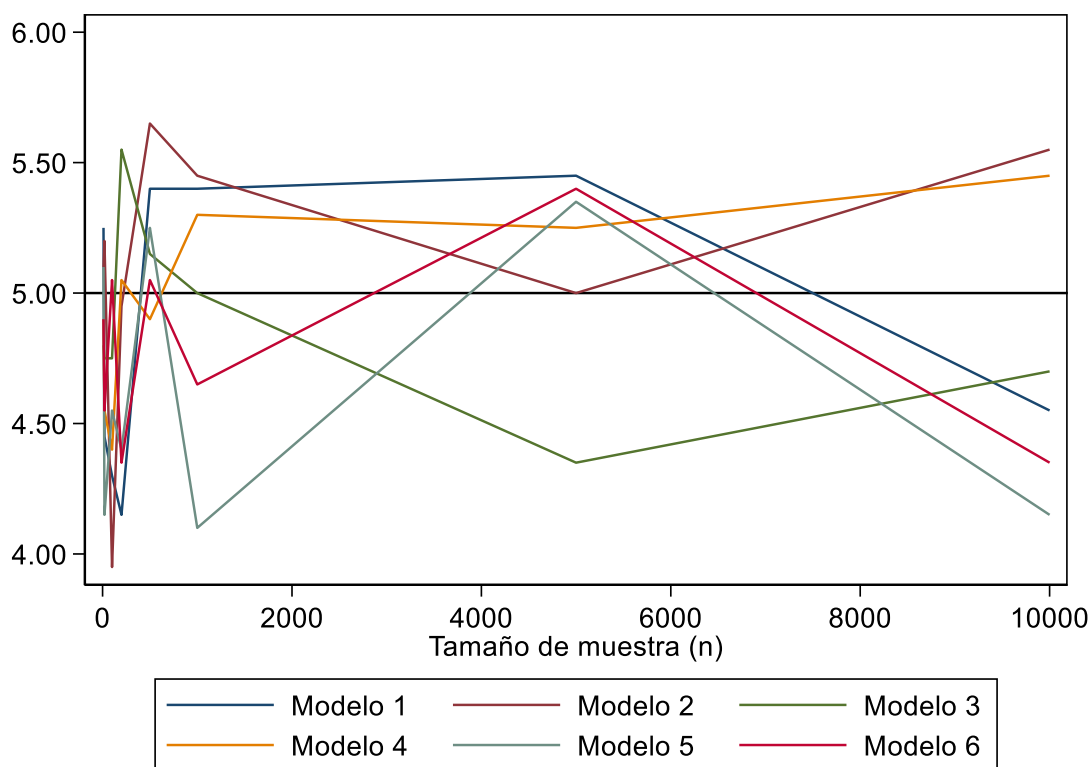
Por otro lado, en la figura 1, se presentan estos mismos resultados para ver más claramente la convergencia. Se puede observar que, comparando entre los distintos modelos, la velocidad de convergencia (a medida que n aumenta) no parece ser muy concluyente. Sin embargo, se puede decir que los modelos con errores normales (modelos 5 y 6) parecieran ser de los más lentos en converger, mientras que el modelo con errores con distribución t-Student y el modelo con errores con una transformación de una distribución Bernoulli (modelos 3 y 4) parecieran ser de los más rápidos.

Tabla 1. Simulaciones mediante estimación por MCC (con 2000 replicaciones).

Resultado	n= 10	n= 20	n= 100	n= 200	n= 500	n= 1000	n= 5000	n= 10000
Modelo 1	5,25	4,45	4,30	4,15	5,40	5,40	5,45	4,55
Modelo 2	5,10	5,20	3,95	4,95	5,65	5,45	5,00	5,55
Modelo 3	4,80	4,75	4,75	5,55	5,15	5,00	4,35	4,70
Modelo 4	4,90	4,55	4,40	5,05	4,90	5,30	5,25	5,45
Modelo 5	5,10	4,15	4,55	4,40	5,25	4,10	5,35	4,15
Modelo 6	4,90	4,55	5,05	4,35	5,05	4,65	5,40	4,35

Fuente: Elaboración propia.

Figura 1. Velocidad de convergencia en simulaciones mediante estimación por MCC (con 2000 replicaciones).



Fuente: Elaboración propia.

SEGUNDA PARTE: Propiedades de Muestra Finita de FGLS.

Considerar el siguiente modelo:

$$y_i = \beta_0 + \beta_1 x_i + u_i, i = 1, 2, \dots, 5N \quad (2)$$

con $\beta_0 = -3$, $\beta_1 = 0,8$, $u_j \sim \mathcal{N}(0, \Omega \otimes I_{N \times N})$, $\Omega = \begin{pmatrix} 4 & 0 & 0 & 0 & 0 \\ 0 & 9 & 0 & 0 & 0 \\ 0 & 0 & 16 & 0 & 0 \\ 0 & 0 & 0 & 25 & 0 \\ 0 & 0 & 0 & 0 & 36 \end{pmatrix}$ y $x_j \sim U[1, 50]$.

Ejercicio 1.

(a) Generar 5000 muestras de $5N = 5$ observaciones de corte transversal a partir del modelo (2). Para cada muestra, estimar por FGLS los parámetros del modelo y realizar un test de hipótesis para contrastar que $H_0: \beta_1 = 0,8$. Reportar tamaño del test al 1% y 5% y el poder del test cuando $\beta_1 = 0$ y $\beta_1 = 0,4$. Adicionalmente, reportar la media, mediana y desvío estándar de las estimaciones de β_0 y β_1 .

(b) **BONUS:** Utilizar la descomposición de Cholesky para encontrar una matriz P que cumpla que se puede escribir a $\Omega = PP'$, aplicar la transformación correspondiente a MCG al modelo (2) y estimar el modelo por MCC, comentar si cambia algún resultado.

Ejercicio 2.

Repetir el punto anterior con $5N = 10$.

Ejercicio 3.

Repetir el punto anterior con $5N = 30$.

Ejercicio 4.

Repetir el punto anterior con $5N = 100$.

Ejercicio 5.

Repetir el punto anterior con $5N = 200$.

Ejercicio 6.

Repetir el punto anterior con $5N = 500$.

En la tabla 2, para cada tamaño de muestra (5, 10, 30, 100, 200, 500), se presentan el tamaño del test con un nivel de significación del 1% y del 5% (considerando la hipótesis nula $H_0: \beta_1 = 0,8$), el poder del test cuando $\beta_1 = 0$ y $\beta_1 = 0,4$ (con un nivel de significación

del 5%), y la media, la mediana y el desvío estándar de las estimaciones de β_0 y β_1 para la estimación por FGLS (y por GLS sólo para el tamaño de muestra $5N=5$) del modelo (2), considerando 5000 replicaciones.

Por un lado, se puede observar que, a medida que el tamaño de muestra aumenta, la media y la mediana de ambos estimadores ($\hat{\beta}_0$ y $\hat{\beta}_1$) tienden a sus valores poblacionales, mientras que sus respectivos desvíos estándar tienden a cero. Por otro lado, se observa que, a medida que el tamaño de muestra aumenta, el tamaño de ambos test (al 1% y al 5%) y el poder de ambos test ($\beta_1=0$ y $\beta_1=0,4$) tiende a 100%, ya que, si bien el estimador tiende al parámetro poblacional (disminuyendo el valor observado del estadístico t), el error estándar del estimador es un valor lo suficientemente pequeño (aumentando, en mayor proporción, el valor observado del estadístico t), provocando que el p-valor sea un valor, extremadamente, pequeño y, por lo tanto, que se rechace la hipótesis nula en gran proporción de las replicaciones.

Tabla 2. Simulaciones mediante estimación por FGLS-GLS (con 5000 replicaciones).

Resultado	5N= 5 (FGLS)	5N= 5 (GLS)	5N= 10 (FGLS)	5N= 30 (FGLS)	5N= 100 (FGLS)	5N= 200 (FGLS)	5N= 500 (FGLS)
Tamaño del test (1%)	11,14	0,90	33,06	67,76	85,58	90,40	93,84
Tamaño del test (5%)	25,76	3,66	47,04	75,96	89,34	92,78	95,42
Poder del test ($\beta_1=0$)	45,12	15,66	82,64	99,48	100,00	100,00	100,00
Poder del test ($\beta_1=0,4$)	31,60	6,32	61,40	91,36	99,50	100,00	100,00
Media $\hat{\beta}_0$	-2,711	-2,688	-2,969	-3,110	-3,063	-3,052	-3,003
Mediana $\hat{\beta}_0$	-3,097	-2,842	-2,872	-3,106	-2,992	-3,045	-3,024
Desvío estándar $\hat{\beta}_0$	27,342	20,962	14,404	7,750	4,304	3,082	1,957
Media $\hat{\beta}_1$	0,785	0,783	0,797	0,802	0,801	0,801	0,801
Mediana $\hat{\beta}_1$	0,788	0,792	0,792	0,800	0,801	0,802	0,801
Desvío estándar $\hat{\beta}_1$	0,960	0,734	0,500	0,264	0,148	0,105	0,068

Fuente: Elaboración propia.

Ejercicio 7.

Describir, detalladamente, las propiedades de muestra finita de FGLS de acuerdo a lo observado en los puntos anteriores. En especial, explicar cómo cambia el tamaño y el poder de los tests a medida que aumenta el tamaño de muestra.

Las propiedades de muestra finita de FGLS, de acuerdo a lo observado en los puntos anteriores, son:

- **Sesgo**: El estimador FGLS puede estar sesgado si el modelo subyacente no se especifica correctamente o si la estructura de correlación verdadera en los datos se impone de manera incorrecta. Sin embargo, a medida que el tamaño de muestra aumenta, este sesgo tiende a disminuir.
- **Eficiencia relativa**: La eficiencia relativa del estimador FGLS en comparación con otros estimadores (como el estimador MCO) puede variar dependiendo de la estructura de correlación verdadera en los datos y del modelo. Sin embargo, en algunos casos, el estimador FGLS puede proporcionar estimaciones más precisas que el estimador MCO, especialmente cuando la estructura de correlación de los errores es ignorada por las estimaciones por MCO.
- **Varianza finita**: La varianza del estimador FGLS depende del tamaño de muestra y de la estructura de correlación verdadera en los datos. A diferencia de las propiedades asintóticas, en muestras finitas, la varianza del estimador FGLS puede no converger a la varianza asintótica y puede ser mayor o menor dependiendo de las características específicas de los datos y del modelo.
- **Robustez**: El estimador FGLS puede ser más robusto que otros estimadores en presencia de violaciones de los supuestos de homocedasticidad y de correlación serial en los datos. Esto significa que el estimador FGLS puede proporcionar estimaciones más precisas, incluso cuando los supuestos clásicos no se cumplan completamente.

TERCERA PARTE: Corrección de la Matriz de Varianzas y Covarianzas en Presencia de Heterocedasticidad, White (1980).

En este ejercicio, se pide evaluar el funcionamiento de la corrección sugerida por White para estimar la matriz de varianzas y covarianzas de los coeficientes estimados por mínimos cuadrados clásicos en muestras finitas. El objetivo final del ejercicio será aprender el cálculo correcto de la matriz de varianzas y covarianzas de los coeficientes estimados por mínimos cuadrados clásicos en presencia de heterocedasticidad. Para esto, considerar el siguiente modelo:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \sqrt{v_i} u_i, i = 1, 2, \dots, n \quad (3)$$

Para $n = 20$, x_1 se determina como una secuencia de 18 puntos espaciados uniformemente entre -1 y 1 con puntos extremos dados por -1 y 1, mientras que x_2 son cuantiles de una distribución normal estándar elegidos aleatoriamente. Las observaciones se repiten tres veces para obtener una muestra de $n = 60$, cinco veces para obtener una muestra de $n = 100$, diez veces para obtener una muestra de $n = 200$, veinte veces para obtener una muestra de $n = 400$ y treinta veces para obtener una muestra de $n = 600$. La generación de los datos de la variable dependiente se hace con $\beta_0 = \beta_1 = \beta_2 = 1$. Hay dos diseños:

Diseño 1: $u_i \sim \mathcal{N}(0, 1)$ y $v_i = e^{0,25x_{1i} + 0,25x_{2i}}$ (normalidad y heterocedasticidad).

Diseño 2: $u_i \sim t_5$ y $v_i = e^{0,25x_{1i} + 0,25x_{2i}}$ (no normalidad y heterocedasticidad).

Todas las simulaciones se basan en 5000 replicaciones. En el modelo de regresión lineal (3), la estimación de MCC es:

$$\hat{\beta} = \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = (X'X)^{-1}X'y,$$

donde X es la matriz de variables explicativas de dimensión $n \times 3$ e y es el vector de observaciones de la variable dependiente de dimensión $n \times 1$. Bajo heterocedasticidad, la matriz de varianzas y covarianzas de los estimadores de MCC es:

$$\text{Var}(\hat{\beta}) = (X'X)^{-1}X'\Omega X(X'X)^{-1}, \text{ con } \Omega = \text{diag}(\sigma_1^2, \dots, \sigma_n^2).$$

Una estimación consistente de esta matriz está dada por la denominada matriz de White (White, 1980):

$$\widehat{\text{Var}}(\hat{\beta}) = (X'X)^{-1}X'\hat{\Omega}X(X'X)^{-1},$$

con $\hat{\Omega} = \text{diag}(\hat{u}_1^2, \dots, \hat{u}_n^2)$ y \hat{u} es el vector de dimensión $n \times 1$ de residuos de la estimación por MCC. Se pide:

Ejercicio (a).

Para cada uno de los diseños 1 y 2, generar 5000 muestras de $n=20$ observaciones a partir del modelo (3). Para cada muestra, estimar por MCC los parámetros del modelo y reportar el sesgo relativo de la estimación de las varianzas de $\tilde{\beta}_0$, $\tilde{\beta}_1$ y $\tilde{\beta}_2$ y los sesgos relativos totales. El sesgo relativo (b_0 , b_1 , b_2) se define como el promedio, a lo largo de las 5000 simulaciones, de la varianza estimada menos la varianza verdadera dividido por la varianza verdadera. El sesgo relativo total es la suma del valor absoluto de (b_0 , b_1 , b_2). El sesgo relativo total es una medida del sesgo agregado de las tres varianzas.

Ejercicio (b).

Repetir el punto anterior con $n=60$.

Ejercicio (c).

Repetir el punto anterior con $n=100$.

Ejercicio (d).

Repetir el punto anterior con $n=200$.

Ejercicio (e).

Repetir el punto anterior con $n=400$.

Ejercicio (f).

Repetir el punto anterior con $n=600$.

Ejercicio (g).

Repetir los puntos (a) a (f), pero, ahora, construir la estimación de la matriz de White usando una matriz diagonal con los errores verdaderos elevados al cuadrado en lugar de utilizar la matriz diagonal con los residuos elevados al cuadrado.

En las tablas 3 y 4, para cada combinación de tamaño de muestra (20, 60, 100, 200, 400, 600) y diseño de los errores (diseño 1 y diseño 2), se presentan los sesgos relativos (b_0 , b_1 , b_2) y el sesgo relativo total para la estimación por MCC del modelo (3), considerando, para la estimación de la matriz de White, los residuos de la estimación por MCC elevados al cuadrado (r^2) y los errores verdaderos elevados al cuadrado (e^2), respectivamente, y considerando 5000 replicaciones. En la figura 1, se presenta el sesgo relativo total para ambos diseños de los errores y ambas estimaciones de la matriz de White.

Por un lado, se puede observar que, para todos los casos, a medida que el tamaño de muestra aumenta, el sesgo relativo total tiende a cero. Por otro lado, se observa que este sesgo es, notablemente, mayor cuando, para la estimación de la matriz de White, se

considera r^2 en comparación a cuando se considera e^2 . Por último, se ve que, para tamaños de muestra pequeños (20, 60, 100), el sesgo relativo total es mayor en los casos del diseño 2, mientras que, para tamaños de muestra grandes (200, 400, 600), este sesgo es mayor en los casos del diseño 1 (excepto para el tamaño de muestra $n=200$ cuando se considera r^2 para la estimación de la matriz de White).

Tabla 3. Simulaciones mediante estimación por MCC-White (r^2) (con 5000 repeticiones).

Diseño	Resultado	n= 20	n= 60	n= 100	n= 200	n= 400	n= 600
Diseño 1	Sesgo relativo b_0	0,230	0,060	0,036	0,017	0,009	0,006
	Sesgo relativo b_1	0,365	0,096	0,052	0,019	0,006	0,000
	Sesgo relativo b_2	0,508	0,113	0,043	0,001	0,027	0,036
	Sesgo relativo total	1,103	0,270	0,132	0,037	0,041	0,042
Diseño 2	Sesgo relativo b_0	0,252	0,064	0,036	0,019	0,009	0,006
	Sesgo relativo b_1	0,431	0,126	0,073	0,032	0,013	0,005
	Sesgo relativo b_2	0,573	0,169	0,084	0,020	0,009	0,026
	Sesgo relativo total	1,256	0,359	0,194	0,070	0,030	0,037

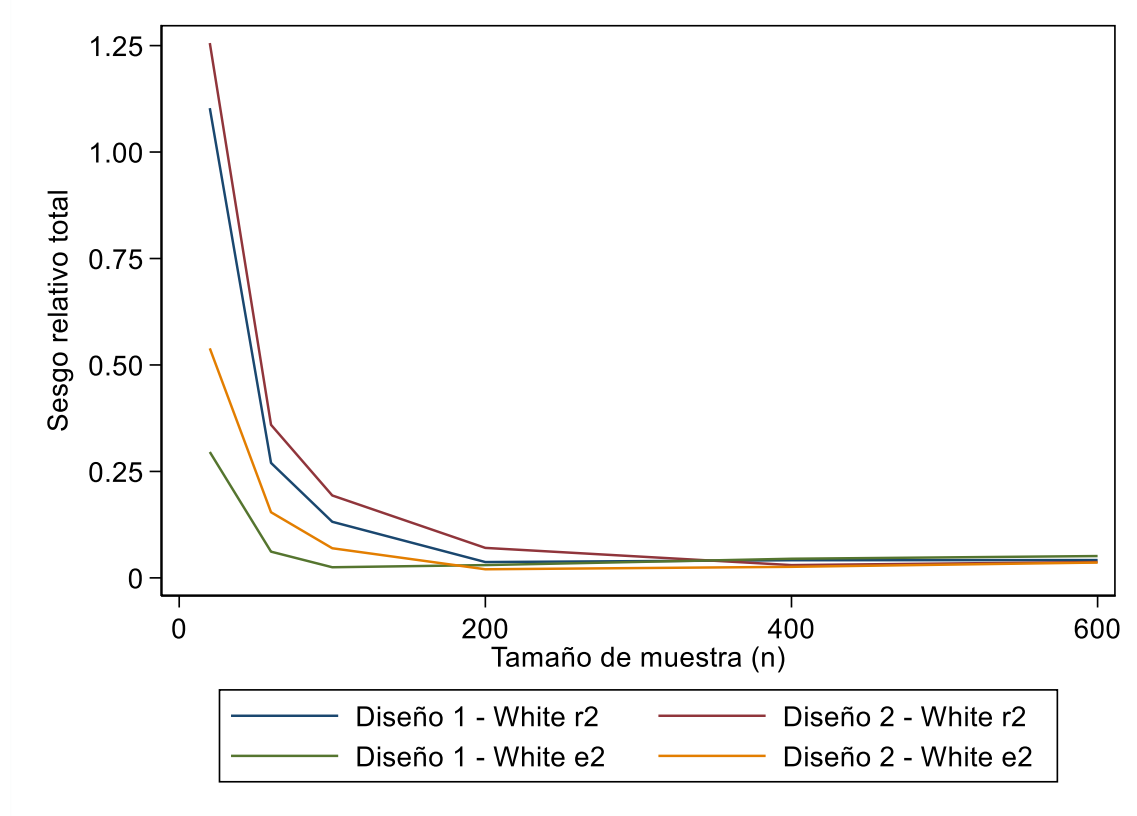
Fuente: Elaboración propia.

Tabla 4. Simulaciones mediante estimación por MCC-White (e^2) (con 5000 repeticiones).

Diseño	Resultado	n= 20	n= 60	n= 100	n= 200	n= 400	n= 600
Diseño 1	Sesgo relativo b_0	0,032	0,008	0,004	0,001	0,001	0,000
	Sesgo relativo b_1	0,099	0,029	0,012	0,000	0,004	0,006
	Sesgo relativo b_2	0,164	0,025	0,009	0,028	0,040	0,045
	Sesgo relativo total	0,296	0,062	0,025	0,030	0,045	0,051
Diseño 2	Sesgo relativo b_0	0,063	0,010	0,004	0,003	0,001	0,001
	Sesgo relativo b_1	0,191	0,059	0,034	0,012	0,003	0,001
	Sesgo relativo b_2	0,285	0,085	0,032	0,005	0,022	0,034
	Sesgo relativo total	0,539	0,154	0,070	0,020	0,026	0,036

Fuente: Elaboración propia.

Figura 2. *Sesgo relativo total (varianza MCC versus varianza White).*



Fuente: Elaboración propia.

Ejercicio (h).

Describir, detalladamente, las propiedades de muestra finita de la estimación de las varianzas de los coeficientes de MCC robustas ante la presencia de heterocedasticidad (con el procedimiento de White) de acuerdo a lo observado en los puntos anteriores.

Las propiedades de muestra finita de la estimación de las varianzas de los coeficientes de MCC robustas ante la presencia de heterocedasticidad (con el procedimiento de White), de acuerdo a lo observado en los puntos anteriores, son:

- Sesgo: Las estimaciones de las varianzas y covarianzas de los coeficientes pueden estar sesgadas debido a la heterocedasticidad. Sin embargo, a medida que el tamaño de muestra aumenta, este sesgo tiende a disminuir.
- Eficiencia: Las estimaciones de las varianzas y covarianzas de los coeficientes corregidas por heterocedasticidad utilizando el procedimiento de White son eficientes en el sentido de que minimizan la varianza de las estimaciones, ya que, ante la presencia de heterocedasticidad, el estimador de MCC no es eficiente.
- Intervalos de confianza y pruebas de hipótesis: Ante la presencia de heterocedasticidad (pero bajo el cumplimiento de los demás supuestos clásicos), el procedimiento de White permite construir intervalos de confianza válidos y realizar pruebas de hipótesis válidas.
- Robustez a la especificación del modelo: Las estimaciones de las varianzas y covarianzas de los coeficientes corregidas por heterocedasticidad utilizando el procedimiento de White son robustas a la especificación del modelo, lo que significa que proporcionan resultados válidos incluso si el modelo subyacente no está completamente especificado.