

¿Qué fue primero: la heterocedasticidad o normalidad?

Un análisis del efecto de la heterocedasticidad sobre las pruebas de normalidad en un diseño experimental con dos factores sin interacción

Isaura Gutiérrez Vargas¹, Aracelly Zárate Cordero¹, Wendy Zárate Cordero¹
isaura.gutierrez@ucr.ac.cr, aracelly.zarate@ucr.ac.cr, wendy.zarate@ucr.ac.cr

RESUMEN

La heterocedasticidad surge cuando los datos no cumplen con el supuesto de homogeneidad de varianzas. El rompimiento de este supuesto puede conllevar a potencias más bajas o un incremento del error tipo I. Sin embargo, este supuesto también es fundamental a la hora de determinar si un conjunto de datos sigue una distribución normal. Con el objetivo de determinar el cumplimiento del supuesto de normalidad ante diferentes niveles de heterocedasticidad, se utilizan simulaciones para encontrar la probabilidad de cometer error tipo I en las pruebas de normalidad Jarque-Bera y Shapiro-Wilk. Como resultado, se obtiene que el incumplimiento del supuesto de homocedasticidad aumenta el error tipo I de ambas pruebas considerablemente. Además, conforme aumenta el número de réplicas, también incrementa la probabilidad de cometer error Tipo I.

PALABRAS CLAVE: simulación, Jarque-Bera, Shapiro-Wilk, error tipo I.

ABSTRACT

Heteroscedasticity arises when the data does not comply with the homogeneity of variances. Breaking this assumption can lead to lower statistical power or an increase in the type I error. However, this assumption is also fundamental when determining if a data set follows the normal distribution. In order to determine compliance with the assumption of normality at different levels of heteroscedasticity, simulations are used to find the probability of committing a type I error in the Jarque-Bera and Shapiro-Wilk normality tests. As a result, it is obtained that the breach of the homoscedasticity assumption increases the type I error of the Jarque-Bera and Shapiro-Wilk tests. In addition, as the number of replicas increases, the probability of committing Type I error also increases.

PALABRAS CLAVE: simulation, Jarque-Bera, Shapiro-Wilk, type I error.

INTRODUCCIÓN

La correcta aplicación de los modelos lineales generales depende de varios supuestos: independencia de las observaciones, datos distribuidos normalmente e igualdad de varianzas u homocedasticidad. En el caso de la homocedasticidad, el cumplimiento de este supuesto garantiza la precisión de los errores estándar y covarianzas asintóticas a lo largo de los parámetros estimados (Rosopa, Schaffer & Schroeder, 2013). Sin embargo, es habitual que muchos de los conjuntos de datos con los que se trabaja incumplan alguno o varios de los supuestos establecidos.

Bajo esta línea, se ha estudiado el efecto que tiene el incumplimiento de los supuestos sobre los resultados del modelo ANOVA. Con respecto a la normalidad, de acuerdo con Lindman (1992), el

¹ Estudiantes de Estadística de la Universidad de Costa Rica

impacto de no cumplir con el supuesto es pequeño y puede ser ignorado. Winer, Brown y Michels (1991), encontraron que la prueba F es robusta a desviaciones moderadas de la normalidad cuando los tamaños de muestra son razonablemente grandes. Además, diversos estudios han aportado evidencia sobre la robustez de los test F bajo no normalidad y homogeneidad de varianzas (Black, Ard, Smith & Schibik, 2010; Schmider, Ziegler, Danay, Beyer & Bühner, 2010).

La heterocedasticidad, por su parte, se define como la situación en la cual la varianza de los errores no es constante. Estudios previos han demostrado que los test F del modelo ANOVA son sensibles a la violación del supuesto de homocedasticidad, por lo que este supuesto tiene consecuencias más graves que la no normalidad. El rompimiento del supuesto de igualdad de varianzas no genera sesgos en los parámetros estimados, sin embargo, sí está asociado a un sesgo en los errores estándar, una menor potencia y errores tipo I inflados (Astivia & Zumbo, 2019; Box, 1954).

Sin embargo, la literatura que explora el efecto que tiene la heterocedasticidad sobre la violación del supuesto de normalidad es escasa. La consecuencia que tiene no cumplir este supuesto no debería ser ignorada, puesto que muchas pruebas de normalidad parten de la existencia de igualdad de varianzas. Por un lado, el estadístico no paramétrico de Shapiro-Wilk se deriva de una muestra de variables aleatorias independientes e idénticamente distribuidas (Shapiro & Wilk, 1965). Por otro lado, la prueba de Jarque-Bera parte del hecho de que existe homogeneidad de varianzas en los datos para derivar su estadístico (Jarque & Bera, 1987).

Si bien ambas prueban cumplen la misma función, tienen usos especializados. Tal como mencionan Flores y Flores (2021); la prueba de Shapiro, es una de las que cuenta con mayor potencia estadística, la cual se recomienda cuando el tamaño de muestra es inferior a cincuenta observaciones. Por el contrario, la prueba de Jarque-Bera presenta mejores resultados con tamaños de muestra grandes y distribuciones simétricas (Pedrosa, Juarros, Robles, Basteiro & Garcia, 2015). Gandica de Roa (2020), llevó a cabo una investigación donde se compararon diversas pruebas de normalidad en condiciones donde se cumplen los supuestos, en la cual se concluye que no hay una prueba que sea mejor que otra en cuanto a potencia y robustez se refiere.

Dado lo anterior, esta investigación se centra en el impacto que tiene la violación del supuesto de homocedasticidad sobre la probabilidad de cometer error tipo I en las pruebas de normalidad usando un modelo factorial sin interacción. Específicamente, se intenta contestar a las siguientes preguntas: ¿Las pruebas de normalidad son válidas ante presencia de heterocedasticidad?, ¿la prueba de Jarque-Bera se ve más distorsionada que la prueba de Shapiro-Wilk? En último lugar, se busca analizar si ante un mayor número de réplicas el efecto de la heterocedasticidad disminuye.

METODOLOGÍA

Para examinar el efecto de la heterocedasticidad sobre las pruebas de normalidad, se realizan simulaciones de conjuntos de datos que provienen de una distribución normal con diferentes variancias. Para realizar la simulación debe tomarse en cuenta varios supuestos que se realizan sobre las variables a utilizar. En primer lugar, se va a suponer que el experimento es un diseño factorial con dos factores, los cuales se identifican con las letras A y B. El factor A cuenta con tres niveles (1, 2, 3) mientras que el factor B tiene dos niveles (a y b) y en total se tienen 6 tratamientos. Los datos son

simulados de forma tal que los factores no presentan interacción y las medias de cada tratamiento son seleccionadas arbitrariamente.

Para dar respuesta a las preguntas planteadas, se diseñan seis escenarios con diferentes niveles de heterocedasticidad. En el primer escenario, se generan datos con homocedasticidad, por lo que se establece que estos provienen de una distribución normal con las medias antes mencionadas y una varianza constante de 5.73 unidades, dicho parámetro fue seleccionado de forma arbitraria.

En el escenario 2, solo uno de los tratamientos presenta una variabilidad alta mientras que en el escenario 3 se decide generar 2 tratamientos con variancias altas. Por otra parte, en el cuarto escenario, tres de los seis tratamientos presentan una baja variabilidad mientras que en los otros tres se presentan variancias elevadas, y en el escenario 5 cuatro tratamientos son generados con variancias desiguales. En el último escenario todos los tratamientos presentan variancias desiguales.

Para demostrar el efecto que tiene la heterocedasticidad sobre la probabilidad de cometer error tipo I se seleccionan las pruebas de Shapiro-Wilk y Jarque-Bera, debido a que estas se basan en diferentes estrategias para determinar la normalidad. La prueba de Shapiro-Wilk se basa en el estadístico W , el cual se define como se muestra:

$$W = \frac{(\sum_{i=1}^n a_i x_{(i)})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

Este estadístico se fundamenta en la idea de que la distribución de una serie de variables aleatorias debería alinearse en una línea recta en un gráfico cuantil-cuantil (Qqplot) y una desviación de esta línea indica falta de normalidad (Sánchez-Espigares, 2018).

La prueba de Jarque-Bera se basa en el estadístico.

$$JB = \frac{n}{6} \left(S^2 + \frac{1}{4} (K - 3)^2 \right)$$

Donde n es el número de observaciones (o grados de libertad en general), S es la asimetría de la muestra, K la curtosis de la muestra

$$S = \frac{\hat{\mu}_3}{\hat{\sigma}^3} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^3}{\left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \right)^{3/2}}$$

$$K = \frac{\hat{\mu}_4}{\hat{\sigma}^4} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4}{\left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \right)^2}$$

Esta prueba a diferencia de la anterior, utiliza el método de momentos debido a que realiza una estimación de la diferencia que existe entre el coeficiente de curtosis y la asimetría (momentos 3 y 4). De manera tal que una distribución es perfectamente normal si el coeficiente de asimetría es 0 y la curtosis tiene un valor de 3. (Sánchez, Grima & Marco, 2018; Siddiqi, 2014).

De esta forma por medio de la simulación se desea estimar la probabilidad de que las pruebas rechacen la hipótesis de normalidad de los residuos cuando estos provienen de una distribución normal, es decir, se cuantifica la probabilidad de cometer error tipo I. Por último, es importante

mencionar que los escenarios que se han establecido son analizados tomando en cuenta diferente cantidad de réplicas. Durante la simulación para el escenario de homocedasticidad se emplean 100000 iteraciones mientras que en los escenarios de heterocedasticidad se realizan 2000 iteraciones, en ambos escenarios se utiliza 5% de significancia.

Para la programación utilizada en este estudio se utiliza el lenguaje R en su versión 4.1.1 (R Core Team, 2017), para la creación de gráficos interactivos se utilizan los paquetes ggplot2: Elegant Graphics for Data Analysis (Wickham, 2016) y plotly (Sievert, 2020.). Para realizar la prueba Jarque-Bera se hace uso del paquete tseries: Time series analysis and computational finance (Trapletti & Hornik, 2022). Por último, para la creación de la aplicación se utilizan los paquetes shiny y shiny themes (Chang, Cheng, Allaire, Sievert, Schloerke, Xie, Allen, McPherson, Dipert & Borges, 2021). El código llevado a cabo para generar la aplicación, así como su respectiva explicación fue incluido dentro de los anexos.

RESULTADOS

En la Tabla 1, se presentan los resultados de la simulación para la probabilidad de cometer error tipo I en la prueba de Shapiro-Wilk. Se puede notar como para el escenario 1, el cual se planteó con igualdad de varianzas, la prueba de Shapiro-Wilk tiene una probabilidad de cometer error tipo I baja, sin importar el número de réplicas. Si se compara con los demás escenarios la diferencia es notable pues, bajo la presencia de heterocedasticidad, es posible notar que la probabilidad de rechazar erróneamente la hipótesis nula aumenta conforme lo hace el número de réplicas.

Además, los escenarios con solo uno o dos tratamientos con varianzas desiguales parecen verse especialmente afectados, porque incluso con una cantidad pequeña de réplicas por tratamiento presentan una probabilidad muy alta de cometer error tipo I. Para el escenario 2, el cual corresponde al caso donde solo un tratamiento tiene una varianza mayor a las demás, se alcanza una probabilidad de 1 de cometer error tipo I con cantidades superiores a 14 réplicas por tratamiento. En el escenario 3 se alcanza la misma probabilidad con 10 réplicas por tratamiento.

En la Tabla 2, se indican los resultados para la prueba de normalidad Jarque-Bera. Al igual que la prueba de Shapiro-Wilk, en casos de homocedasticidad, la prueba se comporta como se esperaba, puesto la probabilidad de cometer error tipo I se mantiene baja al aumentar las réplicas. También se observa que en los escenarios heterocedásticos, conforme aumenta la cantidad de réplicas la probabilidad de cometer error tipo I también crece. El escenario 2 y 3 sigue un comportamiento similar al observado con la prueba de Shapiro-Wilk.

Tabla 1

Probabilidad de cometer error tipo I con la prueba de Shapiro-Wilk

Número de réplicas	3	5	7	10	14	26	35	40	45	50
Escenario 1	0.020	0.037	0.050	0.064	0.047	0.050	0.049	0.055	0.052	0.042
Escenario 2	0.374	0.657	0.814	0.940	0.979	1	1	1	1	1
Escenario 3	0.784	0.963	0.996	1	1	1	1	1	1	1

Escenario 4	0.063	0.169	0.252	0.342	0.457	0.723	0.848	0.910	0.938	0.950
Escenario 5	0.051	0.139	0.177	0.264	0.351	0.634	0.774	0.846	0.905	0.918
Escenario 6	0.093	0.210	0.275	0.380	0.495	0.772	0.897	0.937	0.962	0.967

Tabla 2

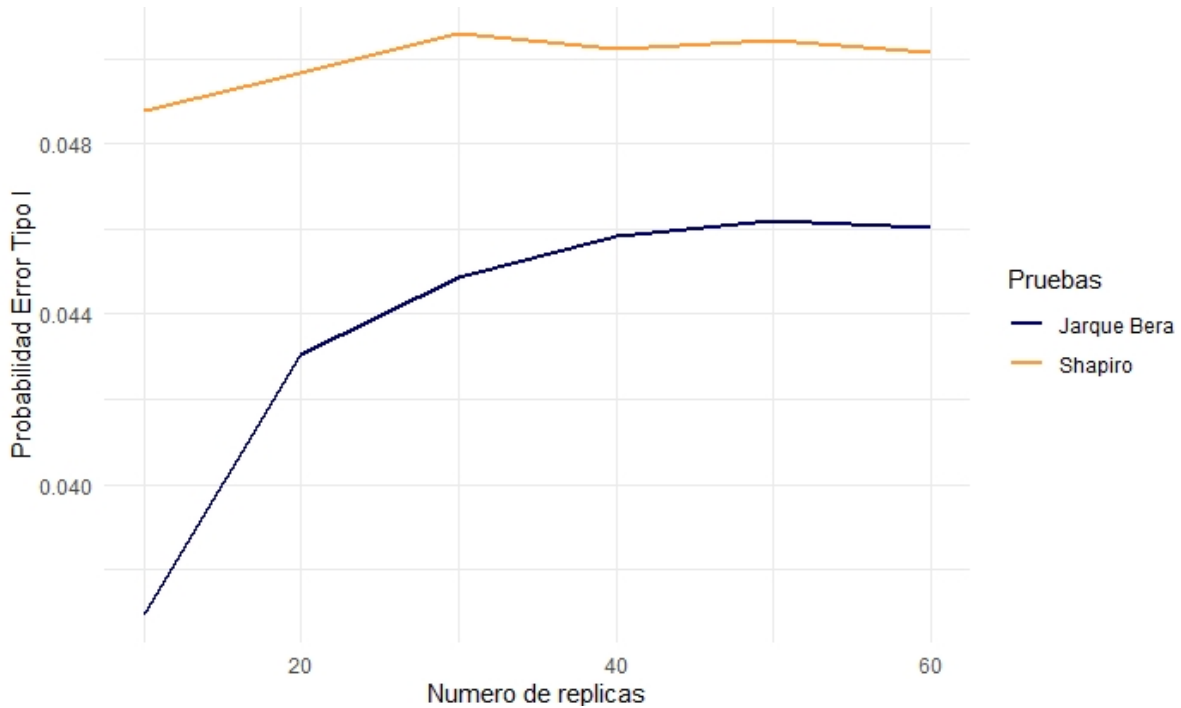
Probabilidad de cometer error tipo I con la prueba de Jarque-Bera

Número de replicas	3	5	7	10	14	26	35	40	45	50
Escenario 1	0.009	0.024	0.027	0.044	0.042	0.056	0.043	0.054	0.052	0.042
Escenario 2	0.382	0.672	0.798	0.946	0.984	1	1	1	1	1
Escenario 3	0.576	0.864	0.973	0.998	1	1	1	1	1	1
Escenario 4	0.062	0.202	0.253	0.389	0.527	0.724	0.835	0.872	0.931	0.935
Escenario 5	0.037	0.122	0.210	0.254	0.348	0.545	0.684	0.719	0.795	0.824
Escenario 6	0.065	0.194	0.309	0.452	0.562	0.790	0.884	0.949	0.953	0.968

Finalmente, en la Figura 1, se puede analizar el resultado de ambas pruebas en escenarios de homocedasticidad. Se observa que ambas pruebas mantienen una probabilidad baja de cometer error tipo I, empero se puede apreciar que la prueba de Jarque-Bera tiende a presentar una probabilidad más baja de cometer error tipo I en comparación con la prueba de Shapiro-Wilk, además, el comportamiento de ambas pruebas es estable al aumentar el número de réplicas.

Figura 1

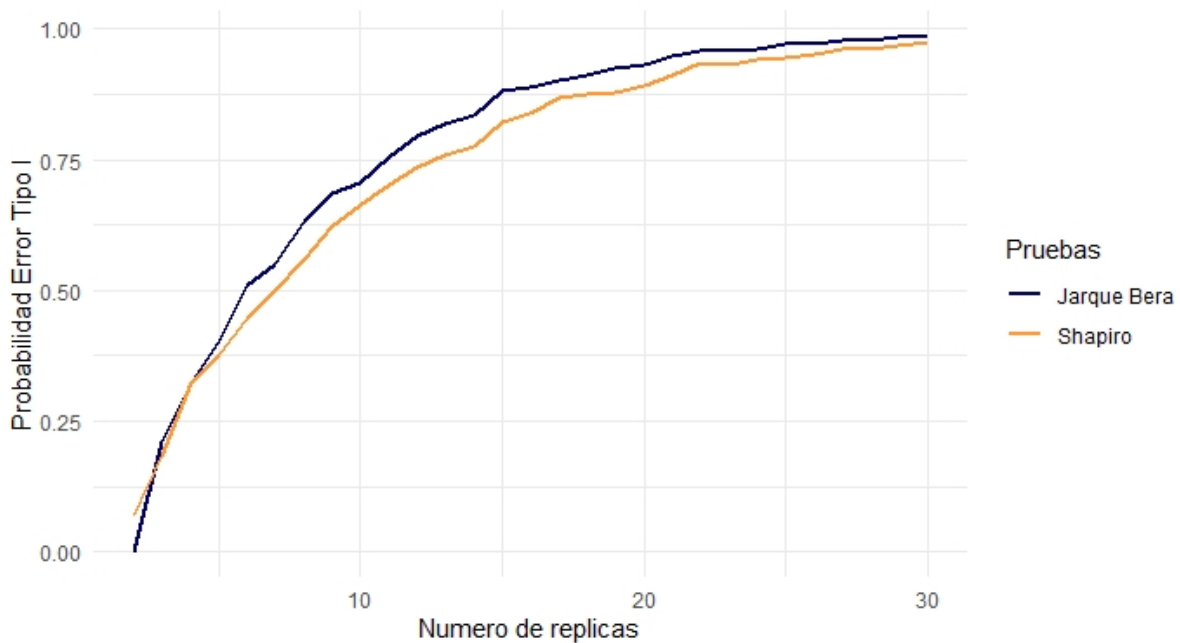
Representación gráfica de la probabilidad de cometer el error tipo 1, comparación de la prueba de Shapiro-Wilk y Jarque-Bera en escenarios de Homocedasticidad.



Por el contrario, al analizar la Figura 2, se evidencia el efecto que tiene la heterocedasticidad en las pruebas para detectar normalidad. Para ambas pruebas la probabilidad de cometer error tipo I aumenta considerablemente conforme lo hace el número de réplicas por tratamiento. En este caso no se observa que una prueba tenga mejores resultados en comparación con la otra y resulta evidente que en escenarios de heterocedasticidad no se puede confiar en los resultados de las pruebas de normalidad dado que es casi seguro que se está cometiendo un error.

Figura 2

Representación gráfica de la probabilidad de cometer el error tipo 1, comparación de la prueba de Shapiro-Wilk y Jarque-Bera en el escenario 3: dos tratamientos con varianzas desiguales



CONCLUSIONES

En este artículo, se ha hecho uso de procedimientos basados en simulaciones para derivar la probabilidad de cometer error Tipo I para las pruebas de normalidad Shapiro-Wilk y Jarque-Bera. Como primer hallazgo, se logró observar que todos los niveles de heterocedasticidad que se probaron aumentan la probabilidad de cometer error tipo 1 en las pruebas de normalidad. En cuanto al rendimiento que presentan ambas pruebas, en línea con lo encontrado por Gandica de Roa (2020), en este análisis se encuentra evidencia de que una prueba no es mejor que la otra, ya que ambas ven su probabilidad de cometer error tipo I alterado.

En segundo lugar, se evidencia que bajo heterocedasticidad, conforme aumenta el tamaño de las réplicas también se tiende a aumentar la probabilidad del error tipo I. Para la prueba de Shapiro-Wilk, este resultado se puede explicar debido a que el efecto de la heterocedasticidad se ve reflejado en la inflación del error estándar, lo cual afecta la prueba pues ante errores estándar más grandes la potencia del contraste disminuye.

En relación con la prueba Jarque-Bera, los autores Jarque y Bera (1987) y Pedrosa et al (2015), afirman que esta prueba es recomendable cuando se cuenta con homocedasticidad y distribuciones simétricas. Estas afirmaciones concuerdan con lo hallado en las simulaciones, puesto que es evidente que en los diferentes escenarios heterocedásticos la probabilidad de rechazar erróneamente la hipótesis nula es alta. Este comportamiento se puede explicar debido a que la heterocedasticidad está afectando la asimetría y el coeficiente de curtosis en la distribución de los residuos.

Uno de los supuestos más cruciales para la correcta aplicación de métodos estadísticos es la normalidad. Por medio del uso de simulaciones podemos ver el comportamiento de las pruebas para diagnosticar normalidad ante una serie de condiciones. No obstante, no es correcto valerse de un único test para concluir si existe normalidad en los datos, se recomienda iniciar la inspección por medio de gráficos de QQ-plot o gráficos de densidad que permitan analizar de manera descriptiva los datos y así obtener conclusiones más robustas.

BIBLIOGRAFÍA

Astivia, O.L., & Zumbo, B.D. (2019). Heteroskedasticity in Multiple Regression Analysis: What it is, How to Detect it and How to Solve it with Applications in R and SPSS. *Practical Assessment, Research and Evaluation*, 24(1), 1-16. <https://doi.org/10.7275/q5xr-fr95>

Black, G., Ard, D., Smith, J., & Schibik, T. (2010). The impact of the Weibull distribution on the performance of the single-factor ANOVA model. *International Journal of Industrial Engineering Computations*, 1, 185-198.

Box, G. E. P. (1954). *Some Theorems on Quadratic Forms Applied in the Study of Analysis of Variance Problems, I. Effect of Inequality of Variance in the One-Way Classification*. *The Annals of Mathematical Statistics*, 25(2), 290–302. doi:10.1214/aoms/1177728786

Chang, W., Cheng, J., Allaire, J., Sievert, C., Schloerke, B., Xie, Y., Allen, J., McPherson, J., Dipert, A. & Borges, B. (2021). Shiny: Web Application.

Flores, C., & Flores, K. (2021). Pruebas para comprobar la normalidad de datos en procesos productivos: Anderson-Darling, Ryan-Joiner, Shapiro-Wilk y Kolmogórov-Smirnov. *Societas*, 23(2), 83-97. <http://portal.amelica.org/ameli/jatsRepo/341/3412237018/3412237018.pdf>

Gandica de Roa, E. (2020). Potencia y Robustez en Pruebas de Normalidad con Simulación Montecarlo. *Revista Cientific*, 5(18), 108-119. http://www.indteca.com/ojs/index.php/Revista_Cientific/article/view/468

Jarque, C. M., & Bera, A. K. (1987). A Test for Normality of Observations and Regression Residuals. *International Statistical Review / Revue Internationale de Statistique*, 55(2), 163. doi:10.2307/1403192

Lindman, H. R. (1992). Analysis of variance in experimental design (S. Fienberg and I. Olkin, Eds.). Springer-Verlag New York, Inc.

Pedrosa, I., Juarros, J., Robles, A., Basteiro, J., & García, E. (2015). Pruebas de bondad de ajuste en distribuciones simétricas, ¿qué estadístico utilizar? *Revista Javeriana*, 14(1), 245-254. <https://revistas.javeriana.edu.co/index.php/revPsycho/article/view/8046>

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

Rosopa, P. J., Schaffer, M. M., & Schroeder, A. N. (2013). *Managing heteroscedasticity in general linear models*. *Psychological Methods*, 18(3), 335–351. doi:10.1037/a0032553

Sánchez, J., Grima, P., & Marco, L. (2018). Graphical comparison of normality tests for unimodal distribution data. *Journal of Statistical Computation and Simulation*, 89(1), 145-154, DOI: 10.1080/00949655.2018.1539085

Shapiro, S.S., & Wilk, M.B. (1965). An Analysis of Variance Test for Normality (Complete Samples). *Biometrika*, 52, 591-611.

Schmider, E., Ziegler, M., Danay, E., Beyer, L., & Bühner, M. (2010). Is it really robust? Reinvestigating the robustness of ANOVA against violations of the normal distribution assumption. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, 6(4), 147–151. <https://doi.org/10.1027/1614-2241/a000016>

Siddiqi, A. (2014), An observatory note on tests for normality assumptions, *Journal of Modelling in Management*, 9(3), 290-305. <https://doi.org/10.1108/JM2-04-2014-0032>

Sievert, C. (2020), *Interactive Web-Based Data Visualization with R, plotly, and shiny*. Chapman and Hall/CRC Florida

Trapletti, A. & Hornik, K. (2022). *tseries: Time Series Analysis and Computational Finance*. R package version 0.10-51., <https://CRAN.R-project.org/package=tseries>.

Wickham et. Al., (2019). Welcome to the tidyverse. *Journal of Open Source software*, 4(43), 1686, <https://doi.org/10.21105/joss.01686>

Winer, B. J., Brown, D. R., & Michels, K. M. (1991). *Statistical principles in experimental design*. McGraw-Hill Humanities.

ANEXO

Simulación Error Tipo I	
Código	Explicación
<pre>sh = function (r1, mu, v) { k=length(mu) n=r1*k y=rnorm(n=n, mean=rep(mu, each=r1), sd=sqrt(rep(v, each=r1))) factor.a=factor(c(rep(1, each=n/3), rep(2, each=n/3), rep(3, each=n/3))) factor.b=factor(rep(c("A", "B"), each=r1, times=3)) mod=lm(y~factor.a*factor.b) p=shapiro.test(residuals(mod))\$p.value return(p) } jb=function (r1, mu, v) { k=length(mu) n=r1*k y=rnorm(n=n, mean=rep(mu, each=r1), sd=sqrt(rep(v, each=r1))) factor.a=factor(c(rep(1, each=n/3), rep(2, each=n/3), rep(3, each=n/3))) factor.b=factor(rep(c("A", "B"), each=r1, times=3)) mod=lm(y~factor.a*factor.b) p=jarque.bera.test(residuals(mod))\$p.value return(p) } #Escenario con Heterocedasticidad r=c(12,12,12,12,12,12,12,12,12) mu=c(input\$mu1,input\$mu2,input\$mu3, input\$mu4, input\$mu5, input\$mu6) v=c(input\$v1,input\$v2,input\$v3, input\$v4, input\$v5, input\$v6) almacen.sh2=c() almacen.jb2=c() for(i in 1:length(r)){ M=1000 sha=c() jab=c() r1=r[i]</pre>	<p>En esta primera parte se crean dos funciones. La primera para obtener el p-value con la prueba de Shapiro-Wilk, dentro de la función se generan datos para dos factores según las medias, réplicas y variancias que se introduzcan en la función.</p> <p>Ahora se crea la segunda función para obtener el p-value con la prueba de Jarque Bera, e igualmente se generan vectores para cada factor según la medias, réplicas y variancias digitadas en la función.</p> <p>Luego se establecen las réplicas por tratamiento, en este punto es necesario realizar una aclaración: para hacer el análisis en el presente trabajo, se fijaron diez tamaños de replicas (3,5,7,10,14,26,35,40,45,50), sin embargo, en el código a la izquierda el numero de replicas se fijo en 12 por comodidad en el uso de la aplicación Shiny. Las medias y variancias son digitados por el usuario.</p> <p>Se procede a crear vectores para almacenar los p-values que se obtengan al realizar las pruebas, cada valor que se guarde en este vector corresponde a la probabilidad de cometer error tipo I según cada número de réplicas que se digiten en el vector r.</p> <p>A continuación, con el uso del ciclo se realizan 1000 iteraciones para estimar la probabilidad de cometer error tipo I, estos valores se guardan en los vectores “sha” y</p>

```

for(j in 1:M) {
  p1=sh(r1,mu,v)
  p2=jb(r1,mu,v)
  sha=c(sha, p1)
  jab=c(jab, p2)
}
almacen.sh2=c(almacen.sh2, mean(sha<0.05))
almacen.jb2=c(almacen.jb2, mean(jab<0.05))
}

Shapiro=mean(sha<0.05)
Jarque=mean(jab<0.05)
tabla=as.data.frame(cbind(Shapiro, Jarque))
tabla
#Escenario con Homocedasticidad
r=c(12,12,12,12,12,12,12,12,12,12)
mu=c(input$mu1,input$mu2,input$mu3, input$mu4,
input$mu5, input$mu6)
var=input$var

almacen.sh=c()
almacen.jb=c()
for(i in 1:length(r)){
  M=1000
  sha=c()
  jab=c()
  r1=r[i]
  for(j in 1:M) {
    p1=sh(r1,mu,var)
    p2=jb(r1,mu,var)
    sha=c(sha, p1)
    jab=c(jab, p2)
  }
  almacen.sh=c(almacen.sh, mean(sha<0.05))
  almacen.jb=c(almacen.jb, mean(jab<0.05))
}

Shapiro=mean(sha<0.05)
Jarque=mean(jab<0.05)
tabla=as.data.frame(cbind(Shapiro, Jarque))
tabla

#Para la generación de los gráficos
#Caso de heterocedasticidad
r=c(2:10)
mu=c(input$mu1,input$mu2,input$mu3, input$mu4,
input$mu5, input$mu6)
v=c(input$v1,input$v2,input$v3, input$v4, input$v5,
input$v6)

almacen.sh2=c()

```

“jab”, este proceso para cada replica digitada, además, se están utilizando las funciones planteadas al inicio

El resultado que finalmente se muestra en Shiny corresponde a la probabilidad de cometer error tipo I pero solo con un tamaño de réplica a la vez (por esta razón se fijó el número de réplicas en 12).

En el caso de la simulación con homocedasticidad, el código funciona exactamente igual que en el caso con heterocedasticidad, la única diferencia es que al usuario solo se le solicita una única variancia, ya que se está suponiendo que todos los tratamientos tienen la misma variancia.

Para representar la probabilidad de cometer el error tipo I, se estableció que las replicas van desde 2 hasta 10, esto se define de esta forma para obtener suficientes puntos y graficar una curva que represente dicha probabilidad, también se solicita al usuario las variancias y las medias para seis tratamientos.

```

almacen.jb2=c()
for(i in 1:length(r)){
  M=1000
  sha=c()
  jab=c()
  r1=r[i]
  for(j in 1:M) {
    p1=sh(r1,mu,v)
    p2=jb(r1,mu,v)
    sha=c(sha, p1)
    jab=c(jab, p2)
  }
  almacen.sh2=c(almacen.sh2, mean(sha<0.05))
  almacen.jb2=c(almacen.jb2, mean(jab<0.05))
}

x1=rep(2:10, times=2)
tab2=data.frame(x1, values=c(almacen.sh2, almacen.jb2),
Pruebas=factor(rep(c("Shapiro", "Jarque Bera"), each=9)))
g1 <- ggplot(tab2, aes(x1, values, colour=Pruebas),
colour=Pruebas)+geom_line()+ylab("Probabilidad Error Tipo
I")+xlab("Numero de
replicas")+theme_minimal()+scale_color_manual(values =
c("#000066", "#FF9933"))
ggplotly(g1)
#Grafico Homocedasticidad
r=c(2:10)
mu=c(input$mu1,input$mu2,input$mu3, input$mu4,
input$mu5, input$mu6)
var=input$var

almacen.sh=c()
almacen.jb=c()
for(i in 1:length(r)){
  M=1000
  sha=c()
  jab=c()
  r1=r[i]
  for(j in 1:M) {
    p1=sh(r1,mu,var)
    p2=jb(r1,mu,var)
    sha=c(sha, p1)
    jab=c(jab, p2)
  }
  almacen.sh=c(almacen.sh, mean(sha<0.05))
  almacen.jb=c(almacen.jb, mean(jab<0.05))
}
x=rep(2:10, times=2)
tab1=data.frame(x, values=c(almacen.sh, almacen.jb),
Pruebas=factor(rep(c("Shapiro", "Jarque Bera"), each=9)))

```

De nuevo se crean vectores para almacenar la probabilidad de cometer el error tipo I con cada una de las réplicas antes establecidas, se realiza un ciclo utilizando las funciones planteadas para estimar dicha probabilidad, efectúan 1000 iteraciones.

Ahora se crea un data frame, con los valores del eje x y las probabilidades obtenidas, dicho data frame es necesario para crear el grafico utilizando la función ggplot. Con ayuda de dicha función se genera un gráfico de curvas, donde cada curva representa la probabilidad de cometer error tipo I de cada una de las pruebas. Para generar el grafico con homocedasticidad, se sigue el procedimiento antes descrito, con la diferencia de que el usuario solo digita una variancia.

```
g2 <- ggplot(tab1, aes(x, values, colour=Pruebas,  
colour=Pruebas)+geom_line()+ylab("Probabilidad Error Tipo  
I")+xlab("Numero de  
replicas")+theme_minimal()+scale_color_manual(values =  
c("#000066", "#FF9933"))  
ggplotly(g2)
```