



FACULTADE DE MATEMÁTICAS

Traballo Fin de Grao

Contrastes de normalidad

Víctor Platas García

2020/2021

UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

GRAO DE MATEMÁTICAS

Traballo Fin de Grao

Contrastes de normalidad

Víctor Platas García

Julio 2021

UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Trabajo propuesto

Área de Coñecemento: Estadística e Investigación Operativa
Título: Contrastes de normalidad
Breve descripción do contido
Muchos procedimientos estadísticos parten del supuesto de que los datos proceden de una distribución normal. Sin embargo, es muy habitual que los datos reales no cumplan esta suposición, lo cual repercutirá en la validez de tales procedimientos. Por este motivo son importantes los contrastes de normalidad. En este trabajo se revisarán los contrastes más conocidos, como el test de Kolmogorov-Smirnov-Lilliefors o el test de Shapiro-Wilk, y se completarán con otros tests, como los basados en la asimetría o la curtosis. Se estudiarán sus propiedades teóricas y se realizarán estudios de simulación para comprobar su funcionamiento en datos simulados.
Recomendacións
Outras observacións

Índice general

Resumen	VIII
Introducción	XI
1. La distribución normal	1
1.1. Contexto histórico	1
1.2. Definición	2
1.3. Teorema Central del Límite	5
1.4. Estadística inferencial	6
1.4.1. Estimación de parámetros de máxima verosimilitud	7
2. El test Chi Cuadrado de Pearson	9
2.1. Contraste de una hipótesis nula simple	9
2.2. Contraste de un modelo paramétrico	11
3. Test basados en la función de distribución	15
3.1. Test de Kolmogorov-Smirnov y versión de Lilliefors	15
3.1.1. Prueba de Kolmogorov-Smirnov	15
3.1.2. Test de Lilliefors	17
3.2. Test de Cramer	17
3.3. Test de Anderson-Darling	18
3.4. Relación Kolmogorov-Cramer-Anderson	19
4. Test basados en los momentos	21
4.1. Asimetría	21
4.1.1. Definición	21
4.1.2. Medidas de asimetría	22
4.1.3. Asimetría en la distribución normal	23
4.1.4. Test de normalidad basado en la asimetría	24

4.2. Curtosis	25
4.2.1. Definición	25
4.2.2. Medidas de curtosis	25
4.2.3. Curtosis en la distribución normal	26
4.2.4. Test de normalidad basado en la curtosis	27
4.3. Test de Jarque-Bera	28
5. Métodos gráficos	31
5.1. Histograma y curva normal	31
5.2. Gráfico de cuantiles teóricos	33
6. Test Shapiro-Wilk y Shapiro-Francia	37
6.1. Test de Shapiro-Wilk	37
6.2. Test de Shapiro-Francia	39
7. Estudio comparativo mediante simulaciones	41
7.1. Resultados de las simulaciones	41
7.2. Conclusiones	50
A. Códigos de R	51
Bibliografía	61

Resumen

Este trabajo se centra en los distintos test de normalidad existentes, comenzando en primer lugar con un breve estudio de la distribución normal. A continuación se presentan los distintos test, profundizando uno por uno en sus diferentes características principales como los estadísticos de contraste, realizando un estudio de cada uno dónde también se analizan las diferencias más importantes entre ellos. Complementariamente, se añade un estudio realizado con *R* donde se analiza la fiabilidad y veracidad de cada test para distintos conjuntos de datos simulados que siguen diversas distribuciones. Finalmente, también se anexan los códigos de *R* utilizados para la realización de las simulaciones.

Abstract

This work focuses on the different existing normality tests, starting firstly with a brief study of the Normal distribution. The tests are presented below, delving one by one into their different main characteristics such as contrast statistics, developing a study of each one of them where the most important differences between them are also analyzed. Complementarily, a study carried out with *R* is added where the reliability and veracity of each test is analyzed for different sets of simulated data that follow diverse distributions. Finally, the *R* codes used to carry out the simulations are also attached.

Introducción

En estadística, los test de normalidad son usados para determinar si un conjunto de datos se ajusta a una distribución normal y para calcular la probabilidad de que una variable del conjunto de datos se distribuya normalmente.

Más exactamente, estas pruebas son una forma de selección de modelo que podemos interpretar de varias maneras, según la interpretación que se haga de la probabilidad:

- En términos de estadística descriptiva, se mide la bondad de ajuste de un modelo normal a los datos. Si dicho ajuste es deficiente, los datos no están bien modelados mediante una distribución normal.
- En las pruebas de hipótesis de la estadística frecuentista, los datos se contrastan con la hipótesis nula de que se distribuyen normalmente.
- En las estadísticas bayesianas, uno no “prueba la normalidad” como tal, sino que calcula la probabilidad de que los datos provengan de una distribución normal con parámetros dados μ, σ y lo compara con la probabilidad de que los datos provienen de otras distribuciones, la mayoría simplemente usando un factor de Bayes (que da la probabilidad relativa de ver los datos dados diferentes modelos), o tomando una distribución previa sobre posibles modelos y parámetros y calculando una distribución posterior dadas las probabilidades calculadas.

Es decir, se utiliza una prueba de normalidad para determinar si los datos de la muestra se han extraído de una población distribuida normalmente (dentro de cierta tolerancia).

Otra forma más informal de probar la normalidad es comparar una curva de probabilidad normal con un histograma de los datos de la muestra. La distribución empírica de los datos (el histograma) debería tener forma de campana y asemejarse a una distribución normal. Si la muestra es pequeña, se podría realizar una regresión de los datos contra los cuantiles de una distribución normal con misma media y varianza que la muestra. De haber poco ajuste a la línea de regresión supondría que podemos estar ante una desviación de la normalidad.

Una herramienta para evaluar la normalidad es la gráfica de probabilidad normal, es decir, una gráfica cuantil-cuantil (QQ-plot) de la distribución normal estándar tal y como veremos más adelante. La correlación entre los datos de la muestra y los cuantiles normales mide cuanto de bien se modelan los datos mediante una distribución normal. Estos gráficos son fáciles de interpretar y también tienen la ventaja de que los valores atípicos se encuentran rápidamente.

Una de las aplicaciones más comunes de los test de normalidad es la de los residuos de un modelo de regresión lineal.[1] Si no tienen una distribución normal, los residuos no deben usarse en ninguna prueba derivada de la distribución normal, como las pruebas t , las F o las chi-cuadrado. Si los residuos no se distribuyen normalmente, entonces la variable dependiente puede tener la forma incorrecta, pueden faltar variables importantes, etc. La corrección de uno o más de estos errores sistemáticos puede producir residuos que se distribuyen normalmente; en otras palabras, la no normalidad de los residuos es a menudo una deficiencia del modelo más que un problema de datos.

En el primer capítulo de este trabajo, se realiza un estudio de la distribución normal, repasando brevemente su contexto histórico, además de recordar sus principales propiedades. A continuación, el Teorema Central del Límite también es enunciado junto con sus características más importantes, para finalmente terminar el capítulo analizando la influencia de la distribución normal en la inferencia estadística. Seguidamente, en el segundo capítulo se realiza un estudio del test Chi Cuadrado de Pearson donde se exponen sus principales características y funcionamiento tanto para una hipótesis nula simple como para un modelo paramétrico. En el tercero, los test de Kolmogorov-Smirnov, Cramér-von-Mises y Anderson-Darling son explicados y se expone la estrecha relación que existe entre ellos pues todos están basados en la función de distribución. Continuando con el cuarto capítulo, la asimetría y la curtosis son definidas y se analiza su influencia en la distribución normal junto con un análisis del funcionamiento del test de Jarque-Bera que emplea ambos parámetros. En el quinto capítulo, se explican los métodos gráficos empleados para estudiar la normalidad como los QQ-Plots, junto con varios ejemplos que ayudan a ver la informalidad de dichos métodos. En el sexto capítulo se da una visión profunda del test de Shapiro-Wilk y de como es construido su estadístico, además de mostrar las diferencias que guarda con su versión de Shapiro-Francia. En el último capítulo, se realiza una simulación en R para cada uno de los test estudiados previamente para ver su comportamiento frente a distintas muestras que siguen unas determinadas distribuciones, de donde se extraen las conclusiones que concluyen el capítulo. Finalmente, en el apéndice se muestran todos los códigos de R utilizados a lo largo del trabajo, así como las explicaciones de las funciones de R empleadas en las simulaciones realizadas.

Capítulo 1

La distribución normal

Para hablar de los test de normalidad, previamente debemos hablar de la distribución como tal, por lo que brevemente repasaremos unas nociones básicas sobre su origen, así como sus principales características y su influencia en el mundo de la estadística y la probabilidad.

1.1. Contexto histórico

El primer estudio realizado sobre la distribución normal fue llevado a cabo por Abraham de Moivre en un artículo de 1733, en una segunda impresión de la segunda edición de su libro *The Doctrine of Chances*, de 1738, analizando la aproximación de la distribución binomial para valores grandes de n . Dicho resultado fue ampliado por Laplace en su libro *Teoría analítica de las probabilidades* (1812), y que en la actualidad recibe el nombre de *Teorema de De Moivre-Laplace*.

Laplace comenzó usando la distribución normal en el análisis de errores de experimentos. Además, en 1805, Legendre introdujo por primera vez el método de mínimos cuadrados, pero Gauss, que sostenía que llevaba usando el método desde 1794, lo justificó en 1809 gracias a que asumió que los errores seguían una distribución normal. Es desde entonces que el nombre de Gauss se asocia a esta distribución dado su continuo uso al analizar diversos datos astronómicos, por lo que frecuentemente se le atribuye un descubrimiento independiente del de De Moivre.

Acerca de la *campana de Gauss*, fue Esprit Jouffret el que usó por primera vez el término "bell surface" (superficie campana) en 1872 para hablar de una distribución normal bivalente con componentes independientes. Finalmente, el nombre de *distribución normal* comenzó a emplearse hacia 1875.

Así pues, la distribución normal vio la luz hace ya más de 200 años, desde sus aparición

a mediados del siglo XVIII en distintos estudios hasta la actualidad, cuyo estudio y formalización ha hecho de esta distribución una de las bases de la estadística que conocemos hoy en día.

1.2. Definición

La distribución normal, también conocida como distribución de Gauss o distribución gaussiana es una distribución de probabilidad de variable continua que suele aparecer frecuentemente tanto en teoría de probabilidades como en la propia estadística.

Para definir formalmente la distribución normal veamos en primer lugar los parámetros que la componen:

- μ : es la media (aunque también puede ser la moda, la mediana o el valor esperado).
- σ : es la desviación típica.
- σ^2 : es la varianza.
- φ : es la función de densidad de probabilidad.

De los parámetros anteriores, destacamos la función de densidad de probabilidad, más comúnmente conocida como *campana de Gauss* que viene dada por:

$$\varphi(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-(x-\mu)^2/2\sigma^2}$$

Y cuya gráfica, también conocida globalmente, es la siguiente:

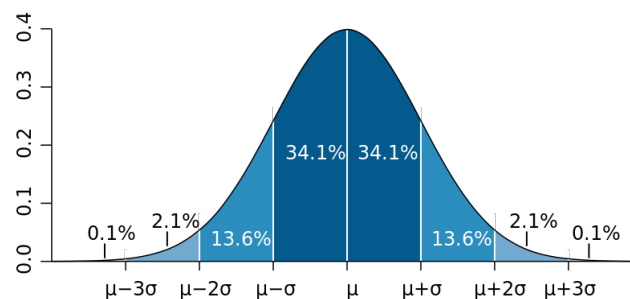


Figura 1.1: Función de densidad de probabilidad

Una vez definida la función de densidad de probabilidad, podemos definir la función de distribución de la distribución normal: [2]

$$\Phi_{\mu, \sigma^2}(x) = \int_{-\infty}^x \varphi_{\mu, \sigma^2}(u) du = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(u-\mu)^2}{2\sigma^2}} du \quad (1.1)$$

La cual, del mismo modo que vimos para la función de densidad, presenta la siguiente gráfica en función de los distintos valores de los parámetros μ y σ^2 :

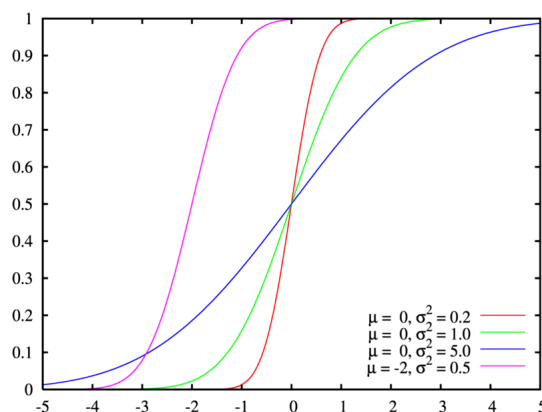


Figura 1.2: Función de distribución

Veamos a continuación las principales propiedades de la distribución normal.

Propiedades

1. Es simétrica respecto de la media μ .
2. El valor de la media μ coincide con el valor de la moda y mediana.
3. Los puntos de inflexión de la función de densidad se dan en $x = \mu - \sigma$ y en $x = \mu + \sigma$.
4. Como podemos observar en la figura 1.1 en el intervalo $[\mu - 2\sigma, \mu + 2\sigma]$ encontramos cerca de el 95,5% de la distribución, de donde se obtienen los límites de las tablas utilizadas habitualmente en la normal estándar.
5. Dado $X \sim N(\mu, \sigma^2)$ y $a, b \in \mathbb{R}$, entonces

$$aX + b \sim N(a\mu + b, a^2\sigma^2)$$

6. Suponiendo $X \sim N(\mu_X, \sigma_X^2)$ y $Y \sim N(\mu_Y, \sigma_Y^2)$ son dos variables aleatorias independientes, entonces:

- La suma de ambas también seguirá una distribución normal, es decir:

$$U = X + Y \sim N(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2)$$

Además, se tiene que dadas dos variables aleatorias independientes cuya suma sigue una distribución normal, podemos asegurar que ambas variables serán normales.

- Al igual que la suma, la diferencia de ambas variables también sigue una distribución normal, es decir:

$$V = X - Y \sim N(\mu_X - \mu_Y, \sigma_X^2 + \sigma_Y^2)$$

- Si las varianzas de las variables X e Y son iguales, entonces U y V son independientes entre sí.
7. Dadas n variables normales estándar independientes X_1, \dots, X_n , la suma de dichas variables $X_1^2 + \dots + X_n^2$ seguirá una distribución χ^2 con n grados de libertad.
 8. Dadas n variables normales estándar independientes X_1, \dots, X_n , la media muestral $\bar{X} = (X_1 + \dots + X_n)/n$ y la varianza muestral $S^2 = ((X_1 - \bar{X})^2 + \dots + (X_n - \bar{X})^2)/(n - 1)$ son independientes.

Adicionalmente, debemos destacar el caso especial cuando $\mu = 0$ y $\sigma = 1$, ya que, en ese caso, la función de distribución recibe el nombre de función de distribución normal estándar, calculada de la siguiente forma sustituyendo dichos valores en la ecuación 1.1:

$$\Phi_{0,1}(x) = \int_{-\infty}^x \varphi_{0,1}(u) du = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du$$

Finalmente, una de las definiciones que también admite la función de distribución viene dada en términos de la función error:

$$\Phi(x) = \frac{1}{2} \left[1 + \operatorname{erf} \left(\frac{x - \mu}{\sigma\sqrt{2}} \right) \right]$$

El complemento de la función de distribución $Q(x)$ definida como $1 - \Phi(x)$ sirve para representar la cola de la probabilidad de la distribución gaussiana y puede venir dada en función del error

$$\Phi^{-1}(p) = \sqrt{2} \operatorname{erf}^{-1}(2p - 1), p \in (0, 1)$$

o, en cambio, expresarse como:

$$\Phi_{\mu, \sigma^2}^{-1}(p) = \mu + \sigma \Phi^{-1}(p) = \mu + \sigma \sqrt{2} \operatorname{erf}^{-1}(2p - 1), p \in (0, 1)$$

Estandarización

Como consecuencia de las propiedades de la distribución normal previamente mencionadas, tenemos que es posible relacionar cualquier variable aleatoria normal con la distribución normal estándar de la siguiente forma.

Dada una variable aleatoria $X \sim N(\mu, \sigma^2)$, entonces

$$Z = \frac{X - \mu}{\sigma}$$

es una variable aleatoria normal estándar, es decir, $Z \sim N(0, 1)$.

Esta transformación, que recibe también los nombres de normalización ó tipificación, sirve de gran ayuda dado que los valores de la función de distribución de la normal estándar están tabulados y de esta forma podemos usar dichos resultados para hallar valores de la función de distribución de variables no estandarizadas a través transformaciones simples.

1.3. Teorema Central del Límite

Consideremos en primer lugar n variables aleatorias $X_i, i = 1, \dots, n$ idénticamente distribuidas e independientes de forma que $E[X_i] = \mu$ y $Var[X_i] = \sigma^2 < \infty \forall i = 1, \dots, n$.

Podemos definir entonces la variable aleatoria S_n como la suma de esas n variables aleatorias, es decir:

$$S_n := X_1 + \dots + X_n = \sum_{i=1}^n X_i$$

Estandarizando S_n obtenemos una nueva variable definida como:

$$Z_n := \frac{S_n - n\mu}{\sigma\sqrt{n}} = \frac{\sum_{i=1}^n X_i - n\mu}{\sigma\sqrt{n}} \quad (1.2)$$

cuya media es igual a 0 y cuya desviación estandar será igual a 1.

Estamos pues en disposición de formalizar el enunciado de dicho teorema.

Teorema 1.1 (Central del Límite). *Sean n variables aleatorias $X_i, i = 1, \dots, n$ independientes e idénticamente distribuidas con $E[X_i] = \mu$ y $Var[X_i] = \sigma^2 < \infty \forall i = 1, \dots, n$ y sea Z_n la distribución definida en 1.2.*

Entonces, la función de distribución de Z_n converge hacia la función de distribución normal estándar cuando $n \rightarrow \infty$, es decir,

$$\lim_{n \rightarrow \infty} P(Z_n \leq z) = \Phi(z) = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

Otra versión que admite el enunciado de este teorema es el que adopta la versión estandarizada de Z_n en función de la media muestral \bar{X} , es decir

$$Z_n = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

Propiedades del Teorema Central del Límite

- Cuando n es suficientemente grande, el *Teorema Central del Límite* nos garantiza una distribución aproximadamente normal.
- Para asegurar la convergencia, es suficiente con que las variables que se suman sean independientes, idénticamente distribuidas, con valor esperado y varianzas finitas.
- La aproximación entre las dos distribuciones es, en general, mayor en el centro de las mismas que en sus extremos o colas.

1.4. Estadística inferencial

En numerosas ocasiones, precisamos de parámetros de la normal para realizar métodos inferencia estadística que nos permitan estimar intervalos de confianza para parámetros como la media o contrastes de hipótesis para valores propuestos para la misma.

Dada una muestra de tamaño n , con media muestral \bar{X} y desviación estándar muestral s , podemos calcular los límites del intervalo de confianza de la media poblacional para el nivel de confianza α elegido, y el p-valor del contraste de hipótesis sobre la hipótesis nula $H_0 : \mu = \mu_0$, siendo μ_0 el valor propuesto para la media poblacional.

El estadístico de contraste sobre la media seguirá una distribución t-Student con $n - 1$ grados de libertad por lo que, en virtud del Teorema Central del Límite, la distribución de este estadístico tiende a la normal (es decir, será aproximadamente normal cuando el tamaño de la muestra sea suficientemente grande).

De la misma forma, se puede realizar un proceso similar para estimar proporciones de un suceso que aconteció en la muestra.

A la hora de comparar parámetros de dos muestras, también encontraremos que para realizar pruebas tipo la prueba t, precisaremos suponer que ambas muestras proceden de poblaciones normalmente distribuidas o, en el caso de la comparación de proporciones de un suceso en dos muestras, también apreciará la normal en la distribución que sigue el estadístico empleado en dicha prueba.

1.4.1. Estimación de parámetros de máxima verosimilitud

La estimación por máxima verosimilitud, conocida también como EMV y, en ocasiones, MLE por sus siglas en inglés, es un método habitual para ajustar un modelo y estimar sus parámetros.

Así pues, supongamos X_1, \dots, X_n variables normalmente distribuidas e independientes con media μ y varianza $\sigma^2 > 0$. Tenemos por tanto que los valores observados de estas n variables aleatorias constituyen una muestra de tamaño n de una población normalmente distribuida. Intentaremos estimar la media poblacional μ y la desviación típica poblacional σ , basándose en los valores observados de esta muestra.

La función de densidad conjunta de estas n variables aleatorias independientes es

$$f(x_1, \dots, x_n; \mu, \sigma) = \prod_{i=1}^n \varphi_{\mu, \sigma^2}(x_i) = \frac{1}{(\sigma\sqrt{2\pi})^n} \prod_{i=1}^n \exp\left(-\frac{1}{2}\left(\frac{x_i - \mu}{\sigma}\right)^2\right), (x_1, \dots, x_n) \in \mathbb{R}^n.$$

La función de verosimilitud basada en las observaciones X_1, \dots, X_n es

$$L(\mu, \sigma) = \frac{C}{\sigma^n} \exp\left(-\frac{\sum_{i=1}^n (X_i - \mu)^2}{2\sigma^2}\right), \mu \in \mathbb{R}, \sigma > 0,$$

con alguna constante $C > 0$.

En el método de máxima verosimilitud, los valores de μ y σ que maximizan la función de verosimilitud se toman como estimadores de los parámetros poblacionales μ y σ .

Normalmente, a la hora de maximizar una función de dos variables consideraríamos las derivadas parciales, pero en este caso, el valor de μ que maximiza la función de verosimilitud con σ fijo no depende de σ .

Sin embargo, tenemos que ese valor de μ se sustituye por μ en la función de verosimilitud con lo que finalmente encontramos el valor de σ que maximiza la expresión resultante. Veámoslo:

En primer lugar observamos que la función de verosimilitud es una función decreciente de $\sum_{i=1}^n (X_i - \mu)^2$ por lo que queremos obtener el valor de μ que minimice esta suma. Sea por tanto $\bar{X}_n = (X_1 + \dots + X_n)/n$ la media muestral basada en las n observaciones. Observamos que

$$\sum_{i=1}^n (X_i - \mu)^2 = \sum_{i=1}^n (X_i - \bar{X}_n)^2 + n(\bar{X}_n - \mu)^2,$$

donde tan sólo el último elemento depende de μ y se minimiza por $\hat{\mu}_n = \bar{X}_n$.

Esta es, por lo tanto, la estimación de máxima verosimilitud de μ basada en las n observaciones X_1, \dots, X_n . Cuando sustituimos esta estimación por μ en la función de verosimilitud, obtenemos

$$L(\bar{X}_n, \sigma) = \frac{C}{\sigma^n} \exp\left(-\frac{\sum_{i=1}^n (X_i - \bar{X}_n)^2}{2\sigma^2}\right), \sigma > 0$$

Para obtener el estimador de máxima verosimilitud tanto de σ^2 como de σ debemos calcular la log-función de verosimilitud, esto es, el logaritmo de la función de verosimilitud, que denotaremos como ℓ :

$$\ell(\bar{X}_n, \sigma) = \log C - n \log \sigma - \frac{\sum_{i=1}^n (X_i - \bar{X}_n)^2}{2\sigma^2}, \sigma > 0,$$

de donde se obtiene

$$\frac{\partial}{\partial \mu} \ell(\bar{X}_n, \sigma) = -\frac{n}{\sigma^3} \left(\sigma^2 - \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2 \right), \sigma > 0.$$

Observamos que esta derivada tendrá valor positivo, cero o negativo según σ^2 esté entre 0 y $\hat{\sigma}_n^2 := \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2$, sea igual a dicha cantidad o sea mayor que esa cantidad respectivamente.

Esta media de cuadrados de residuos es pues el estimador de máxima verosimilitud de σ^2 , y su raíz cuadrada es el estimador de máxima verosimilitud de σ basado en las n observaciones.

Observemos finalmente que si hay solamente una observación, es decir, que $n=1$ o $X_1 = \dots = X_n$, lo cual solo ocurre con probabilidad cero, entonces $\hat{\sigma}_n^2 = 0$ lo que refleja el hecho de que en estos casos la función de verosimilitud es ilimitada cuando σ decrece hasta cero.

Capítulo 2

El test Chi Cuadrado de Pearson

La prueba Chi-cuadrado, también llamada Ji cuadrado (χ^2), fue publicada en el año 1900 por el matemático británico Karl Pearson y se considera una prueba no paramétrica que mide la discrepancia entre una distribución observada y otra teórica (bondad de ajuste), indicando en qué medida las diferencias existentes entre ambas, de haberlas, se deben al azar en el contraste de hipótesis.

Una de las mayores ventajas de este test es que puede ser utilizado para trabajar tanto con distribuciones discretas como continuas como, por ejemplo, la distribución normal.

Para una correcta realización de la prueba necesitaremos unos pequeños requerimientos tal y como veremos más adelante en profundidad:

- Necesitaremos que los datos se encuentren agrupados por categorías ó clases, es decir, que si nuestros datos no se encuentran agrupados originalmente, deberemos construir una tabla de frecuencias o histograma.
- También necesitaremos que el tamaño muestral sea suficientemente grande para que la prueba sea válida, lo cual supone una desventaja frente a otro tipo de test.

2.1. Contraste de una hipótesis nula simple

Sea por lo tanto una muestra aleatoria simple X_1, \dots, X_n donde todas las X_i siguen una misma distribución F y todas son observaciones independientes.

Si la distribución F se corresponde a una variable aleatoria discreta, denotemos A_1, \dots, A_k sus valores posibles. La idea del test consiste en realizar un recuento de las frecuencias de cada A_i en la muestra, y compararlas con las frecuencias que cabría esperar si fuera cierta la hipótesis nula, esto es, si la distribución fuera F_0 .

En el caso de que F sea una distribución continua, se divide la recta real en k conjuntos que denotaremos de la misma forma como A_1, \dots, A_k . De esta forma, podemos calcular las frecuencias de observaciones que pertenecen a cada A_i , y llevar a cabo el test igual que en el caso discreto, pues este test trabaja únicamente con k valores y sus frecuencias observadas y esperadas bajo F_0 .

Por tanto, consideremos N_j el número de observaciones muestrales que se encuentran en A_j . Suponiendo que la distribución fuese F_0 , podemos calcular la probabilidad de cada A_j bajo esta distribución, la cual denotaremos como p_j^0 . De ser cierta la hipótesis nula $H_0 : F = F_0$, cada N_j tendría una distribución binomial con n intentos y probabilidad de éxito p_j^0 .

Además, también bajo H_0 , se tendría $E(N_j) = np_j^0$, de modo que el número de observaciones que cabría esperar en el conjunto A_j sería np_j^0 por lo que denotaremos frecuencias observadas a los N_1, \dots, N_k y frecuencias esperadas a np_1^0, \dots, np_k^0 .

Como estadístico de contraste usaremos el siguiente:

$$D = \sum_{j=1}^k \frac{(N_j - np_j^0)^2}{np_j^0}$$

que usaremos para rechazar H_0 si $D > c$ siendo c cierta constante, aceptando la hipótesis nula en caso contrario.

Dicho estadístico se aproxima a una distribución Chi-cuadrado con $(k - 1)$ grados de libertad cuando el tamaño muestral $n \rightarrow \infty$ por lo que la constante c que considerábamos antes vendrá dada por $c = \chi_{k-1, \alpha}^2$ y, por ende, el test consistirá en:

$$\text{Rechazar } H_0 : F = F_0 \text{ si } D = \sum_{j=1}^k \frac{(N_j - np_j^0)^2}{np_j^0} > \chi_{k-1, \alpha}^2$$

Además, en el test chi-cuadrado, si la distribución es continua la forma en la que escogemos los conjuntos A_1, \dots, A_k pueden suponer diferencias importantes por lo que a la hora de formar dichos conjuntos debemos tener en cuenta que: [3]

- Para que la aproximación de la distribución por la chi-cuadrado sea buena, es conveniente que las frecuencias esperadas sean grandes por lo que es recomendable que

$$np_j^0 \geq 5, \forall j \in 1, \dots, k$$

- Para poder distinguir si la distribución verdadera es F_0 , conviene que se divida en muchos conjuntos, por lo que se aconseja que $k \geq 5$

Observamos por lo tanto que ambas condiciones implican que el tamaño muestral n sea superior a 25, tal y como comentábamos al inicio del capítulo. Además, generalmente se respeta la primera condición, ya que sin ella la aproximación de la distribución no sería válida mientras que la segunda condición únicamente es utilizada para obtener resultados significativos ya que no afecta a la validez de la prueba por lo que usualmente se toma $k = \sqrt{n}$

2.2. Contraste de un modelo paramétrico

También se puede considerar una hipótesis nula del tipo

$$H_0 : F \in F_\theta / \theta \in \Theta$$

donde se plantea si la distribución pertenece al modelo paramétrico representado por F_θ donde θ representa los parámetros del modelo.

Podemos considerar de la misma forma que para el caso de hipótesis nula simple las mismas observaciones independientes con la misma distribución F con $A_i, i = 1, \dots, k$ la división de los conjuntos en el caso continuo y N_j la frecuencia de A_j .

Considerando el estimador $\hat{\theta}$ del parámetro θ podremos calcular las probabilidades $p_j(\hat{\theta})$ de cada A_j bajo la distribución $F_{\hat{\theta}}$ y con frecuencias esperadas $np_j(\hat{\theta}), i = 1, \dots, k$, por lo que podremos calcular el estadístico chi-cuadrado de la misma forma, con la diferencia de que ahora su aproximación será una distribución chi-cuadrado pero con menos grados de libertad como consecuencia de de la estimación de los parámetros.

Por lo tanto,

$$D = \sum_{j=1}^k \frac{(N_j - np_j(\hat{\theta}))^2}{np_j(\hat{\theta})} \approx \chi_{k-q-1}^2$$

donde q será el número de parámetros estimados que en el caso de la normal serán 2 si estimamos la media y la varianza.

Con la intención de entender mejor el funcionamiento del estadístico chi-cuadrado, presentamos el siguiente ejemplo realizado en R , en el que se generan 30 números aleatorios que siguen una distribución normal y después se realiza la prueba, primero bajo hipótesis de que los datos pertenecen a una normal estándar, luego realizando una estimación de la media y varianza y, por último, empleando la función `pearson.test()`, cuyo funcionamiento se explica en profundidad en el anexo, pues también será la utilizada en las simulaciones del Capítulo 7.

Por lo tanto, veamos en el siguiente código las diferencias de la implementación del estadístico y el resultado proporcionado por la función.

```
> set.seed(123456)
> n=30
> x=rnorm(n)
> nclas=5
>
> # Test de N(0,1)
> ext=qnorm((0:nclas)/nclas)
> xag=cut(x,breaks=ext)
>
> obs=table(xag)
> esp=n/nclas
>
> est=sum((obs-esp)^2/esp);est
[1] 5.666667
> niv=1-pchisq(est,df=nclas-1);niv
[1] 0.2254631
>
>
> # Test de Normalidad
> med=mean(x)
> dt=sd(x)
> ext=qnorm((0:nclas)/nclas,mean=med,sd=dt)
> xag=cut(x,breaks=ext)
>
> obs=table(xag)
> esp=n/nclas
>
> est=sum((obs-esp)^2/esp);est
[1] 2.666667
> niv=1-pchisq(est,df=nclas-3);niv
[1] 0.2635971
>
> # Con el paquete nortest
> library(nortest)
> pearson.test(x,n.classes=5)
```

```
Pearson chi-square normality test
```

```
data: x
```

```
P = 2.6667, p-value = 0.2636
```

Como podemos observar, el p-valor y el valor del estadístico coinciden cuando estimamos media y varianza y al usar *pearson.test()*, mientras que el resultado varía al suponer la normalidad de la muestra.

Capítulo 3

Test de normalidad basados en la función de distribución

En este capítulo nos ocuparemos de los principales test de normalidad que se basan en la función de distribución, comenzando por el test de Kolmogorov-Smirnov y su variante de Lilliefors, seguido de las pruebas de Cramer y Anderson y para finalizar una breve sección donde se muestra la relación entre ellos.

3.1. Test de Kolmogorov-Smirnov y versión de Lilliefors

En primer lugar, comenzaremos describiendo la prueba que, más tarde, el profesor de universidad Hubert Lilliefors, usó como base para introducir el test que lleva su nombre.

3.1.1. Prueba de Kolmogorov-Smirnov

La prueba de Kolmogórov-Smirnov (también conocida como prueba K-S), cuyo nombre proviene de los matemáticos rusos Andrey Kolmogorov y Nikolai Smirnov, es una prueba no paramétrica que determina la bondad de ajuste de dos distribuciones de probabilidad entre sí.

Podemos plantearnos un problema de contraste sobre la distribución de una variable aleatoria absolutamente continua. Empezamos con un contraste donde la hipótesis nula sea simple, es decir,

$$H_0 : F = F_0$$

siendo F_0 una distribución completamente específica. Se trata de ver si la distribución coincide con cierta función, F_0 , o no. Para realizar el contraste emplearemos el estadístico

$$D_n = \sup_{x \in \mathbb{R}} |F_n(x) - F_0(x)|$$

de forma que cuando D_n sea grande rechazaremos la hipótesis $H_0 : F = F_0$. [4]

Para determinar la distribución del estadístico de Kolmogorov-Smirnov, supondremos que la distribución verdadera es F , sin necesidad de centrarse en el valor F_0 de un problema de contraste de hipótesis.

Además podemos observar que:

$$D_n = \sup_{x \in \mathbb{R}} |F_n(x) - F_0(x)| = \max(D_n^+, D_n^-)$$

donde $D_n^+ = \sup_{x \in \mathbb{R}} (F_n(x) - F(x))$ y $D_n^- = \sup_{x \in \mathbb{R}} (F(x) - F_n(x))$

Algunas ventajas que presenta esta prueba son:

- Es fácil de calcular y usar, y no requiere agrupación de los datos.
- El estadístico es independiente de la distribución de frecuencias esperada, solo depende del tamaño de la muestra.

Visualizaremos mejor la realización de dicha prueba con el siguiente ejemplo:

Ejemplo: Sean las mediciones de la altura de un tipo de planta las siguientes:

48,1, 47,8, 45,1, 46,3, 45,4, 47,2, 46,6, 46

Sabiendo que tanto media y desviación típica son conocidas ($\mu = 40$ y $\sigma = 3$), veamos si nuestra muestra se adecúa a una distribución normal

Solución: Estandarizamos los valores medidos y calculamos los valores de D_n

X_n	Z_n	$F_n(X)$	$F_0(X)$	$ F_n(X) - F_0(X) $
48,1	2,7	0,9554	0,125	0,8304
47,8	2,6	0,9641	0,250	0,7141
45,1	1,7	0,9772	0,375	0,6022
46,3	2,1	0,9821	0,500	0,4821
45,4	1,8	0,9861	0,625	0,3611
47,2	2,4	0,9918	0,750	0,2418
46,6	2,2	0,9953	0,875	0,1203
46	2	0,9965	1,000	0,0035

Como podemos observar, tenemos que el valor de $D_n = 0,8304$ y, si observamos los valores tabulados para $\alpha = 0,05$ y $n = 8$ obtenemos que $D_n = 0,8304 > 0,45427$ por lo que deberemos rechazar la hipótesis de la normalidad.

3.1.2. Test de Lilliefors

Para proceder con el test de Lilliefors debemos en primer lugar estimar la media y la varianza de la población en base a nuestros datos.

El estadístico del test de Lilliefors es calculado de la misma forma que en la prueba de Kolmogorov-Smirnov, con la única modificación que en este caso calcularemos la máxima diferencia entre la función de distribución empírica y la función de distribución acumulativa (FDC) de la distribución normal con la media y varianza estimadas.

Por lo tanto, consideramos la distancia entre la distribución empírica (estimador no paramétrico de F) y el estimador paramétrico $F_{\hat{\theta}}$. [5]

$$\sup_{x \in \mathbb{R}} |F_n(x) - F_{\hat{\theta}}(x)|$$

Como hemos visto, ambos test son muy similares, pero una de las principales diferencias reside en las tablas de los valores críticos para ambas pruebas, que serán utilizadas para determinar si se rechaza o no la hipótesis nula (H_0)

3.2. Test de Cramer

El test de Cramér-von Mises fue introducido por los investigadores Harald Cramér y Richard Edler von Mises se emplea para juzgar la bondad de una función de distribución

acumulada F_n comparada con una función de distribución empírica F_0 , o para comparar dos distribuciones empíricas.

El estimador de Cramér von Mises acepta o rechaza la hipótesis nula H_0 : “Los datos provienen de una distribución conocida F_0 ”, mediante el cálculo de la distancia entre las dos funciones, de modo que H_0 debe aceptarse para valores de CVM_n pequeños:

$$CVM_n = n \int_{-\infty}^{\infty} [F_n(x) - F_0(x)]^2 dF_0(x) \quad (3.1)$$

donde n denota el número de datos.

De forma alternativa, se puede demostrar que el estadístico se puede escribir de la siguiente forma:

$$T = CVM_n = \frac{1}{12n} + \sum_{i=1}^n \left(\frac{2i-1}{2n} - F_0(X_i) \right)^2$$

Test de Watson

Una pequeña modificación del anterior estadístico viene dada por este resultado, el cual emplea el estadístico: [6]

$$U^2 = T - n\left(\bar{F} - \frac{1}{2}\right)^2$$

, donde

$$\bar{F} = \frac{1}{n} \sum_{i=1}^n F(X_i)$$

3.3. Test de Anderson-Darling

El test de Anderson-Darling, nombrado así por *Theodore Wilbur Anderson* y *Donald A. Darling*, sirve para comprobar si una muestra procedente de una serie de datos proviene de una distribución determinada.

Esta prueba se basa en la comparación de la distribución de probabilidades acumulada empírica (resultado de los datos) con la distribución de probabilidades acumulada teórica (definida en H_0).

Así pues, y de la misma forma que en los test anteriores, nuestras hipótesis serán que la muestra sigue una distribución de probabilidad específica (H_0) o, si por el contrario, no lo hace (H_1). En nuestro caso, usaremos esta prueba para comprobar si la muestra sigue una distribución normal.

Este test, que es una variación de la *prueba de Kolmogorov-Smirnov* explicada anteriormente, le da más peso a las colas de lo que lo hace el dicho test K-S.

En el test de Kolmogorov-Smirnoff los valores críticos no dependen de la distribución específica a la que estamos sometiendo el test, mientras que el de Anderson-Darling sí emplea dicha distribución a la hora de calcular los valores críticos.

Como consecuencia, el test de Anderson-Darling tiene la ventaja de tener una mayor sensibilidad mientras que, por otro lado, los valores críticos variarán dependiendo de la distribución que hayamos escogido, por lo que deberán ser calculados para cada distribución, aunque hoy en día los valores críticos están tabulados para algunas distribuciones conocidas como la normal, la uniforme o la exponencial, entre otras.

El estadístico de Anderson-Darling está definido como:

$$A^2 = -n - S \quad (3.2)$$

donde

$$S = \sum_{i=1}^n \frac{(2i-1)}{n} [\ln(F(Y_i)) + \ln(1 - F(Y_{n+1-i}))]$$

y donde F es la función de distribución acumulativa de distribución específica que estamos tratando de comprobar que siguen nuestros datos.[7]

También debemos destacar que los datos de nuestra muestra (Y_i) deben estar ordenados

Por tanto, rechazaremos la hipótesis de que nuestra muestra sigue una determinada distribución si el valor del estadístico 3.2 es mayor que el valor crítico que, recordemos, depende de la distribución.

3.4. Relación Kolmogorov-Cramer-Anderson

Puesto que los tres test presentados en este capítulo se basan en la función de distribución, prestaremos atención a las similitudes de ambos test, así como un breve repaso a sus principales diferencias.

En primer lugar, debemos remontarnos a 1928 cuando Cramér propuso cuantificar la evidencia en contra de la hipótesis nula de que la función de distribución de nuestra muestra seguía una función de distribución determinada ($H_0 : F = F_0$).

Para dichos cálculos, usó la siguiente cantidad:

$$\int_{\mathbb{R}} [F_n(x) - F_0(x)]^2 dK(x)$$

donde K es una función no negativa. Esta cantidad le sirvió para medir la discrepancia entre F_n y F_0 , ponderadas por $K(x)$. Cramér, correctamente, siguió el criterio de rechazar H_0 cuando esta discrepancia era muy grande.

A la vez que Cramér, von Mises llegó a resultados similares aunque de manera independiente.

Más tarde, al rededor de 1937, Smirnov consiguió generalizar el criterio de Cramér y von Mises como: [8]

$$W_n^2 = n \int_{\mathbb{R}} [F_n(x) - F_0(x)]^2 \psi(x) dF_0(x) \quad (3.3)$$

donde ψ era una función que ponderaba las diferencias entre las distribuciones F_n y F_0 .

La expresión anterior se trata de un estadístico, pues F_n está definida a partir de la muestra aleatoria y puede denotarse también como:

$$W_n^2 = E_{F_0}[\psi(X)(F_n - F(X))^2]$$

donde E_{F_0} significa que la esperanza se calcula suponiendo cierta H_0 , es decir, que la distribución de X es F_0 . W_n^2 cuantifica el valor esperado de la diferencia cuadrática entre F_n y F_0 ponderada por $\psi(x)$.

Si escogiésemos $\psi(x) = \psi(F_0(x))$, es decir que la función que pondera las diferencias dependa de $F_0(x)$ entonces W_n^2 es libre de distribución bajo H_0 .

Así pues, sustituyendo $\psi(x) = 1$ en 3.3 obtendríamos el estadístico de Cramér-von Mises 3.1.

En cambio, Anderson y Darling propusieron utilizar para su estadístico la siguiente función:

$$\psi(x) = \frac{1}{F_0(x)(1 - F_0(x))}$$

Esta elección de ψ es lo que implica que las diferencias en las colas de la distribución tienen mayor peso. Además, y de la misma forma que en el caso anterior, basta con sustituir la definición de Anderson y Darling de $\psi(x)$ en 3.3 para obtener el estadístico de su test.

Pese a que los test de Cramér-von Mises y Anderson -Darling guardan una gran similitud con respecto al de Kolmogorov-Smirnov (K-S) y, por ende, al de Lilliefors, una de las principales diferencias que presentan es que en el test K-S los valores críticos son independientes de la distribución empírica escogida mientras que para las otras dos pruebas no lo son.

Esta última diferencia implica que a la hora de realizar los tests, podamos tener tabulados los valores críticos del test K-S mientras que no podremos contar con esa ventaja en el resto de pruebas, pues dichos valores variarán según las distintas distribuciones.

Capítulo 4

Test de normalidad basados en los momentos

En este capítulo definiremos y analizaremos en profundidad los conceptos de asimetría y curtosis y su influencia en la distribución normal. Brevemente repasaremos los test de normalidad basados en estos dos parámetros para finalmente estudiar el test de Jarque-Bera que combina ambos conceptos.

4.1. Asimetría

Veamos qué es la asimetría estadística junto con algunos valores que la definen.

4.1.1. Definición

Las medidas de asimetría permiten medir el grado de simetría (o asimetría) que presenta una distribución de probabilidad de una variable aleatoria sin la necesidad de realizar su representación gráfica.

Puesto que la simetría se mide en torno a un eje de simetría, consideramos como tal una recta paralela al eje de ordenadas que pasa por la media de la distribución. Por lo tanto, definimos a una distribución como simétrica si existe el mismo número de valores a la derecha que a la izquierda de la media y, por lo tanto, tiene el mismo número de desviaciones con signo positivo que con signo negativo.

En cambio, diremos que hay asimetría positiva (o a la derecha) si la cola a la derecha de la media es más larga que la de la izquierda, es decir, si hay valores más separados de la media a la derecha. Del mismo modo, afirmaremos que hay asimetría negativa (o a la izquierda) si la cola a la izquierda de la media es más larga que la de la derecha, es decir,

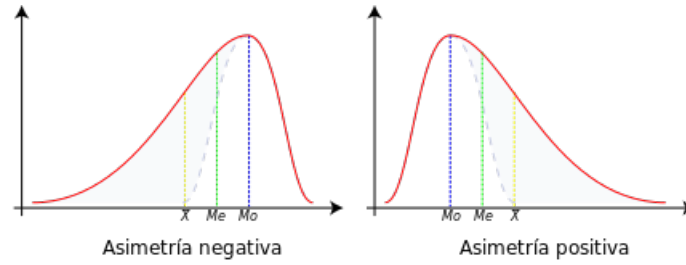


Figura 4.1: Representación gráfica de ambos tipos de asimetría.

si hay valores más separados de la media a la izquierda.

4.1.2. Medidas de asimetría

A continuación presentamos algunas de las medidas de asimetría más conocidas y utilizadas.

Coefficiente de asimetría de Fisher

Esta medida, que es una de las más empleadas, está basada en el tercer momento estándar dado que nos interesa mantener el signo de las desviaciones con respecto a la media, para obtener si son mayores las que ocurren a la derecha de la media que las de la izquierda.

Aún así no se debe tomar el momento estándar con respecto a la media de orden 1 debido a que una simple suma de todas las desviaciones siempre es cero. Como ejemplo de esto tenemos que si los datos están agrupados en k clases obtenemos que: [9]

$$\sum_{i=1}^k f_i(x_i - \mu) = \sum_{i=1}^k f_i x_i - \mu \sum_{i=1}^k f_i = \mu - \mu = 0$$

donde x_i representa la marca de la clase i -ésima y f_i es la frecuencia relativa de dicha clase. Por lo tanto, se toman las desviaciones al cubo y el coeficiente de asimetría de Fisher queda definido como:

$$S = \gamma_1 = \frac{\mu_3}{\sigma^3} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^3}{\left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2\right)^{3/2}} \quad (4.1)$$

donde μ_3 es el tercer momento entorno a la media y σ la desviación típica.

Así pues los valores de γ_1 determinan los distintos tipos de asimetría que sufre nuestra distribución:

- $S > 0$: la distribución es asimétrica positiva o a la derecha.
- $S < 0$: la distribución es asimétrica negativa o a la izquierda.

Por otro lado, sabemos que si la distribución es simétrica entonces $S = 0$ pero el recíproco es falso, es decir, si $S = 0$ no podemos afirmar que la distribución es simétrica.

Coefficiente de asimetría de Pearson

El coeficiente de asimetría de Pearson está basado en que en aquellas distribuciones que son simétricas, la media de la distribución coincide con la moda.

Dicho coeficiente sólo se debe utilizar en distribuciones uniformes, unimodales y moderadamente asimétricas y se define como:

$$A_p = \frac{\mu - moda}{\mu}$$

donde μ es el momento ordinario de orden 1, es decir la media.

Por lo tanto, si estamos ante una distribución simétrica la media coincidirá con la moda y, consecuentemente, $A_p = 0$. En cambio si la distribución tiene asimetría positiva, el valor de la media estará por encima de la moda, lo que significa que el valor de $A_p > 0$.

Coefficiente de asimetría de Bowley-Yule

Este coeficiente se basa en la posición de los cuartiles y la mediana y está definido como:

$$A_{BY} = \frac{Q_{3/4} + Q_{1/4} - 2Me}{Q_{3/4} - Q_{1/4}}$$

Ahora bien, si tenemos en cuenta que el primer y el tercer cuartil están a la misma distancia de la media si la distribución es simétrica, obtenemos que, en caso de simetría, $A_{BY} = 0$, mientras que si la asimetría es positiva, $A_{BY} > 0$.

4.1.3. Asimetría en la distribución normal

Una vez definidos los distintos coeficientes de asimetría veamos que el valor de éste cuando nos encontramos en una distribución es exactamente 0.

Para simplificar los cálculos y pues cualquier tipo de distribución normal tendrá un coeficiente de asimetría igual a 0 usaremos la distribución normal estándar.

Proposición 4.1. *El coeficiente de asimetría de la distribución normal estándar es 0*

Demostración. En primer lugar, sea $X \sim N(0, 1)$. Tenemos que en el caso de la distribución normal estándar, la desviación típica tiene valor 1 y la media 0, por lo que el coeficiente de asimetría quedará de la siguiente forma:

$$S = \mu_3(X) = E[(X - E[X])^3] = E[X^3] = \alpha_3(X)$$

Así pues, nos bastará con calcular el momento de tercer orden de la distribución normal en torno a la media, es decir, el valor de la tercera derivada de la Función Generatriz de Momentos de X para $t = 0$. Por lo tanto y volviendo a que nos encontramos trabajando con la distribución estándar, tenemos que $\varphi_X(t) = e^{\frac{t^2}{2}}$.

De donde la primera derivada será:

$$\varphi'_X(t) = \frac{2t}{2} 2t\varphi_X(t) = t\varphi_X(t)$$

, y puesto que estamos centrándonos en $t = 0$,

$$\varphi'_X(t=0) = 0$$

Para la segunda derivada tenemos que

$$\varphi''_X(t) = 1\varphi_X(t) + t\varphi'_X(t) = \varphi_X(t) + t^2\varphi_X(t) = (1 + t^2)\varphi_X(t)$$

por lo que, haciendo $t = 0$ obtenemos que $\varphi''_X(t=0) = 1$ para la tercera derivada:

$$\varphi'''_X(t) = 2t\varphi_X(t) + (1 + t^2)\varphi'_X(t) = (3t + t^3)\varphi_X(t)$$

donde nuevamente haciendo $t=0$,

$$\varphi'''_X(t) = 0$$

por lo tanto,

$$\mu_3(X) = \alpha_3(X) = S(X) = 0$$

□

4.1.4. Test de normalidad basado en la asimetría

Una vez definido el concepto de asimetría, podremos usarlo para determinar si una variable aleatoria se ajusta a una distribución normal o no.

Puesto que la distribución gaussiana es simétrica en torno a la media y que una de las formas más habituales de desviarse respecto de la normalidad es por falta de simetría, podemos construir un método de contraste basándonos en este concepto.

Por lo tanto, si la distribución de los datos es normal, el coeficiente de asimetría seguirá una distribución asintótica normal de media igual a cero y varianza $\frac{6}{n}$, por lo que emplearemos como estadístico de contraste el siguiente:

$$\sqrt{\frac{n}{6}}S \sim N(0, 1)$$

Por lo tanto, se rechazará la hipótesis nula de normalidad cuando el valor del estadístico anterior sea muy grande (positiva o negativamente), en comparación con los cuantiles de la normal estándar.

4.2. Curtosis

4.2.1. Definición

Definimos la curtosis de una variable aleatoria como una característica de la forma que adquiere la función de distribución de dicha variable.

Un valor grande en la curtosis significa que existe una gran acumulación de valores de la variable tanto cerca de la media como lejos de ella, a la vez que implica una menor frecuencia de los valores intermedios. Esta característica explica algunas formas de la función de distribución, con colas más gruesas, centro más apuntado y menor proporción de valores entre el pico y las colas.

No debemos caer en el error de asumir que cuanto mayor sea la curtosis, mayor será la varianza o viceversa, pues no es cierto.

4.2.2. Medidas de curtosis

Para calcular un valor de la curtosis empleamos el siguiente indicador:

$$K = \frac{\mu_4}{\sigma^4} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4}{\left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2\right)^2} \quad (4.2)$$

Este coeficiente viene definido por el cuarto momento centrado con respecto a la media (μ_4) y la desviación estándar (σ).

Puesto que en la distribución normal se verifica que $\mu_4 = 3\sigma^4$, existe una segunda definición para el exceso de curtosis:

$$g_2 = \frac{\mu_4}{\sigma^4} - 3 = K - 3$$

de forma que este coeficiente valdrá 0 cuando estemos trabajando con una distribución normal y se tome a esta como referencia para la curtosis.

Así pues, si tomamos como referencia la distribución normal, una distribución podrá ser:[10]

- **Leptocúrtica:** se da cuando $K > 3$ y $g_2 > 0$ y la distribución será más apuntada y con colas más gruesas que la dist. Normal.
- **Platicúrtica:** se da cuando $K < 3$ y $g_2 < 0$ y la distribución será menos apuntada y con colas menos gruesas que la dist. Normal.
- **Mesocúrtica:** se da cuando $K = 3$ y $g_2 = 0$ y la distribución seguirá una distribución normal.

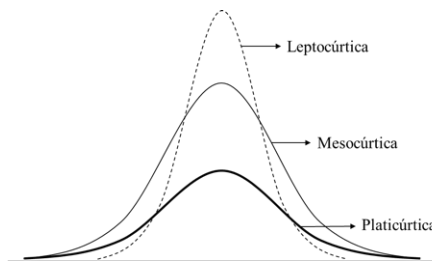


Figura 4.2: Representación gráfica de los tipos de curtosis.

Por último tenemos que existe otra forma de medir la curtosis, examinando dicho parámetro de la suma de variables aleatorias, es decir, suponiendo Y la suma de n variables aleatorias independientes con la misma distribución X , se tendrá:

$$Kurt[Y] = \frac{Kurt[X]}{n}$$

4.2.3. Curtosis en la distribución normal

Del mismo modo que hicimos con la asimetría, centraremos nuestros cálculos en la distribución normal estándar, para comprobar que dicha distribución tiene un coeficiente de curtosis $K = 3$ o, en su caso, $g_2 = 0$.

Proposición 4.2. *El coeficiente de curtosis K de la distribución normal estándar $X \sim N(0, 1)$ es 3, y, por lo tanto, $g_2 = 0$.*

Demostración. El coeficiente de curtosis $K = \frac{\mu_4}{\sigma^4}$ es el momento central de cuarto orden en torno a la media, pero dado que la distribución normal estándar tiene media igual a 0 y desviación típica 1, coincidirá con la cuarta derivada de la Función Generatriz de Momentos para el valor $t = 0$.

Aprovechando los cálculos de la demostración de la Proposición 4.1 tenemos que la cuarta derivada de la Función Generatriz de Momentos será:

$$\varphi_X^{IV}(t) = (3+3t^2)\varphi_X(t) + (3t+t^3)\varphi_X'(t) = (3+3t^2)\varphi_X(t) + (3t+3t^3)t\varphi_X'(t) = (3+6t^2+t^4)\varphi_X(t)$$

donde, para $t = 0$,

$$K = \mu_4(X) = \alpha_4(X) = \varphi_X^{IV}(t=0) = 3$$

De donde, por la propia definición de $g_2 = 0$:

$$g_2 = \frac{\mu_4}{\sigma^4} - 3 = K - 3 = 3 - 3 = 0$$

□

Por lo tanto, y a la vista del resultado anterior, la distribución normal posee la forma tipo en cuanto a aplastamiento o apuntamiento y sirve de modelo para las demás distribuciones por lo que numerosos autores consideran que el valor referencia del coeficiente de curtosis es 3 y no 0.

4.2.4. Test de normalidad basado en la curtosis

Tras definir el concepto de curtosis, podremos usarlo para determinar si una variable aleatoria se ajusta a una distribución normal o no.

Así pues, podremos rechazar la hipótesis de la normalidad si la curtosis de nuestra muestra es mucho mayor o mucho menor que cero, que es el valor que se le asigna a la distribución normal como vimos en la Proposición 4.2.

Además, si la distribución de los datos muestrales es normal, entonces la curtosis tendrá distribución asintótica normal de media cero y varianza $\frac{24}{n}$, por lo que se puede emplear como estadístico de contraste el siguiente:

$$\sqrt{\frac{n}{24}}g_2 \sim N(0, 1)$$

Por lo tanto, se rechazará la hipótesis nula de normalidad cuando el estadístico anterior sea muy grande (tanto en positivo como en negativo), en comparación con los cuantiles de $N(0, 1)$.

4.3. Test de Jarque-Bera

La prueba de Jarque-Bera vió la luz cuando dos economistas de la Australian National University, Carlos M. Jarque y Anil K. Bera, idearon un estadístico para un test de normalidad que empleaba tanto la asimetría, como la curtosis.

Para realizar dicho test, se emplea el siguiente estadístico que, como decimos, tiene como base la asimetría y la curtosis de la muestra:

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

donde S y K son los coeficientes de asimetría y curtosis definidos en 4.1 y 4.2 respectivamente.

Este estadístico se distribuye asintóticamente como una distribución Chi-cuadrado con dos grados de libertad y puede usarse para probar la hipótesis nula de que los datos pertenecen a una distribución normal, lo que implica que la asimetría y el exceso de curtosis son nulos ($S = 0$ y $K = 3$).

La prueba también es utilizada en modelos de regresión para probar la normalidad de los residuos. Para ello se utilizan los residuos estimados obtenidos por mínimos cuadrados. Finalmente, los puntos críticos para muestras pequeñas se pueden calcular vía Monte Carlo. [11]

Como ejemplo del test de Jarque-Bera podemos implementar el estadístico en R y realizar la prueba para una muestra generada aleatoriamente.

Ejemplo: *Para realizar el test de Jarque-Bera definimos una función que nos devuelva el valor crítico del estadístico.*

```
JB=function(x){ med=mean(x)
v=mean((x-med)^2)
S=mean((x-med)^3)/v^1.5
K=mean((x-med)^4)/v^2
estadistico=(n/6)*(S^2+0.25*(K-3)^2)
return(1-pchisq(estadistico,df=2))}
```

Una vez implementado el test, generamos una nuestra muestra aleatoria de tamaño $n=30$ bajo hipótesis de normalidad y comprobamos si el valor que nos devuelve la función nos permite aceptar la hipótesis nula.

```
> x=rnorm(30)
> JB(x)
[1] 0.7517366
```

La función nos devuelve un valor crítico grande, por lo que podemos aceptar la hipótesis nula tal y como preveíamos

Capítulo 5

Métodos gráficos

Menos eficaces que los métodos expuestos previamente, los métodos gráficos también suponen una forma de comprobar si un conjunto de datos se adecúa a una distribución normal. A continuación presentamos los más relevantes.

5.1. Histograma y curva normal

Una forma mucho más informal que el resto de test para comprobar la normalidad consiste en realizar un histograma muestra para luego superponer la curva de la función de distribución con misma media y desviación estándar que los datos a estudiar.

Realizando simulaciones con tamaño muestral $n = 500$ con el software *R* observamos que suponiendo que nuestra distribución es normal, el histograma se ajusta a la función de densidad, pero para distintas distribuciones obtenemos que las diferencias serán bastante evidentes.

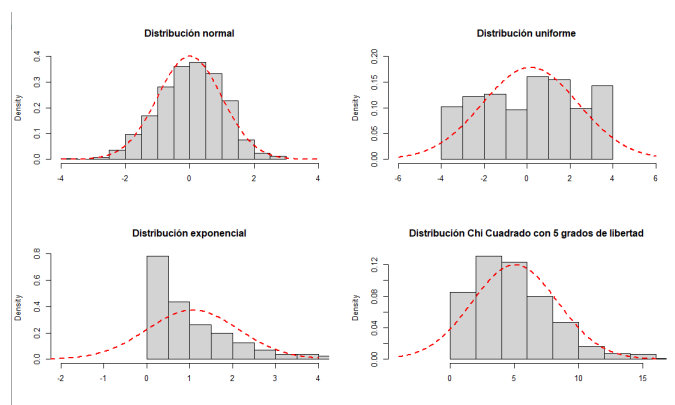


Figura 5.1: Histogramas de distintas distribuciones y función de densidad normal.

Dichas diferencias, son mucho más visibles cuando se trabaja con distribuciones que tienen una sola cola, pero con distribuciones como la binomial que observamos a continuación, podríamos llegar a concluir erróneamente que sigue una normal, pues parece que el histograma se adecúa de forma bastante fiable a la densidad de la normal.

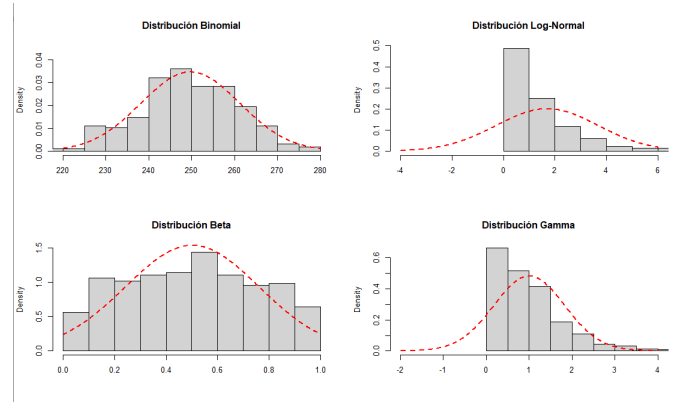


Figura 5.2: Histogramas de distintas distribuciones y función de densidad normal.

Este tipo de test conlleva un gran número de desventajas dada su poca formalidad y más aún si nuestra muestra es de un tamaño reducido ya que, como podemos observar en el ejemplo siguiente, simulando una muestra que sigue una distribución normal con distintos tamaños muestrales, obtenemos que para muestras pequeñas los histogramas no se corresponden con la función de densidad de la normal, pese a que sí siguen dicha distribución. No obstante, cuanto mayor se hace el tamaño de la muestra el histograma va adecuándose mejor a la curva.

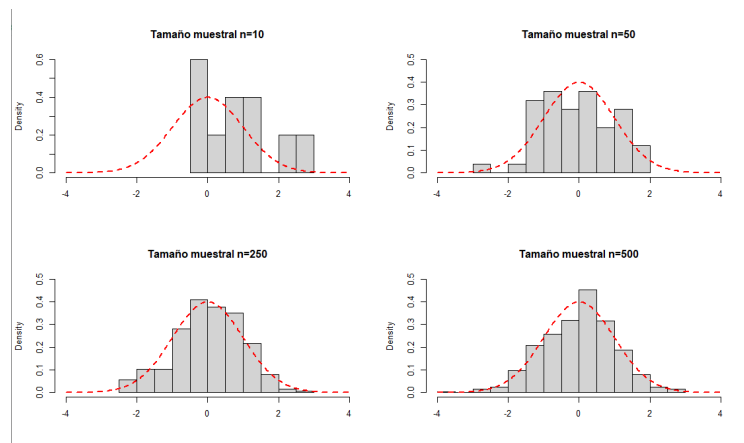


Figura 5.3: Simulación de histogramas.

5.2. Gráfico de cuantiles teóricos

Los gráficos de cuantiles teóricos, más comunmente llamados QQ-plots son un método gráfico para el diagnóstico de las diferencias entre la distribución de una población de la que hemos extraído una muestra aleatoria y una distribución usada para la comparación, que en nuestro caso será la distribución normal.[12]

Para testar la normalidad de una muestra de tamaño n , se dibujan n puntos de la forma (x, y) que se corresponden con los cuantiles de nuestra muestra (coordenada y) y el mismo cuantil de la distribución normal teórica (coordenada x).

Pese a que hay más formas de calcularlos, para los cuantiles de la distribución que usamos como comparación (dist. normal) se emplea habitualmente la fórmula:

$$\frac{k}{n+1}$$

Por tanto, si nuestra muestra se ajusta a una distribución normal, los puntos del QQ-plot quedarán en una línea diagonal que se corresponde con la gráfica de $y = x$ o muy próximos a ella.

Además, los QQ-plots también son muy utilizados para comparar las formas de ambas distribuciones además de proporcionarnos una primera idea de si propiedades como localización, escala o simetría son similares.

Generalmente, los QQ-plot suponen un método más potente que el comparar los histogramas y la función de distribución, pero pueden resultar más complicados de interpretar dado que, en cierto modo, son una forma gráfica de realizar un test de bondad de ajuste, en vez de realizarse de forma numérica.

Algunas de las interpretaciones que podemos extraer de un QQ-plot son:

- Como comentamos, si los puntos siguen la gráfica $y = x$ entonces ambas distribuciones pueden considerarse idénticas y, por tanto, consideraremos nuestra muestra como normal.
- Si los puntos siguen una línea recta pero que no es la gráfica de $y = x$, entonces nuestra distribución puede ser una transformación lineal de la normal.
- Si los puntos siguen una gráfica más plana que $y = x$ entonces los datos de la distribución del eje horizontal estarán más dispersos, es decir que los datos de nuestra muestra estarán mas compactos que los de la normal.
- Si los puntos siguen una forma arqueada o de "S" entonces nuestra muestra tendrá una mayor asimetría que la normal o las colas tendrán mucho más peso.

Por último podemos relacionar los QQ-plots con los gráficos de probabilidad ya que dichos gráficos emplean la esperanza del k -ésimo orden de la distribución frente a los QQ-plots que usan los cuantiles, por lo que solo existe una diferencia importante entre ambos al trabajar con muestras de tamaño n pequeño.

Fueron simuladas tres muestras de tamaño 100 que siguen una distribución normal, normal logarítmica y T-Student con 20 grados de libertad respectivamente y cuyos gráficos son los siguientes:

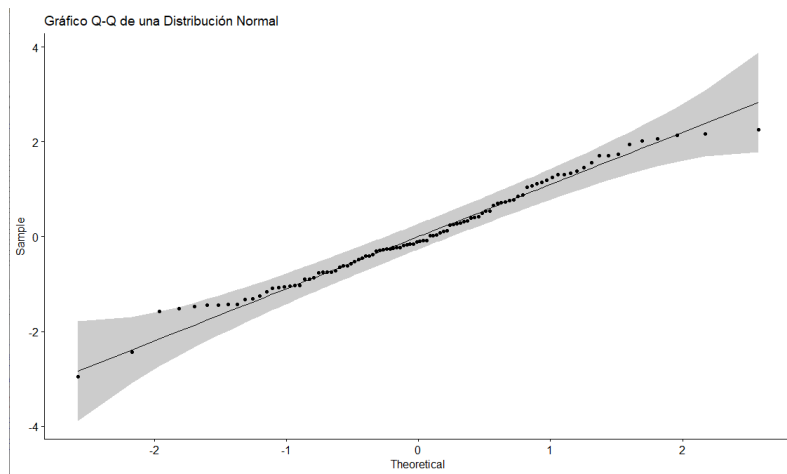


Figura 5.4: Ejemplo de QQ-Plot de una distribución normal.

Como podemos observar, los datos se adecúan y ajustan a la gráfica de $y = x$ dado que, evidentemente, siguen una distribución normal.

A continuación, veremos lo que ocurre con la distribución normal logarítmica, donde cabe esperar que los cuantiles de nuestros datos apenas coincidan con los de la normal, desviándose completamente de la trayectoria de la diagonal.

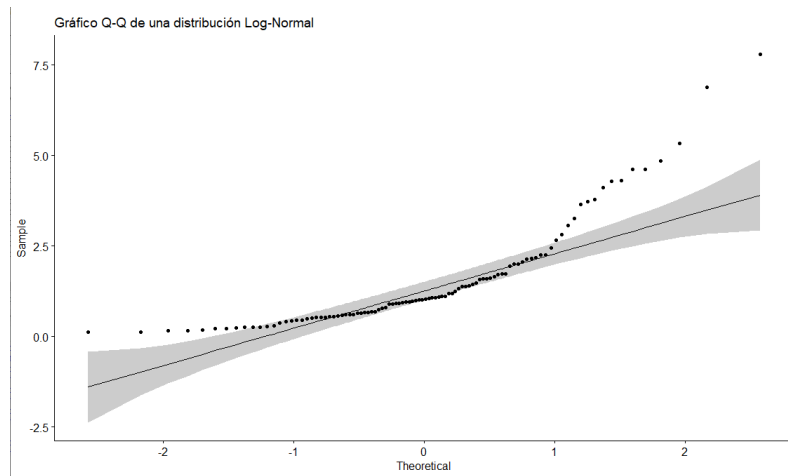


Figura 5.5: Ejemplo de QQ-Plot de una distribución log-normal.

Pero no siempre se obtienen gráficas con un desfase de los cuantiles de la muestra y los de la normal tan grande. Veamos un ejemplo de esta situación con el caso de la distribución T-Student con 20 grados de libertad:

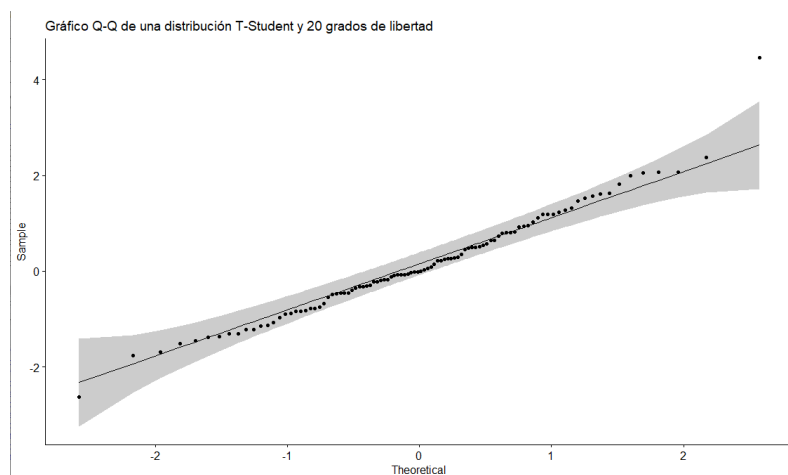


Figura 5.6: Ejemplo de QQ-Plot de una distribución T-Student.

Observamos que en la gráfica los puntos se adecúan bastante bien y siguen la diagonal, lo que puede dar lugar a pensar que nos encontramos ante una muestra que sigue una distribución normal del mismo modo que en la figura 5.4. En cambio, nuestra muestra sigue una distribución T-Student con 20 grados de libertad, con lo que comprobamos que los métodos gráficos pueden ser confusos y dar lugar a más errores que los estudiados previamente.

Capítulo 6

Test de Shapiro-Wilk y variante de Shapiro-Francia

En este capítulo, estudiaremos el test de Shapiro-Wilk junto con su variante de Shapiro-Francia, además de que observaremos las diferencias de resultados entre ambos test, realizando ambas pruebas a una misma muestra.

6.1. Test de Shapiro-Wilk

En 1965 los estadísticos Samuel Shapiro y Martin Wilk publicaron el test que lleva hoy su nombre y que, además, es considerado uno de los más potentes.

De la misma forma que los test anteriores, contrasta las hipótesis de que el conjunto de datos de estudio proviene de una distribución normal (H_0) o que, en cambio sigue cualquier otra distribución (H_1).

Para realizar el test, en primer lugar debemos considerar los datos estandarizados:

$$Z_i = \frac{X_i - \hat{X}}{S}, \quad i \in 1, \dots, n$$

Y, una vez estandarizados podemos construir el estadístico de la siguiente manera:[13]

$$W = \sum_{i=1}^{[n/2]} a_{i,n} (Z_{(n-i+1):n} - Z_{i:n})$$

donde $Z_{1:n} < \dots < Z_{n:n}$ es la muestra ordenada de los datos estandarizados y a_i , n constantes. Se trata de calcular las distancias entre los datos de la muestra ordenada, simétricos respecto de la mediana, es decir, la distancia entre el primero y el último, el

segundo y el penúltimo, y así sucesivamente, de forma genérica, se comparan el $Z_{i:n}$ y el $Z_{(n-i+1):n}$.

La finalidad es comparar estas distancias con las que habría en una muestra de observaciones normales.

Además, el coeficiente de determinación guarda una estrecha relación con el QQ-Plot, ya que lo podemos interpretar como el coeficiente de determinación de los puntos representados en el QQ-Plot. Por lo tanto, rechazaremos la normalidad cuando el estadístico de Shapiro-Wilk sea pequeño, ya que eso será debido a que los puntos del QQ-Plot se alejan de la recta esperada

Para visualizar de mejor forma la aplicación del test de Shapiro-Wilk nos ayudamos del siguiente ejemplo.

Ejemplo: *Se han medido los tiempos de ejecución de cierta tarea por parte de seis trabajadores, con los siguientes resultados en horas:*

6,55, 5,58, 4,33, 2,53, 5,49, 8,64, 3,87, 4,82

Se quiere saber si las mediciones provienen de una población normal.

Solución: *Ciertos parámetros de estas distribuciones, como media, desviación típica u otros, que en las secciones anteriores se determinaron con el propósito de facilitar la comparación, ahora son irrelevantes, puesto que el test de Shapiro-Wilk comienza con una estandarización de los datos.*

Ayudándonos de R, utilizaremos la función shapiro.test:

```
x = c(6.55, 5.58, 4.33, 2.53, 5.49, 8.64, 3.87, 4.82)
shapiro.test(x)
```

Ejecutando lo anterior obtenemos que:

```
Shapiro-Wilk normality test
```

```
data: x
W = 0.9827, p-value = 0.9643
```

El test nos devuelve un p-valor muy grande, 0.9643, que es claramente mayor que 0.05 ya que estamos suponiendo un nivel de significación del 5 por ciento, por lo tanto, no tenemos evidencias suficientes para rechazar la hipótesis de normalidad.

6.2. Test de Shapiro-Francia

Introducido por Samuel Shapiro y R. S. Francia en 1972, es una simplificación del test de Shapiro-Wilk, por lo que su funcionamiento es muy similar y son pocas las diferencias entre ambos.

Así pues, sea $x_{(i)}$ el i -ésimo valor ordenado de nuestra muestra (véase como ejemplo que si nuestra muestra fuese 5, -1, 2, 0 nuestro $x_2 = 0$), tendremos que el estadístico está definido como:[14]

$$W' = Cor^2(x_{(i)}, m_i) = \frac{\sigma_{(x,m)}}{\sigma_x \sigma_m} = \frac{\sum_{i=1}^n (x_{(i)} - \bar{x})(m_i - \bar{m})}{\sqrt{(\sum_{i=1}^n (x_{(i)} - \bar{x})^2)(\sum_{i=1}^n (m_i - \bar{m})^2)}}$$

Además, se tiene que dicho estadístico es el cuadrado del coeficiente de correlación entre los estadísticos de orden muestrales (es decir, nuestros cuantiles en un gráfico QQ-plot) $x_{(1)}, \dots, x_{(n)}$ y los valores esperados de los estadísticos de orden de la distribución $N(0, 1)$ m_1, \dots, m_n , es decir, nuestros cuantiles teóricos en el QQ-plot.

Del mismo modo que hicimos con el test de Shapiro-Wilk, podemos seguir el mismo ejemplo que el realizado para dicho test y ver que ocurre para la variante de Shapiro-Francia.

Ejemplo: Sean los mismos tiempos de ejecución de cierta tarea por parte de seis operarios, con los mismos resultados en horas:

6,55, 5,58, 4,33, 2,53, 5,49, 8,64, 3,87, 4,82

Veamos si siguen una distribución normal.

Solución: Ayudándonos de R, utilizaremos la función `sf.test`, de la librería `nortest` :

```
library(nortest)
x = c(6.55, 5.58, 4.33, 2.53, 5.49, 8.64, 3.87, 4.82)
sf.test(x)
```

Que, tras ejecutar, nos dará el siguiente resultado:

Shapiro-Francia normality test

```
data: x
W = 0.96053, p-value = 0.826
```

El test nos devuelve un p-valor grande, 0.826, que es claramente mayor que 0.05 ya que estamos suponiendo un nivel de significación del 5 por ciento igual que para el test de Shapiro-Wilk, por lo tanto, no tenemos evidencias suficientes para rechazar la hipótesis de normalidad.

Capítulo 7

Estudio comparativo mediante simulaciones

En el siguiente capítulo se comparan los resultados obtenidos en las simulaciones realizadas en R , donde para diferentes distribuciones de probabilidad se realizan los distintos test estudiados previamente y se calcula la potencia de cada método aplicado a cada distribución para distintos tamaños muestrales.

Para cada distribución fueron realizadas 1000 simulaciones al 5 por ciento de significación por lo que cuanto más cercano sea el tanto por uno de rechazo a 0.05 más significativo será el resultado de que se acepta la hipótesis nula.

Dichos resultados han sido tabulados de forma que en cada tabla se observan las potencias de los distintos test para tamaños muestrales diferentes de muestras que siguen una distribución específica. Adicionalmente, los datos de las tablas son representados mediante gráficas donde las potencias de los test se superponen para tener una forma más visual de poder compararlos.

Finalmente, se concluye el capítulo con las conclusiones extraídas de las simulaciones.

7.1. Resultados de las simulaciones

Para comenzar y como ejemplo, podemos observar lo que ocurre al aplicar los distintos test a una distribución normal.

Porcentajes de rechazo de los diferentes test de normalidad para muestras simuladas que siguen una **distribución Normal**

Test hechos al 5% de significación y 1000 simulaciones para cada tamaño muestral

	Tamaño 10	Tamaño 30	Tamaño 50	Tamaño 100	Tamaño 500
Test Chi Cuadrado de Pearson	0.074	0.057	0.055	0.038	0.033
Test de Lilliefors	0.054	0.049	0.039	0.054	0.044
Test de Cramer-von-Mises	0.047	0.048	0.047	0.045	0.053
Test de Anderson-Darling	0.043	0.049	0.043	0.048	0.058
Test de Jarque-Bera	0.052	0.060	0.048	0.045	0.049
Test de Shapiro-Wilk	0.043	0.068	0.042	0.051	0.051
Test de Shapiro-Francia	0.057	0.066	0.052	0.050	0.052

Figura 7.1: Tantos por uno de rechazo para una distribución normal.

Al tratarse de una distribución normal, la totalidad de los test dan resultados muy similares y cercanos al 0.05 de rechazo, pues todos detectan la normalidad de la distribución, pero veamos que ocurre cuando realizamos el test a una muestra que sigue otra distribución.

Tantos por uno de rechazo de los diferentes test de normalidad para muestras simuladas que siguen una **distribución Exponencial**

Test hechos al 5% de significación y 1000 simulaciones para cada tamaño muestral

	Tamaño 10	Tamaño 30	Tamaño 50	Tamaño 100	Tamaño 500
Test Chi Cuadrado de Pearson	0.412	0.834	0.980	1	1
Test de Lilliefors	0.316	0.784	0.956	1	1
Test de Cramer-von-Mises	0.382	0.912	0.982	1	1
Test de Anderson-Darling	0.408	0.940	0.994	1	1
Test de Jarque-Bera	0.342	0.846	0.970	1	1
Test de Shapiro-Wilk	0.452	0.964	1.000	1	1
Test de Shapiro-Francia	0.438	0.952	0.996	1	1

Figura 7.2: Tantos por uno de rechazo para una distribución exponencial.

Claramente el tanto por uno de rechazo se dispara con respecto al ejemplo anterior pues estamos trabajando con una distribución exponencial, pero podemos sacar algunas conclusiones de los datos anteriores ya que podemos observar que cuanto mayor es el

tamaño muestral, mayor es el rechazo de la hipótesis nula de que nuestra distribución sigue una normal.

Además, observamos que la prueba de Shapiro-Wilk es la que mejor y más rápido detecta que los datos no siguen una distribución normal, aunque no hay una gran diferencia con respecto al resto de pruebas dado que esta distribución solo tiene una cola, por lo que es más fácil de detectar para los test. Veamos que ocurre con otras distribuciones.

A continuación, analizaremos que ocurre en el caso de que nuestra muestra siga una distribución T-Student. Dado que esta distribución se aproxima a una distribución normal cuanto mayor es el número de sus grados de libertad, por lo que la lógica nos hace pensar que el rechazo de la hipótesis nula irá siendo menor según aumenten los grados de libertad.

Veamos que ocurre con dos distribuciones T-Student con 5 y 50 grados de libertad respectivamente y veremos que diferencias hay entre ambas.

Tantos por uno de rechazo de los diferentes test de normalidad para muestras simuladas que siguen una distribución T-Student con 5 grados de libertad					
Test hechos al 5% de significación y 1000 simulaciones para cada tamaño muestral					
	Tamaño 10	Tamaño 30	Tamaño 50	Tamaño 100	Tamaño 500
Test Chi Cuadrado de Pearson	0.086	0.096	0.106	0.194	0.576
Test de Lilliefors	0.098	0.136	0.210	0.318	0.894
Test de Cramer-von-Mises	0.100	0.188	0.268	0.424	0.958
Test de Anderson-Darling	0.104	0.196	0.284	0.460	0.980
Test de Jarque-Bera	0.126	0.254	0.424	0.626	0.996
Test de Shapiro-Wilk	0.110	0.208	0.344	0.544	0.992
Test de Shapiro-Francia	0.130	0.256	0.402	0.606	0.994

Figura 7.3: Tantos por uno de rechazo para una distribución T-Student con 5 grados de libertad.

Como podemos observar, para tamaños muestrales pequeños, muchos de los test son incapaces de detectar que nos encontramos ante una distribución no normal, incluso algunos, como el Chi-Cuadrado de Pearson precisa de un tamaño muestral muy grande para obtener un resultado significativo que permita rechazar la hipótesis nula. Esta situación se complica, naturalmente, cuando aumentan los grados de libertad:

Tantos por uno de rechazo de los diferentes test de normalidad para muestras simuladas que siguen una distribución T-Student con 50 grados de libertad					
Test hechos al 5% de significación y 1000 simulaciones para cada tamaño muestral					
	Tamaño 10	Tamaño 30	Tamaño 50	Tamaño 100	Tamaño 500
Test Chi Cuadrado de Pearson	0.080	0.044	0.066	0.052	0.040
Test de Lilliefors	0.070	0.042	0.070	0.056	0.058
Test de Cramer-von-Mises	0.068	0.040	0.066	0.072	0.066
Test de Anderson-Darling	0.064	0.044	0.068	0.068	0.070
Test de Jarque-Bera	0.052	0.054	0.076	0.064	0.116
Test de Shapiro-Wilk	0.050	0.032	0.060	0.058	0.086
Test de Shapiro-Francia	0.062	0.056	0.080	0.060	0.104

Figura 7.4: Tantos por uno de rechazo para una distribución T-Student con 50 grados de libertad.

Tal y como se preveía absolutamente ninguno de los test es capaz de dar unos resultados que permitan rechazar la hipótesis nula.

A continuación pondremos de manifiesto una de las mayores desventajas del test de Jarque-Bera que, recordemos, está basado en la asimetría y la curtosis de la función de distribución por lo que, cuando nos encontramos con distribuciones similares a la normal en esos parámetros, su veracidad se ve afectada notablemente.

Tantos por uno de rechazo de los diferentes test de normalidad para muestras simuladas que siguen una distribución Binomial					
Test hechos al 5% de significación y 1000 simulaciones para cada tamaño muestral					
	Tamaño 10	Tamaño 30	Tamaño 50	Tamaño 100	Tamaño 500
Test Chi Cuadrado de Pearson	0.116	0.258	0.676	1.000	1.000
Test de Lilliefors	0.106	0.252	0.476	0.898	1.000
Test de Cramer-von-Mises	0.076	0.154	0.258	0.828	1.000
Test de Anderson-Darling	0.076	0.148	0.256	0.776	1.000
Test de Jarque-Bera	0.046	0.046	0.036	0.036	0.032
Test de Shapiro-Wilk	0.070	0.104	0.164	0.338	1.000
Test de Shapiro-Francia	0.060	0.108	0.150	0.298	1.000

Figura 7.5: Tantos por uno de rechazo para una distribución binomial.

Mientras que la mayoría de los test realizados aumentan el rechazo a la hipótesis nula según aumenta el tamaño muestral siguiendo con la tendencia de los ejemplos anteriores, el test de Jarque-Bera es incapaz de detectar que no se trata de una distribución normal y su porcentaje de rechazo sigue próximo al 5% o incluso por debajo, por lo que acepta la hipótesis de normalidad para cualquier tamaño muestral. Además, observamos de nuevo que para muestras pequeñas, la mayoría de los test realizados arrojan resultados poco significativos, por lo que el tamaño muestral juega un papel clave a la hora de la realización de las pruebas de normalidad.

Algo similar acontece cuando analizamos lo que ocurre en una distribución uniforme, aunque en menor medida.

Tantos por uno de rechazo de los diferentes test de normalidad para muestras simuladas que siguen una **distribución Uniforme**
Test hechos al 5% de significación y 1000 simulaciones para cada tamaño muestral

	Tamaño 10	Tamaño 30	Tamaño 50	Tamaño 100	Tamaño 500
Test Chi Cuadrado de Pearson	0.085	0.109	0.213	0.463	1
Test de Lilliefors	0.058	0.131	0.267	0.613	1
Test de Cramer-von-Mises	0.075	0.235	0.433	0.838	1
Test de Anderson-Darling	0.080	0.294	0.572	0.954	1
Test de Jarque-Bera	0.015	0.002	0.017	0.722	1
Test de Shapiro-Wilk	0.086	0.395	0.745	0.998	1
Test de Shapiro-Francia	0.050	0.169	0.465	0.977	1

Figura 7.6: Tantos por uno de rechazo para una distribución uniforme.

Obtenemos que el test de Jarque-Bera sigue dando resultados poco fiables si la muestra no es de la magnitud adecuada, mientras que Shapiro-Wilk sigue apareciendo como la prueba que mejores porcentajes de rechazo tiene.

De la misma forma que analizamos el impacto del aumento en los grados de libertad de una distribución T-Student para ver como afectaba esos cambios a la fiabilidad de los test, podemos realizar un estudio similar con la distribución Chi Cuadrado, tal y como se expone a continuación

Tantos por uno de rechazo de los diferentes test de normalidad para muestras simuladas que siguen una **distribución Chi cuadrado con 5 grados de libertad**
Test hechos al 5% de significación y 1000 simulaciones para cada tamaño muestral

	Tamaño 10	Tamaño 30	Tamaño 50	Tamaño 100	Tamaño 500
Test Chi Cuadrado de Pearson	0.158	0.308	0.458	0.872	1
Test de Lilliefors	0.146	0.412	0.568	0.894	1
Test de Cramer-von-Mises	0.166	0.514	0.732	0.976	1
Test de Anderson-Darling	0.184	0.580	0.792	0.984	1
Test de Jarque-Bera	0.190	0.516	0.718	0.986	1
Test de Shapiro-Wilk	0.200	0.668	0.880	0.994	1
Test de Shapiro-Francia	0.212	0.614	0.840	0.994	1

Figura 7.7: Tantos por uno de rechazo para una distribución chi cuadrado con 5 grados de libertad.

Se observa que todos los test rechazan de manera bastante fiable la hipótesis de normalidad en cuanto se aumenta el tamaño de la muestra mínimamente. En cambio, al aumentar los grados de libertad de la distribución simulada, los resultados obtenidos sufren alguna variación sustancial.

Tantos por uno de rechazo de los diferentes test de normalidad para muestras simuladas que siguen una **distribución Chi Cuadrado con 50 grados de libertad**
Test hechos al 5% de significación y 1000 simulaciones para cada tamaño muestral

	Tamaño 10	Tamaño 30	Tamaño 50	Tamaño 100	Tamaño 500
Test Chi Cuadrado de Pearson	0.094	0.034	0.094	0.082	0.288
Test de Lilliefors	0.062	0.066	0.098	0.152	0.566
Test de Cramer-von-Mises	0.070	0.068	0.138	0.188	0.694
Test de Anderson-Darling	0.068	0.076	0.152	0.222	0.778
Test de Jarque-Bera	0.050	0.112	0.176	0.250	0.892
Test de Shapiro-Wilk	0.066	0.102	0.192	0.290	0.904
Test de Shapiro-Francia	0.070	0.102	0.186	0.276	0.882

Figura 7.8: Tantos por uno de rechazo para una distribución chi cuadrado con 50 grados de libertad.

Los porcentajes de rechazo sufren un gran recorte y ninguno es capaz de arrojar un resultado veraz que rechace la hipótesis nula sin necesidad de estar ante un volumen muy

grande de la muestra. El mas notorio es el caso del test Chi-Cuadrado de Pearson, cuya fiabilidad cae en picado al aumentar los grados de libertad mientras que el resto de test, pese a que rechazan de peor forma la normalidad, siguen aportando resultados fiables para tamaños muestrales adecuados.

Por último, analizamos el buen comportamiento de los test para dos distribuciones (Gamma y Normal-Logarítmica) donde observamos que todos son capaces de rechazar la hipótesis nula con porcentajes de rechazo significativos.

En primer lugar, podremos observar en el estudio de la distribución Gamma que el test de Shapiro-Wilk requiere de un tamaño muestral mucho menor que el resto de test para alcanzar porcentajes de rechazo de entorno al 90 %, lo cual pone de manifiesto una mayor potencia y fiabilidad de esta prueba frente a los demás que, aunque rechazan la normalidad, tienen un comportamiento algo peor.

Tantos por uno de rechazo de los diferentes test de normalidad para muestras simuladas que siguen una distribución Gamma					
Test hechos al 5% de significación y 1000 simulaciones para cada tamaño muestral					
	Tamaño 10	Tamaño 30	Tamaño 50	Tamaño 100	Tamaño 500
Test Chi Cuadrado de Pearson	0.268	0.610	0.842	0.996	1
Test de Lilliefors	0.210	0.602	0.810	0.994	1
Test de Cramer-von-Mises	0.238	0.730	0.912	1.000	1
Test de Anderson-Darling	0.258	0.822	0.950	1.000	1
Test de Jarque-Bera	0.234	0.698	0.910	1.000	1
Test de Shapiro-Wilk	0.266	0.900	0.994	1.000	1
Test de Shapiro-Francia	0.264	0.862	0.978	1.000	1

Figura 7.9: Tantos por uno de rechazo para una distribución gamma.

Para terminar, los resultados simulados de la distribución Log-Normal, vamos a poder observar una gran fiabilidad de los resultados arrojados por todas las pruebas ya que pese a realizar el test con muestras pequeñas, los porcentajes de rechazo son notablemente mayores que con el resto de distribuciones.

Tantos por uno de rechazo de los diferentes test de normalidad para muestras simuladas que siguen una distribución Logarítmica Normal					
Test hechos al 5% de significación y 1000 simulaciones para cada tamaño muestral					
	Tamaño 10	Tamaño 30	Tamaño 50	Tamaño 100	Tamaño 500
Test Chi Cuadrado de Pearson	0.558	0.944	0.998	1	1
Test de Lilliefors	0.468	0.926	0.994	1	1
Test de Cramer-von-Mises	0.566	0.978	1.000	1	1
Test de Anderson-Darling	0.592	0.986	1.000	1	1
Test de Jarque-Bera	0.498	0.948	0.998	1	1
Test de Shapiro-Wilk	0.628	0.994	1.000	1	1
Test de Shapiro-Francia	0.608	0.990	1.000	1	1

Figura 7.10: Tantos por uno de rechazo para una distribución normal logarítmica.

Tal y como comentábamos previamente, no precisamos de que nuestra muestra sea de un tamaño muy grande para obtener resultados significativos de cualquier test, ya que prácticamente todos arrojan porcentajes de rechazo de casi el 50 % para el tamaño muestral más pequeño y que se disparan en cuanto aumentamos ligeramente la magnitud de la muestra.

Para ver de forma más visual las potencias de los test recurrimos a unas gráficas donde superponemos los resultados que previamente hemos tabulado, en ellas observamos como los métodos con mejores potencias, son los que más destacan en las gráficas dado que tienen un mayor porcentaje de rechazo de la hipótesis nula. Para poder seguir las gráficas nos ayudaremos de la siguiente leyenda:

	Test de Pearson
	Test de Jarque-Bera
	Test de Shapiro-Wilk
	Test de Shapiro-Francia
	Test de Lilliefors
	Test de Cramer
	Test de Anderson

Figura 7.11: Leyenda.

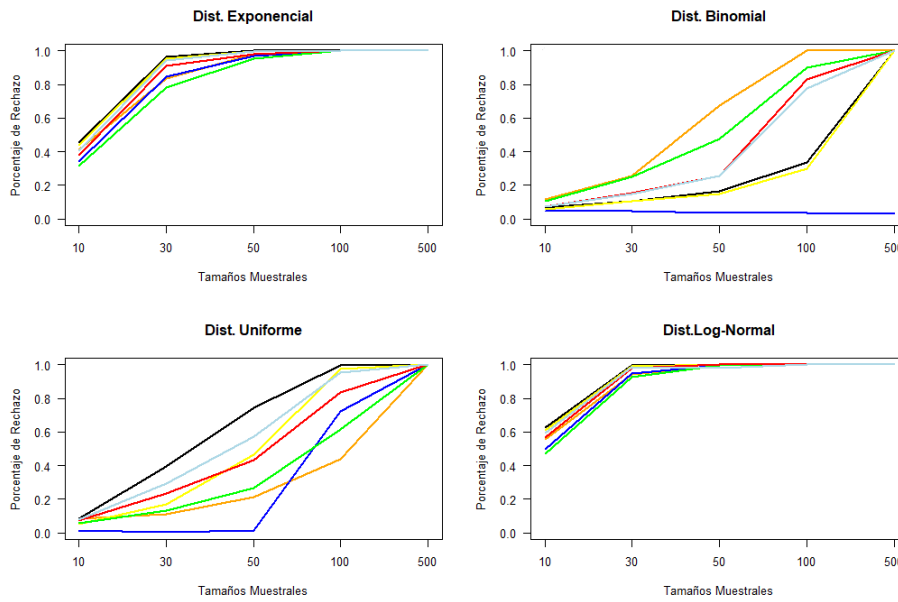


Figura 7.12: Bloque I de gráficas

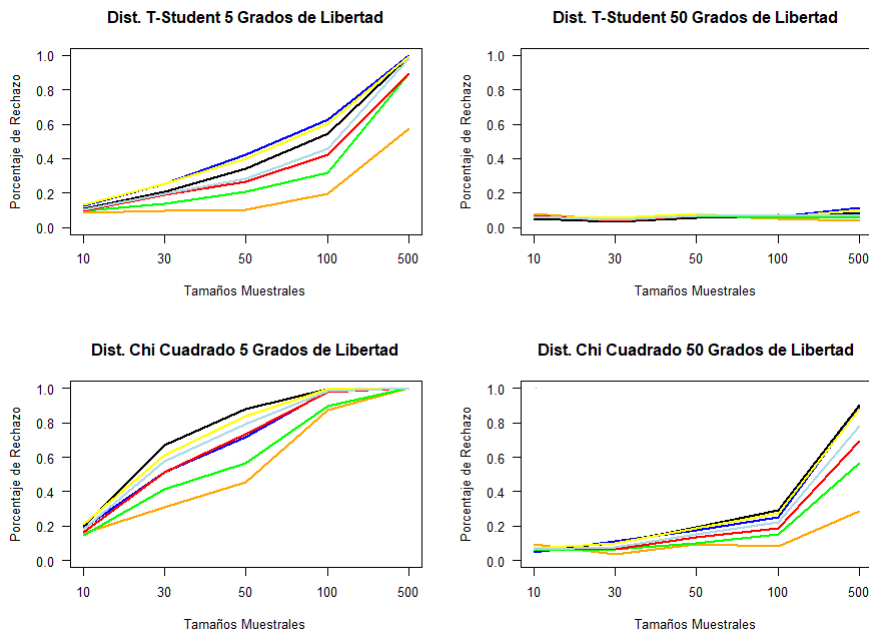


Figura 7.13: Bloque II de gráficas

Una vez revisadas las simulaciones realizadas, podemos extraer algunas conclusiones sobre los diversos test realizados.

7.2. Conclusiones

Analizados los resultados arrojados del comportamiento de los test ante diferentes distribuciones, podemos extraer algunas conclusiones sobre cuales de ellos arrojan resultados mas fiables y tienen una mayor potencia respecto a los demás.

Observamos que, salvo para la distribución binomial, el test de Shapiro-Wilk es el que mejor porcentaje de rechazo arroja para cada tamaño muestral, con lo que podemos concluir que es el test que más potencia tiene de los estudiados. Además su variante de Shapiro-Francia también presenta muy buenos resultados para la mayoría de distribuciones.

Respecto al test de Jarque-Bera, tal y como analizábamos previamente, su fiabilidad se ve trastocada cuando se enfrenta a distribuciones con parámetros de asimetría y curtosis similares a la normal, por lo que puede arrojar resultados erróneos que nos hagan aceptar la hipótesis de normalidad cuando esta no sea cierta.

Por otro lado, el test Chi Cuadrado de Pearson se presenta como el más fiable en la distribución binomial, pero su rendimiento baja notoriamente en el resto de distribuciones, sobre todo en las T-Student y Chi Cuadrado, independientemente de sus grados de libertad que, como vimos, también afectan a los resultados obtenidos.

El resto de test arrojaron resultados muy similares, por lo que el test de Shapiro-Wilk y su variante de Shapiro-Francia se presuponen como los test más fiables a la hora de realizar un contraste de normalidad cuando desconozcamos la distribución de la muestra a la que le apliquemos el test.

Por último, recalcar el papel fundamental que juega el tamaño muestral a la hora de realizar las pruebas, ya que es evidente, tanto en las tablas como en las gráficas, que cuanto mayor es el volumen de la muestra, mejores y más fiables son los resultados de los porcentajes de rechazo independientemente del test realizado.

Apéndice A

Códigos de R

En el siguiente anexo se presentan todas las líneas de código utilizadas tanto en las simulaciones como en la obtención de ejemplos a lo largo de todo el trabajo, junto con las explicaciones de las funciones y paquetes utilizados.

A continuación se presentan los códigos utilizados en *R* para obtener los resultados simulados.

En primer lugar, destacaremos las funciones necesarias para realizar los diferentes test, ya que más tarde observaremos que el código es el mismo para las distintas distribuciones, con la única variación de los comandos necesarios para la generación de números aleatorios que siguen cada una de las distribuciones específicas, dichas funciones proceden de los paquetes *normtest* y *nortest*.

Así pues:

```
ad.test(x) # Para realizar el test de Anderson-Darling
cvm.test(x) # Para realizar el test de Cramer-von-Mises
sf.test(x) # Para realizar el test de Shapiro-Francia
shapiro.test(x) # Para realizar el test de Shapiro-Wilk
lillie.test(x) # Para realizar el test de Lilliefors
pearson.test(x) # Para realizar el test Chi-Cuadrado de Pearson
jb.norm.test(x) # Para realizar el test de Jarque-Bera
```

A continuación se explica brevemente el funcionamiento de las funciones anteriores, junto con los parámetros que requieren.

La función *ad.test()* emplea el estadístico de prueba $A = -n - \frac{1}{n} \sum_{i=1}^n (2i-1)(\ln(F(x_{(i)})) + \ln(1 - F(x_{(n-i+1))))$, tal y como definimos en el capítulo 3 donde \bar{x} y s son la media y la desviación estándar de los valores de los datos. El p-valor se calcula a partir del estadístico modificado $Z = A(1,0 + \frac{0,75}{n} + \frac{2,25}{n^2})$ de acuerdo con la tabla 4.9 en Stephens (1986). [15]

La función *cvm.test()* requiere de varios parámetros para su correcto funcionamiento:

- *x*: el vector que contiene los valores de nuestra muestra.
- *null*: es la distribución específica que queremos comprobar que sigue nuestra muestra
- *estimated*: es un valor lógico, de ser "TRUE"(que viene por defecto) asumirá que los valores de media y varianza son estimados, de ser "FALSE"los da por conocidos.

Las funciones *shapiro.test()* y *sf.test()* tienen un funcionamiento idéntico cuya única diferencia es la misma que en el propio método ya que para realizar el test de Shapiro-Francia sí estimará media y varianza. En ambos casos el p-valor viene calculado por la fórmula dada por Royston (1993). [16]

La función *pearson.test()* requiere de tres parámetros para su correcto funcionamiento:

- *x*: el vector que contiene los valores de nuestra muestra.
- *n.classes*: el número de intervalos en los que se dividirá la muestra (por defecto se dividirá en $k = \lceil (2(n^{2/5})) \rceil$ intervalos con $k = \lceil (2(n^{2/5})) \rceil$)
- *adjust*: es un valor lógico, de ser "TRUE"(que viene por defecto) la función realizará el test a una distribución chi-cuadrado con $n.classes-3$ grados de libertad (es decir, como si estuviésemos estimando la media y la varianza) y de ser "FALSE"serán $n.classes-1$ grados de libertad

Dicha función usará como estadístico $P = \frac{\sum(C_i - E_i)^2}{E_i}$ donde C_i son las frecuencias observadas y E_i las esperadas.

La función *jb.norm.test(x)* requiere de los siguientes parámetros:

- *x*: el vector que contiene los valores de nuestra muestra.
- *n.repl*: el número de replicaciones aleatorias por el método de Monte Carlo.

Esta función se basa en el siguiente estadístico: $JB = \frac{n}{6} (S^2 + \frac{1}{4}(K - 3)^2)$ tal y como está explicado en el capítulo 4, con S el coeficiente de asimetría y de la muestra y K el de curtosis. Por último, el p-valor se calcula mediante simulación de Monte Carlo.

A continuación, se muestra como fueron generadas las variables aleatorias para cada una de las simulaciones. Dichas muestras son números aleatorios que siguen una distribución específica tal como sigue:

```

x=rnorm(muestra) # Dist. Normal
y=runif(muestra) # Dist. Unif
z=rexp(muestra) # Dist. Expo
chi1=rchisq(muestra,5) #Chi 5 grados
chi2=rchisq(muestra,50) #Chi 50 grados
t1=rt(muestra,5) # TStudent 5 grados
t2=rt(muestra,50) # TStudent 50 grados
bi=rbinom(muestra,muestra,0.5) # Binomial
l=rlnorm(muestra) # Log-normal
b=rbeta(muestra,1.5,1.5) # Beta
g=rgamma(muestra,1.5,1.5) # Gamma

```

Como podemos observar, el tamaño muestral viene determinado por la variable *muestra* cuyo valor se modifica dentro de un bucle como veremos a continuación.

Como ejemplo, veremos como sigue el código para la realización del test a la muestra *x* que sigue una distribución normal, ya que el código es exactamente el mismo y se sigue de la misma forma para el resto.

```

set.seed(123456)

n=c(10,30,50,100,500) # Tamaño muestral
ns=1000 # Numero de muestras simuladas

# Librerías necesarias para los distintos test
library(normtest)
library(nortest)

# Creación de vectores donde irán los p-valores de cada test
pv_anderson_norm=c()
pv_cvm_norm=c()
pv_sf_norm=c()
pv_lil_norm=c()
pv_pear_norm=c()
pv_sw_norm=c()
pv_jb_x=c()

# Creación de vectores de potencias de cada test

```

```
pot_anderson_norm=c()
pot_cvm_norm=c()
pot_sf_norm=c()
pot_lil_norm=c()
pot_pear_norm=c()
pot_sw_norm=c()
pot_jb_x=c()

# Bucle de la realización de los test a las muestras
for (i in 1:length(n)){
  muestra=n[i]
  for (is in 1:ns){
    x=rnorm(muestra,0,1)

    anderson_norm=ad.test(x)
    cvm_norm=cvm.test(x)
    sf_norm=sf.test(x)
    lil_norm=lillie.test(x)
    pear_norm=pearson.test(x)
    sw_norm=shapiro.test(x)
    jb_norm=jb.norm.test(x)

    # Vectores con los p-valores
    pv_anderson_norm[is]= anderson_norm$p.value
    pv_cvm_norm[is]=cvm_norm$p.value
    pv_sf_norm[is]=sf_norm$p.value
    pv_lil_norm[is]=lil_norm$p.value
    pv_pear_norm[is]=pear_norm$p.value
    pv_sw_norm[is]=sw_norm$p.value
    pv_jb_x[is]=jb_norm$p.value

  }

  # Vectores con las potencias de cada test
  pot_anderson_norm[i]= sum(pv_anderson_norm<0.05)/ns
  pot_cvm_norm[i]=sum(pv_cvm_norm<0.05)/ns
```

```

pot_sf_norm[i]=sum(pv_sf_norm<0.05)/ns
pot_lil_norm[i]=sum(pv_lil_norm<0.05)/ns
pot_pear_norm[i]=sum(pv_pear_norm<0.05)/ns
pot_sw_norm[i]=sum(pv_sw_norm<0.05)/ns
pot_jb_x[i]=sum(pv_jb_x<0.05)/ns
}

# Creación de un data frame con los valores de las potencias
Potencias_Normal= data.frame(rbind(pot_pear_norm,pot_lil_norm,pot_cvm_norm,
pot_anderson_norm,pot_jb_x,pot_sw_norm,pot_sf_norm))

rownames(Potencias_Normal)<-c("Test Chi Cuadrado de Pearson","Test de Lilliefors",
"Test de Cramer-von-Mises","Test de Anderson-Darling","Test de Jarque-Bera",
"Test de Shapiro-Wilk","Test de Shapiro-Francia")

colnames(Potencias_Normal)<-c("Tamaño 10", "Tamaño 30", "Tamaño 50", "Tamaño 100",
"Tamaño 500")

#Visualización del data frame
View(Potencias_Normal)

```

Una vez generado el data frame con los datos correspondientes, usamos el siguiente código para la creación de las tablas y las gráficas para representar los resultados obtenidos.

De la misma forma que para el código anterior, únicamente se presenta el código empleado para una distribución normal, pues la única variación que hay es el uso de los data frames correspondientes a las distintas distribuciones.

```
# PARA LAS GRAFICAS
```

```

# Division del área de trabajo para que salgan 4 gráficas en un mismo plot
Conf2x2 = matrix(c(1:4), nrow=2, byrow=TRUE)
layout(Conf2x2)
layout.show(4)

# GRAFICA EXPONENCIAL
plot(1:5,xlab="Tamaños Muestrales",ylab="Porcentaje de Rechazo",main=
"Dist. Exponencial", ylim=c(0,1),xaxt="n",las=2)

```

```
axis(1, at = 1:5, labels = n)
lines(1:5,pot_pear_exp,lwd=2,col="orange")
lines(1:5,pot_jb_exp,lwd=2,col="blue")
lines(1:5,pot_sw_exp,lwd=2)
lines(1:5,pot_sf_exp,lwd=2,col="yellow")
lines(1:5,pot_lil_exp,lwd=2,col="green")
lines(1:5,pot_cvm_exp,lwd=2,col="red")
lines(1:5,pot_anderson_exp,lwd=2,col="light blue")

# GRAFICA BINOMIAL
plot(1:5,xlab="Tamaños Muestrales",ylab="Porcentaje de Rechazo",main=
"Dist. Binomial", ylim=c(0,1),xaxt="n",las=2)
axis(1, at = 1:5, labels = n)
lines(1:5,pot_pear_binom,lwd=2,col="orange")
lines(1:5,pot_jb_binom,lwd=2,col="blue")
lines(1:5,pot_sw_binom,lwd=2)
lines(1:5,pot_sf_binom,lwd=2,col="yellow")
lines(1:5,pot_lil_binom,lwd=2,col="green")
lines(1:5,pot_cvm_binom,lwd=2,col="red")
lines(1:5,pot_anderson_binom,lwd=2,col="light blue")

# GRAFICA UNIFORME
plot(1:5,xlab="Tamaños Muestrales",ylab="Porcentaje de Rechazo",main=
"Dist. Uniforme", ylim=c(0,1),xaxt="n",las=2)
axis(1, at = 1:5, labels = n)
lines(1:5,pot_pear_unif,lwd=2,col="orange")
lines(1:5,pot_jb_unif,lwd=2,col="blue")
lines(1:5,pot_sw_unif,lwd=2)
lines(1:5,pot_sf_unif,lwd=2,col="yellow")
lines(1:5,pot_lil_unif,lwd=2,col="green")
lines(1:5,pot_cvm_unif,lwd=2,col="red")
lines(1:5,pot_anderson_unif,lwd=2,col="light blue")

# GRAFICA LOG
plot(1:5,xlab="Tamaños Muestrales",ylab="Porcentaje de Rechazo",main=
"Dist.Log-Normal", ylim=c(0,1),xaxt="n",las=2)
```

```

axis(1, at = 1:5, labels = n)
lines(1:5,pot_pear_log,lwd=2,col="orange")
lines(1:5,pot_jb_log,lwd=2,col="blue")
lines(1:5,pot_sw_log,lwd=2)
lines(1:5,pot_sf_log,lwd=2,col="yellow")
lines(1:5,pot_lil_log,lwd=2,col="green")
lines(1:5,pot_cvm_log,lwd=2,col="red")
lines(1:5,pot_anderson_log,lwd=2,col="light blue")

# GRAFICA TSTUDENT5
plot(1:5,xlab="Tamaños Muestrales",ylab="Porcentaje de Rechazo",main=
"Dist. T-Student 5 Grados de Libertad", ylim=c(0,1),xaxt="n",las=2)
axis(1, at = 1:5, labels = n)
lines(1:5,pot_pear_t5,lwd=2,col="orange")
lines(1:5,pot_jb_t5,lwd=2,col="blue")
lines(1:5,pot_sw_t5,lwd=2)
lines(1:5,pot_sf_t5,lwd=2,col="yellow")
lines(1:5,pot_lil_t5,lwd=2,col="green")
lines(1:5,pot_cvm_t5,lwd=2,col="red")
lines(1:5,pot_anderson_t5,lwd=2,col="light blue")

# GRAFICA TSTUDENT 50
plot(1:5,xlab="Tamaños Muestrales",ylab="Porcentaje de Rechazo",main=
"Dist. T-Student 50 Grados de Libertad", ylim=c(0,1),xaxt="n",las=2)
axis(1, at = 1:5, labels = n)
lines(1:5,pot_pear_t50,lwd=2,col="orange")
lines(1:5,pot_jb_t50,lwd=2,col="blue")
lines(1:5,pot_sw_t50,lwd=2)
lines(1:5,pot_sf_t50,lwd=2,col="yellow")
lines(1:5,pot_lil_t50,lwd=2,col="green")
lines(1:5,pot_cvm_t50,lwd=2,col="red")
lines(1:5,pot_anderson_t50,lwd=2,col="light blue")

# GRAFICA CHI CUADRADO 5
plot(1:5,xlab="Tamaños Muestrales",ylab="Porcentaje de Rechazo",main=
"Dist. Chi Cuadrado 5 Grados de Libertad", ylim=c(0,1),xaxt="n",las=2)

```

```

axis(1, at = 1:5, labels = n)
lines(1:5,pot_pear_c5,lwd=2,col="orange")
lines(1:5,pot_jb_c5,lwd=2,col="blue")
lines(1:5,pot_sw_c5,lwd=2)
lines(1:5,pot_sf_c5,lwd=2,col="yellow")
lines(1:5,pot_lil_c5,lwd=2,col="green")
lines(1:5,pot_cvm_c5,lwd=2,col="red")
lines(1:5,pot_anderson_c5,lwd=2,col="light blue")

# GRAFICA CHI CUADRADO 50
plot(1:5,xlab="Tamaños Muestrales",ylab="Porcentaje de Rechazo",main=
"Dist. Chi Cuadrado 50 Grados de Libertad", ylim=c(0,1),xaxt="n",las=2)
axis(1, at = 1:5, labels = n)
lines(1:5,pot_pear_c50,lwd=2,col="orange")
lines(1:5,pot_jb_c50,lwd=2,col="blue")
lines(1:5,pot_sw_c50,lwd=2)
lines(1:5,pot_sf_c50,lwd=2,col="yellow")
lines(1:5,pot_lil_c50,lwd=2,col="green")
lines(1:5,pot_cvm_c50,lwd=2,col="red")
lines(1:5,pot_anderson_c50,lwd=2,col="light blue")

# PARA LAS TABLAS

# Uso de la librería necesaria
library(gt)

# Implementación de los datos en las tablas
tab_1_ad<-gt(Potencias_Normal, rowname_col = "Test",rownames_to_stub = TRUE)
tab_2_ad<- tab_1_ad %>%
tab_header(
title = md("Tantos por uno de rechazo de los diferentes test de
normalidad para muestras simuladas que siguen una distribución Normal"),
subtitle = md("Test hechos al 5% de significación y 1000 simulaciones para cada
tamaño muestral")
) %>%

```

```

tab_style(
  style = list(
    cell_fill(color= "light blue",alpha=0.3),
    cell_text(weight = "normal")
  ),
  locations = cells_body(
    columns = TRUE,
    rows = c(1,3,5,7))) %>%
  tab_style(
    style = list(
      cell_fill(color= "light blue",alpha=0.3),
      cell_text(weight = "normal")
    ),
    locations = cells_stub(
      rows = c(1,3,5,7)))
tab_2_ad

```

Para el apartado de métodos gráficos, el siguiente código fue empleado para obtener los histogramas que aparecen en dicho capítulo:

```

set.seed(123456)

# Creación de muestras con distintos tamaños
x= seq(-4, 4, length = 500)
normal10=rnorm(10,0,1)
normal50=rnorm(50,0,1)
normal250=rnorm(250,0,1)
normal500=rnorm(500,0,1)

Conf2x2 = matrix(c(1:4), nrow=2, byrow=TRUE)
layout(Conf2x2)
layout.show(4)

# Creación de los histogramas con la función de densidad de la normal superpuesta
hist(normal10, freq = F,xlim = c(-4, 4),main="Tamaño muestral n=10",xlab="")
lines(x, dnorm(x, mean = 0, sd = 1), col = "red", lty = 2, lwd = 2)

```

```
hist(normal50, freq = F,xlim = c(-4, 4),ylim =c(0, 0.5),main="Tamaño muestral
n=50",xlab="")
```

```
lines(x, dnorm(x, mean = 0, sd = 1), col = "red", lty = 2, lwd = 2)
```

```
hist(normal250, freq = F,xlim = c(-4, 4),ylim =c(0, 0.5),main="Tamaño muestral
n=250",xlab="")
```

```
lines(x, dnorm(x, mean = 0, sd = 1), col = "red", lty = 2, lwd = 2)
```

```
hist(normal500, freq = F,xlim = c(-4, 4),ylim =c(0, 0.5),main="Tamaño muestral
n=500",xlab="")
```

```
lines(x, dnorm(x, mean = 0, sd = 1), col = "red", lty = 2, lwd = 2)
```

Por último, para la obtención de los QQ-Plots del capítulo 6 de Métodos gráficos, se empleó el paquete *ggplot2* junto con la librería *ggpubr* tal como sigue:

```
# GRAFICAS QQPLOT PARA CADA DISTRIBUCION
```

```
set.seed(123456)
```

```
# Instalación del paquete
```

```
install.packages(ggplot2)
```

```
# Uso de la librería requerida
```

```
library(ggpubr)
```

```
# Creación de los QQ-Plots
```

```
ggqqplot(x,add.params = list(color = "red"),conf.int = TRUE,
title= "Gráfico Q-Q de una Distribución Normal")
```

```
ggqqplot(t2,add.params = list(color = "red"),conf.int = TRUE,
title= "Gráfico Q-Q de una distribución T-Student y 20 grados de libertad")
```

```
ggqqplot(l,add.params = list(color = "red"),conf.int = TRUE,
title= "Gráfico Q-Q de una distribución Log-Normal")
```

Bibliografía

- [1] Portney, L.G. & Watkins, M.P. (2000). *Foundations of clinical research: applications to practice*. New Jersey: Prentice Hall Health. pp. 516–517.
- [2] Weisstein, Eric W. «*Normal Distribution Function*»
- [3] Plackett, R.L. (1983). «*Karl Pearson and the Chi-Squared Test*». *International Statistical Review* (International Statistical Institute (ISI)): 59-72
- [4] Vélez Ibarrola, R. y García Pérez, A. (1997). *Principios de Inferencia Estadística*. Universidad Nacional de Educación a Distancia (UNED).
- [5] Lilliefors, H. (June 1967), *On the Kolmogorov–Smirnov test for normality with mean and variance unknown* , *Journal of the American Statistical Association*, Vol. 62. pp. 399–402.
- [6] Watson, G.S. (1961) *Goodness-Of-Fit Tests on a Circle*, *Biometrika*, 48 (1/2), 109-114
- [7] Anderson, T. & Darling, D. (1954) *A test of goodness of fit*. *Journal of the American Statistical Association*, 49(268), 765-769.
- [8] Darling, D. (1957) *The Kolmogorov-Smirnov, Crámer-von Mises tests*. *The Annals of Mathematical Statistics*, 28(4), 823-838
- [9] Fco. Javier Martín-Pliego López (2007) *Introducción a la Estadística Económica y Empresarial. Teoría y Práctica.*, Editorial Thomson, (Madrid)
- [10] DeCarlo, Lawrence T. (1997). *On the Meaning and Use of Kurtosis*
- [11] Jarque, Carlos M.; Bera, Anil K. (1981). «*Efficient tests for normality, homoscedasticity and serial independence of regression residuals: Monte Carlo evidence*». *Economics Letters* 7 (4): 313-318.
- [12] Cleveland, W.S. (1994) *The Elements of Graphing Data*, Hobart Press.

- [13] Shapiro, S. S.; Wilk, M. B. (1965). «*An analysis of variance test for normality (complete samples)*». *Biometrika* 52 (3-4): 591-611
- [14] Shapiro, S. & Francia, R. (1972) *An approximate analysis of variance test for normality*. *Journal of the American Statistical Association*, 67, 215-216.
- [15] Stephens, M.A. (1986): *Tests based on EDF statistics*. In: D'Agostino, R.B. and Stephens, M.A., eds.: *Goodness-of-Fit Techniques*. Marcel Dekker, New York.
- [16] Royston, P. (1993): *A pocket-calculator algorithm for the Shapiro-Francia test for non-normality: an application to medicine*. *Statistics in Medicine*, 12, 181–184.