



Universidad de Oviedo

Análisis comparativo de tests de uniformidad

Ignacio Fernández Suárez

Supervisado por: Raúl Pérez Fernández

UNIVERSIDAD DE OVIEDO

Facultad de Ciencias

Grado en Matemáticas

Mayo de 2023

# Índice general

<b>1. Introducción</b>	<b>1</b>
1.1. Motivación . . . . .	1
1.2. Objetivos . . . . .	2
1.3. Estructura del trabajo . . . . .	3
<b>2. La distribución uniforme</b>	<b>5</b>
2.1. Distribución uniforme discreta . . . . .	5
2.2. Distribución uniforme continua . . . . .	6
2.2.1. Estimación paramétrica en la distribución uniforme continua . . . . .	8
2.3. Distribución uniforme estándar . . . . .	13
<b>3. Pruebas para la uniformidad en el intervalo unidad</b>	<b>18</b>
3.1. Pruebas gráficas . . . . .	18
3.1.1. Histogramas . . . . .	18
3.1.2. Diagramas de cajas . . . . .	19
3.1.3. Gráficos Q-Q . . . . .	19
3.2. Tests de uniformidad . . . . .	21
3.3. Tests tradicionales adaptados a la uniformidad en $[0,1]$ . . . . .	22
3.3.1. Test de la Razón de Verosimilitudes . . . . .	22
3.3.2. Test de Kolmogorov-Smirnov . . . . .	24
3.3.3. Test Chi-Cuadrado de Pearson . . . . .	25

3.3.4.	Test de Cramer-von Mises . . . . .	26
3.3.5.	Test de Anderson-Darling . . . . .	27
3.3.6.	Test de Kuiper . . . . .	28
3.4.	Tests específicos para la uniformidad . . . . .	29
3.4.1.	Tests de Zhang . . . . .	29
3.4.2.	Test Greenwood . . . . .	33
3.4.3.	Test Q . . . . .	34
3.4.4.	Test de Durbin-Kolmogorov . . . . .	35
3.4.5.	Test Kendall-Sherman . . . . .	37
3.4.6.	Test del Producto de Probabilidades de Pearson . . . . .	38
3.4.7.	Test de Sukhatme . . . . .	39
<b>4.</b>	<b>Experimentación</b>	<b>42</b>
4.1.	La distribución bajo $H_0$ , la uniforme estándar . . . . .	43
4.1.1.	La potencia en función del tamaño muestral . . . . .	43
4.2.	La familia uniforme . . . . .	45
4.3.	La familia $\beta(p, q)$ . . . . .	47
4.3.1.	La potencia en función del tamaño muestral . . . . .	48
4.3.2.	La potencia en función de los parámetros de forma . . . . .	49
4.4.	Familia Trapezoidal . . . . .	60
4.4.1.	La potencia en función del intervalo modal . . . . .	62
4.4.2.	La potencia en función del tamaño de muestra . . . . .	64
4.5.	Familia Triangular . . . . .	66
4.5.1.	La potencia en función de la moda . . . . .	67
<b>5.</b>	<b>Conclusiones</b>	<b>70</b>
	<b>Bibliografía</b>	<b>73</b>

# Capítulo 1

## Introducción

### 1.1. Motivación

En matemáticas, o cualquier otra ciencia experimental, la modelización es una parte fundamental en la mayoría de trabajos de investigación. Con ello, se obtiene un modelo de un fenómeno real el cual puede tener una resolución analítica (y proporcionar una solución exacta o aproximada basada en resultados asintóticos) o una resolución a raíz de una simulación (cuando no es posible obtener una solución exacta, o la aproximación no es buena). Dentro de los modelos que se estudian en matemáticas, hablaremos de los modelos estocásticos, que se diferencian de los deterministas por contar con una componente aleatoria. Estos modelos pueden resolverse mediante la conocida simulación estocástica o método de simulación de Monte Carlo [22].

Para llevar este tipo de simulación a cabo, es necesario tener una cantidad  $n$  de números aleatorios, que representarán una muestra, siendo  $n$  el tamaño muestral. Dichos números pueden ser aleatorios puros (donde no existe ningún patrón que permita deducirlos), pseudo-aleatorios (los que, aunque parezcan estar generados al azar, se generan a partir de un algoritmo determinista) y cuasi-aleatorios (que vienen determinados por una

secuencia regular).

Paquetes estadísticos, como **R Studio**, utilizan números pseudo-aleatorios procedentes de una uniforme estándar para generar números aleatorios de cualquier otra distribución. En el siguiente capítulo veremos cómo. De aquí nace, principalmente, la importancia de saber si una muestra dada se distribuye uniformemente en el intervalo  $[0,1]$  o no.

Para ver si, efectivamente una muestra procede de una uniforme estándar, recurriremos a la inferencia estadística, más concretamente a los contrastes de hipótesis, y en nuestro caso, pruebas de bondad de ajuste.

## 1.2. Objetivos

Los tests de bondad de ajuste son usados para determinar si una muestra proviene de cierta distribución. Su hipótesis nula viene determinada por suponer que la función de distribución empírica de nuestra muestra se ajusta a la función de distribución teórica que queremos estudiar. En nuestro caso nos interesaría resolver esos contrastes, donde la hipótesis nula compare la función de distribución empírica con la distribución uniforme en el intervalo  $[0,1]$ .

El principal objetivo de este trabajo será introducir y comparar tanto tests generales de bondad de ajuste aplicados a uniformidad, como tests específicamente diseñados para la uniformidad. Para ello se realizará, en el último capítulo, una experimentación basándose en las potencias de varios tests introducidos previamente con varias familias de alternativas.

### 1.3. Estructura del trabajo

Este trabajo se ha distribuido de la siguiente manera:

- **Capítulo 1: Introducción**

Donde se presenta la importancia de este trabajo y los objetivos que se esperan conseguir.

- **Capítulo 2: La distribución uniforme**

Capítulo dedicado a la definición y casuística de la familia uniforme. Se introduce la distribución uniforme discreta y continua, hablando de la estimación de sus parámetros y de diferentes estimadores aplicados a esta distribución. Por último, se particulariza al caso de la distribución uniforme estándar, la uniforme en el intervalo  $[0,1]$ . Se dan una serie de resultados importantes para nuestro trabajo.

- **Capítulo 3: Pruebas para la uniformidad en el intervalo unidad**

En este capítulo se tratan tanto pruebas gráficas que ayudan a intuir la uniformidad, como tests de bondad de ajuste para detectarla. Se introducen tres métodos gráficos y catorce tests de uniformidad, entre ellos seis son comúnmente conocidos, se les llamará tests tradicionales, y otros ocho que están diseñados específicamente para testear uniformidad.

- **Capítulo 4: Experimentación**

Se compararán los tests introducidos en el Capítulo 3 haciendo un análisis de la potencia. Les proporcionaremos diversas distribuciones alternativas, haciendo variar los parámetros en cada una y tomando diferentes tamaños de muestra. Se darán unas breves conclusiones para cada tabla presentada.

- **Capítulo 5: Conclusiones**

Finalmente se hará un resumen de las conclusiones obtenidas en el capítulo anterior, determinando así las mejores y peores características de cada test.

## Capítulo 2

# La distribución uniforme

Una de las distribuciones más simples, pero a la vez más útiles en el análisis estadístico, es la distribución uniforme. Se entiende por distribución uniforme la que representa un fenómeno en el cual todos los valores de un conjunto (ya sea discreto o continuo) son igualmente probables de suceder.

A partir de este punto del trabajo, a cualquier variable aleatoria que sea uniforme (o supongamos de partida que es uniforme) la denotaremos por  $U$ .

### 2.1. Distribución uniforme discreta

Como se ha introducido antes, una variable discreta  $U$  que tome valores en un conjunto finito  $\{u_1, u_2, \dots, u_n\}$  y asigne la misma probabilidad a todos ellos se dirá que sigue una distribución uniforme discreta. Asumimos, por simplicidad que,  $u_1 < u_2 < \dots < u_n$ . Esta distribución uniforme suele denotarse por  $\mathcal{U}_{\{u_1, \dots, u_n\}}$ .

Ahora, sabiendo que todos los sucesos son equiprobables, se define la función de masa de probabilidad de  $U$  como

$$P(U = u) = \frac{1}{n} \quad \forall u \in \{u_1, u_2, \dots, u_n\}$$

Además, puede definirse la función de distribución (acumulada) (CDF, por sus siglas del inglés: Cumulative Distribution Function) para todo  $k \in \{1, \dots, n\}$

$$F_U(u_k) = P(U \leq u_k) = \sum_{u_k \leq u} P(U = u_k) = \sum_{i=1}^k P(U = u_i) = \frac{k}{n}$$

Asímismo, definimos momentos como la esperanza

$$E(U) = \sum_{i=1}^n u_i P(U = u_i) = P(U = u) \sum_{i=1}^n u_i = \frac{1}{n} \sum_{i=1}^n u_i$$

o la varianza

$$\begin{aligned} Var(U) &= \sum_{i=1}^n (u_i - E(U))^2 P(U = u_i) = P(U = u) \sum_{i=1}^n (u_i - E(U))^2 \\ &= \frac{1}{n} \sum_{i=1}^n (u_i - E(U))^2 \end{aligned}$$

Uno de los ejemplos más recurrentes por cualquier profesor de Estadística y Probabilidad es un dado perfecto de 6 caras. El soporte de la variable serían los números  $\{1, 2, 3, 4, 5, 6\}$ , con probabilidad de obtener cualquier cara igual a  $1/6$ . La esperanza y varianza en este experimento serían:

$$\begin{aligned} E(U) &= \frac{1}{6} \sum_{i=1}^6 u_i = \frac{21}{6} \\ Var(U) &= \frac{1}{6} \sum_{i=1}^6 \left(u_i - \frac{21}{6}\right)^2 \approx 2.916 \end{aligned}$$

## 2.2. Distribución uniforme continua

También podemos decir que una variable continua  $U$  sigue una distribución uniforme en el intervalo  $[a, b]$  si reparte la probabilidad igualitariamente (o uniformemente) a lo largo de ese intervalo, con  $a$  y  $b \in \mathbb{R}$  cualesquiera pero fijos. Es decir, la densidad tendrá un valor constante en  $[a, b]$ , dependiendo dicho valor de la longitud del intervalo. Se denotará simplemente

como  $\mathcal{U}(a, b)$ .

De esta interpretación podemos deducir su función de densidad (PDF, por sus siglas del inglés: Probability Density Function):

$$f(u) = \begin{cases} \frac{1}{b-a} & \text{si } u \in [a, b] \\ 0 & \text{si } u \notin [a, b] \end{cases}$$

así como su función de distribución, integrando  $f(u)$  en  $(-\infty, u)$ :

$$F(u) = \begin{cases} 0 & \text{si } u < a \\ \frac{u-a}{b-a} & \text{si } a \leq u < b \\ 1 & \text{si } u \geq b \end{cases}$$

Para valores de  $u$  entre  $a$  y  $b$ , podemos interpretarla aritméticamente como el área del rectángulo de base  $u - a$  y altura  $\frac{1}{b-a}$ , tal y como aparece en la Figura 2.1.

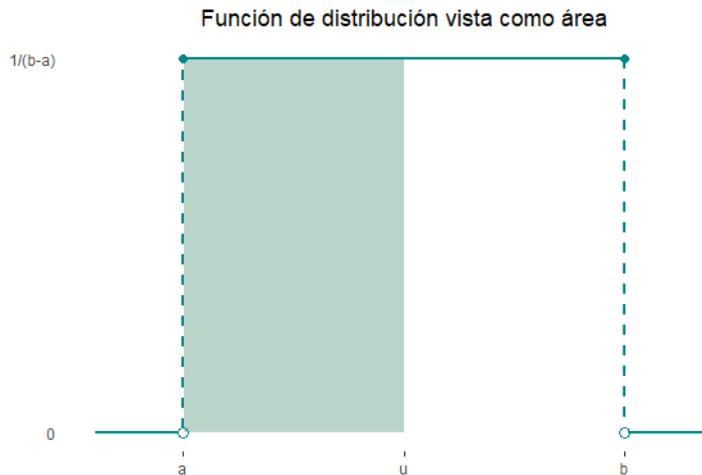


Figura 2.1: Función de distribución de una  $\mathcal{U}(a, b)$  vista como área del rectángulo.

A partir de su definición, podemos calcular la esperanza de  $U$

$$E(U) = \int_a^b \frac{1}{b-a} \cdot u \, du = \frac{1}{b-a} \left[ \frac{u^2}{2} \right]_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{(b-a)(b+a)}{2(b-a)} = \frac{b+a}{2}$$

Efectivamente, parece lógico que, si se reparte uniformemente la probabilidad en  $[a, b]$ , la esperanza sea el punto medio del intervalo, es decir  $\frac{b+a}{2}$ .

Por otro lado, la varianza de una variable uniforme sería

$$\begin{aligned} Var(U) &= \int_a^b (u - E(U))^2 \cdot \frac{1}{b-a} du = \int_a^b \left(u - \frac{b+a}{2}\right)^2 \cdot \frac{1}{b-a} du \\ &= \frac{1}{3(b-a)} \left[ \left(u - \frac{b+a}{2}\right)^3 \right]_a^b = \left[ \frac{u^3}{3} + \frac{(-b-a)u^2}{2} + \frac{(b+a)^2 u}{4} \right]_a^b \\ &= \frac{\frac{b^3+3a^2b}{12} - \frac{3ab^2+a^3}{12}}{b-a} = \frac{(b-a)^2}{12} \end{aligned}$$

Se puede demostrar que la probabilidad de que  $U \sim \mathcal{U}(a, b)$  pertenezca a un subintervalo de  $[a, b]$  es independiente de la ubicación de este intervalo, pero dependiente de la longitud del mismo.

En efecto, sea  $[c, c + \varepsilon]$  un subintervalo de  $[a, b]$  con  $\varepsilon > 0$  fijo, se tiene

$$P(U \in [c, c + \varepsilon]) = \int_c^{c+\varepsilon} \frac{1}{b-a} du = \frac{c + \varepsilon}{b-a} - \frac{c}{b-a} = \frac{\varepsilon}{b-a}$$

lo cual no depende de  $c$ , pero sí de la longitud  $\varepsilon$ . ■

### 2.2.1. Estimación paramétrica en la distribución uniforme continua

Existen ocasiones en las que no se conocen los extremos del intervalo. En estos casos, podemos aplicar la inferencia estadística para estimarlos puntualmente, donde los dos métodos más comunes son el método de los momentos, y el método de máxima verosimilitud.

#### Estimador de $\theta$ por el método de los momentos

##### Caso general

Es uno de los métodos de estimación más antiguos y más sencillos. Sea  $X$

una variable aleatoria con distribución  $F_\theta$ , con  $\theta \in \Theta \subset \mathbb{R}^k$ .

Se definen sus  $k$ -ésimos momentos poblacionales como

$$\mu_k = E(X^k) = \begin{cases} \int_{\mathbb{R}} x^k f(x) dx & \text{si } X \text{ es continua} \\ \sum_i x_i^k P(x_i) & \text{si } X \text{ es discreta} \end{cases} \quad (2.1)$$

Tomamos ahora una m.a.s.  $(X_1, X_2, \dots, X_n)$  y definimos sus  $k$ -ésimos momentos muestrales como

$$m_k = \frac{1}{n} \sum_{i=1}^n X_i^k \quad (2.2)$$

Este método consiste en igualar los primeros momentos muestrales y poblacionales tantas veces como parámetros necesitemos estimar, así tendremos un sistema con mismo número de ecuaciones que de incógnitas. En nuestro caso, como mucho tendremos dos incógnitas (los dos extremos del intervalo).

### Caso uniforme

Sea  $U \sim \mathcal{U}(\theta_1, \theta_2)$  una variable continua, donde  $\theta = (\theta_1, \theta_2) \in \Theta \subset \mathbb{R}^2$ , con  $\theta_1 < \theta_2$ , y sea  $(U_1, U_2, \dots, U_n)$  una m.a.s. suya. Calculamos los dos primeros momentos poblacionales

$$\begin{aligned} \mu_1 &= E(U) = \frac{\theta_2 + \theta_1}{2} \\ \mu_2 &= E(U^2) = \int_{\theta_1}^{\theta_2} \frac{u^2}{\theta_2 - \theta_1} du = \frac{1}{\theta_2 - \theta_1} \left[ \frac{u^3}{3} \right]_{\theta_1}^{\theta_2} \\ &= \frac{(\theta_2)^3 - (\theta_1)^3}{3(\theta_2 - \theta_1)} = \frac{(\theta_1)^2 + (\theta_2)^2 + \theta_1\theta_2}{3} \end{aligned}$$

Calculamos también los dos primeros momentos muestrales

$$\begin{aligned} m_1 &= \frac{1}{n} \sum_{i=1}^n u_i = \bar{U} \\ m_2 &= \frac{1}{n} \sum_{i=1}^n u_i^2 = \overline{U^2} \end{aligned}$$

y planteamos el sistema

$$\begin{cases} \mu_1 = m_1 \\ \mu_2 = m_2 \end{cases} \iff \begin{cases} \frac{\theta_2 + \theta_1}{2} = \bar{U} \\ \frac{(\theta_1)^2 + (\theta_2)^2 + \theta_1 \theta_2}{3} = \bar{U}^2 \end{cases}$$

Por sustitución, tomando  $\theta_1 = 2\bar{U} - \theta_2$  y sustituyendo en la segunda ecuación, llegamos a

$$\begin{aligned} (2\bar{U} - \theta_2)^2 + \theta_2^2 + \theta_2(2\bar{U} - \theta_2) &= 3\bar{U}^2 \iff \\ \theta_2^2 - 2\bar{U}\theta_2 + 4\bar{U}^2 - 3\bar{U}^2 &= 0 \end{aligned}$$

por lo que

$$\begin{aligned} \theta_2 &= \frac{2\bar{U} \pm \sqrt{4\bar{U}^2 - 16\bar{U}^2 + 12\bar{U}^2}}{2} = \frac{2\bar{U} \pm \sqrt{12\bar{U}^2 - 12\bar{U}^2}}{2} \\ &= \frac{2\bar{U} \pm \sqrt{12S^2}}{2} = \bar{U} \pm S\sqrt{3} \end{aligned}$$

y así

$$\theta_1 = 2\bar{U} - \bar{U} \pm S\sqrt{3} = \bar{U} \pm S\sqrt{3} = \theta_2$$

donde  $S = \sqrt{\frac{\sum_{i=1}^n (u_i - \bar{U})^2}{n}}$  es la desviación típica muestral.

Teniendo en cuenta que  $\theta_1 \neq \theta_2$  y que  $\theta_1 < \theta_2$ , obtendríamos

$$\theta_1 = \bar{U} - S\sqrt{3}, \quad \theta_2 = \bar{U} + S\sqrt{3}.$$

## Estimador máximo veosímil de $\theta$

### Caso general

Sea  $X$  una variable aleatoria, y sea  $(X_1, X_2, \dots, X_n)$  una m.a.s. tal que su distribución pertenece a  $\mathcal{F}_\theta = \{F_\theta, \theta \in \Theta \subset \mathbb{R}^k\}$ . Sea además,  $\vec{x} = (x_1, x_2, \dots, x_n)$  una realización muestral de la m.a.s. Se define *la función de verosimilitud asociada a dicha realización muestral* como la probabilidad o densidad conjunta de la muestra, denotada por  $L(\vec{x}; \theta) = f_\theta(\vec{x}) = \prod_{i=1}^n f(x_i; \theta)$ , donde  $\vec{x}$  es aleatorio pero fijo, siendo entonces  $\theta$  la variable de

la función.

Un estimador  $\hat{\theta}(\vec{x})$  se dice *estimador máximo verosímil* de  $\theta$  si

$$L(\vec{x}; \hat{\theta}(\vec{x})) = \sup_{\theta \in \Theta} L(\vec{x}; \theta) \quad (2.3)$$

y suele abreviarse por las siglas *EMV*.

Estos estimadores gozan de unas propiedades aconsejables a la hora de estimar, como por ejemplo

- **Invarianza.** Si  $\hat{\theta}$  es el *EMV* de  $\theta$ , entonces  $h(\hat{\theta})$  es *EMV* de  $h(\theta)$ , con  $h$  una función medible biyectiva.
- **Consistencia.** Sea  $(X_1, \dots, X_n)$  una m.a.s. de una variable  $X$ , que cumpla las condiciones de regularidad de Fréchet-Cramér-Rao (FCR) y alguna otra condición general, si  $\hat{\theta}_n = \hat{\theta}(X_1, \dots, X_n)$  es *EMV* de  $\theta$ , se tiene que  $\hat{\theta}_n \xrightarrow[n \rightarrow \infty]{} \theta$ . Formalmente, las condiciones de regularidad de Fréchet-Cramér Rao son:

1. el espacio paramétrico  $\Theta$  es un intervalo abierto;
2. el soporte de la variable,  $\mathbb{X}$ , no depende del parámetro  $\theta$ ;
3. la función de probabilidad o densidad conjunta de  $(X_1, \dots, X_n)$  definida por  $f(x_1, \dots, x_n; \theta) = f(x_1; \theta) \cdot \dots \cdot f(x_n; \theta)$  es dos veces derivable para todo  $\theta \in \Theta$  y para casi toda muestra de observaciones  $(x_1, \dots, x_n)$ , y esas derivadas son integrables en  $\mathbb{X}^n$
4. para todo  $\theta \in \Theta$  se tiene, si  $X$  es una variable continua

$$\begin{aligned} & \int_{\mathbb{X}} \dots \int_{\mathbb{X}} \frac{\partial}{\partial \theta} f(x_1, \dots, x_n; \theta) dx_1 \dots dx_n \\ &= \frac{\partial}{\partial \theta} \int_{\mathbb{X}} \dots \int_{\mathbb{X}} f(x_1, \dots, x_n; \theta) dx_1 \dots dx_n \end{aligned}$$

y que

$$\begin{aligned} & \int_{\mathbb{X}} \cdots \int_{\mathbb{X}} T(x_1, \dots, x_n) \cdot \frac{\partial}{\partial \theta} f(x_1, \dots, x_n; \theta) dx_1 \dots dx_n \\ &= \frac{\partial}{\partial \theta} \int_{\mathbb{X}} \cdots \int_{\mathbb{X}} T(x_1, \dots, x_n) \cdot f(x_1, \dots, x_n; \theta) dx_1 \dots dx_n \end{aligned}$$

Si  $X$  fuera discreta se reemplaza la integral por un sumatorio.

- **Insegadez asintótica.** Bajo ciertas condiciones generales (FCR), siendo  $\hat{\theta}_n$  *EMV* de  $\theta$ , se tiene que  $E[\hat{\theta}_n] \xrightarrow{n \rightarrow \infty} \theta$ .

- **Normalidad asintótica.** Bajo las condiciones ya expuestas en puntos anteriores, se cumple que

$$\frac{\hat{\theta}_n - \theta}{1/\sqrt{I(X_1, \dots, X_n; \theta)}} \xrightarrow{n \rightarrow \infty} \mathcal{N}(0, 1).$$

donde  $I(X_1, \dots, X_n; \theta) = E \left[ \left( \frac{\partial}{\partial \theta} \log f(X_1, \dots, X_n; \theta) \right)^2 \right]$  es la cantidad de información de Fisher.

Hay casos donde encontrar el supremo de  $L(\vec{x}; \theta)$  puede ser realmente complicado, así, en muchas ocasiones se toma la transformación logarítmica de  $L$ , es decir, maximizaremos  $\log L(\vec{x}, \theta)$ .

Esto puede hacerse ya que, por ser el logaritmo neperiano una función estrictamente creciente, el máximo de ambas funciones coincide.

### Caso uniforme

Sea  $U \sim \mathcal{U}(\theta_1, \theta_2)$  una variable continua, donde  $\theta = (\theta_1, \theta_2) \in \Theta \subset \mathbb{R}^2$ , con  $\theta_1 < \theta_2$ , y sea  $(U_1, U_2, \dots, U_n)$  una m.a.s. suya. A partir de esta, obtenemos  $\vec{u} = (u_1, u_2, \dots, u_n)$  una realización muestral. Entonces, hallamos la función de verosimilitud:

$$L(\vec{u}; \theta) = \prod_{i=1}^n f(u_i; \theta) = \prod_{i=1}^n \frac{1}{\theta_2 - \theta_1} \mathbb{I}_{\{\theta_1 < u_i < \theta_2\}} = \frac{1}{(\theta_2 - \theta_1)^n} \mathbb{I}_{\{u_{(1)} > \theta_1\}} \mathbb{I}_{\{u_{(n)} < \theta_2\}} \quad (2.4)$$

donde  $u_{(1)}$  y  $u_{(n)}$  son el menor y el mayor elemento de la realización muestral, respectivamente. Los estadísticos de orden se denotan con letras mayúsculas, luego los denotaremos por  $U_{(1)}$  y  $U_{(n)}$

En el caso uniforme, no haría falta ni calcular la transformación logarítmica ni hallar su máximo, el *EMV* se deduce estudiando el espacio paramétrico.

Se observa que la función de la Ecuación (2.4), como función de  $\theta_1$ , es estrictamente creciente, y se cumple que  $u_{(1)} > \theta_1$ . Resulta claro que el valor que hace máxima esa función sería  $\hat{\theta}_1 = U_{(1)}$ , es decir, el mínimo de la muestra.

Análogamente, la función de la Ecuación (2.4), como función de  $\theta_2$ , es estrictamente decreciente. Así, entre todos los valores del parámetro que cumplen  $u_{(n)} < \theta_2$ , el supremo se alcanza en  $\hat{\theta}_2 = U_{(n)}$ , es decir, el máximo de la muestra.

Luego, el estimador máximo verosímil de  $\theta = (\theta_1, \theta_2)$  serán los estadísticos ordenados de orden 1 y  $n$  respectivamente

$$\hat{\theta} = (U_{(1)}, U_{(n)}) \quad (2.5)$$

El *EMV* para el caso uniforme cumple la propiedad de invarianza bajo transformaciones del parámetro, pero no las propiedades asintóticas vistas en el caso general. Esto es debido a que algunas de las condiciones de regularidad fallan, por ejemplo la dependencia del soporte de la variable del parámetro a estimar.

### 2.3. Distribución uniforme estándar

La distribución uniforme estándar es un caso particular que hace referencia a la distribución uniforme continua en el intervalo  $[0,1]$ .

Si  $U \sim \mathcal{U}(0, 1)$ , se tiene que su función de densidad

$$f(u) = \begin{cases} 1 & \text{si } u \in [0, 1] \\ 0 & \text{si } u \notin [0, 1] \end{cases}$$

coincide con la función indicador en el  $[0,1]$ , denotada por  $\mathbb{I}_{[0,1]}$ .

Así, su función de distribución es

$$F(u) = \begin{cases} 0 & \text{si } u < 0 \\ u & \text{si } 0 \leq u < 1 \\ 1 & \text{si } u \geq 1 \end{cases}$$

Por tanto, su esperanza siempre será

$$E(U) = \frac{1}{2}$$

y su varianza

$$\text{Var}(U) = \frac{1}{12}$$

Esta distribución tiene varias propiedades importantes para la estadística, en concreto, algunos resultados útiles para este trabajo son los siguientes.

**Proposición 2.1.** Si  $U \sim \mathcal{U}(0, 1)$ , la variable  $1 - U \sim \mathcal{U}(0, 1)$ .

*Demostración:*

Sea  $U \sim \mathcal{U}(0, 1)$ , y sea  $V = 1 - U$ . Entonces su función de distribución viene dada por

$$\begin{aligned} F_V(v) &= P(V \leq v) = P(1 - U \leq v) = P(U \geq 1 - v) = 1 - P(U \leq 1 - v) \\ &= 1 - F_U(1 - v) = \begin{cases} 1 - 0 & \text{si } 1 - v < 0 \\ 1 - (1 - v) & \text{si } 0 \leq 1 - v < 1 \\ 1 - 1 & \text{si } 1 - v \geq 1 \end{cases} = \begin{cases} 1 & \text{si } v > 1 \\ v & \text{si } 0 < v \leq 1 \\ 0 & \text{si } v \leq 0 \end{cases} \end{aligned}$$

que coincide con la de una uniforme estándar. ■

Como se ha introducido en la Motivación de este trabajo, las muestras procedentes de una uniforme estándar tienen una importancia destacada en la simulación de distribuciones discretas o continuas. El resultado que garantiza que esto sea posible es el siguiente, en el caso continuo.

**Proposición 2.2.** *Sea  $X$  una variable aleatoria absolutamente continua con función de distribución  $F_X(x)$ , entonces la variable  $U = F_X(X)$  tiene distribución  $\mathcal{U}(0, 1)$ .*

*Demostración:*

Es claro que  $U$  solo tomará valores en  $[0, 1]$ , ya que está definida como la función de distribución de otra variable, es decir, una probabilidad.

Si  $F_X(x)$  es estrictamente creciente, la inversa de su función de distribución está bien definida y viene dada por

$$F_X^{-1}(u) = x \iff F_X(x) = u$$

Sin embargo, nada nos asegura que esto sea así. En caso de tener una función que sea constante en algún intervalo  $[x_1, x_2] \subset [0, 1]$ , la inversa  $F_X^{-1}$  no estaría bien definida ya que, cualquier  $x$  tal que  $x_1 \leq x \leq x_2$  verifica que  $F_X(x) = u$ . Para evitar este problema se define, para  $0 < u < 1$

$$F_X^{-1}(u) = \inf_{x \in [0, 1]} \{F_X(x) \geq u\} \quad (2.6)$$

Esta definición sirve también para el caso en el que  $F_X$  sea estrictamente creciente, así, usando esta definición se tiene

$$\begin{aligned} P(U \leq u) &= P(F_X(X) \leq u) \\ &= P(F_X^{-1}[F_X(X)] \leq F_X^{-1}(u)) \\ &= {}^1P(X \leq F_X^{-1}(u)) \\ &= F_X(F_X^{-1}(u)) \\ &= {}^2u \end{aligned}$$

Claramente tenemos que  $P(U \leq u) = 0$  para  $u \leq 0$  y  $P(U \leq u) = 1$  si  $u \geq 1$ , demostrando así que  $U$  sigue una distribución uniforme estándar. ■

Al igual que podemos obtener una distribución uniforme estándar partiendo de cualquier distribución continua simplemente conociendo su función de distribución, podemos hacer el proceso inverso. Si partimos de una variable uniforme en el intervalo unidad, podemos obtener cualquier distribución continua a través de la inversa de su función de distribución o función cuantil.

Para el caso discreto podríamos hacer algo análogo. Supongamos que queremos generar una muestra de una v.a. discreta  $X$  con función de masa de probabilidad  $P(X = x_j) = p_j$ , con  $j \in \{0, 1, 2, \dots\}$ . Generamos un número aleatorio  $u$  procedente de una uniforme estándar, y se define

$$X = \begin{cases} x_0 & \text{si } u < p_0 \\ x_1 & \text{si } p_0 \leq u < p_0 + p_1 \\ \vdots & \\ x_j & \text{si } \sum_{i=0}^{j-1} p_i \leq u < \sum_{i=0}^j p_i \\ \vdots & \end{cases}$$

lo que es equivalente a hallar la inversa de  $P(u)$ .

Como  $u$  procede de una distribución  $\mathcal{U}(0, 1)$ ,  $P(a \leq u < b) = b - a$ ,

---

<sup>1</sup>En caso de que  $F_X$  sea estrictamente creciente siempre es cierto. Si  $F_X$  es plana en algún intervalo, hay que tener en cuenta la definición de la Ecuación (2.6).

<sup>2</sup>Por continuidad de  $F_X$ .

entonces

$$P(X = x_j) = P\left(\sum_{i=0}^{j-1} p_i \leq u < \sum_{i=0}^j p_i\right) = \sum_{i=0}^j p_i - \sum_{i=0}^{j-1} p_i = p_j$$

y así  $X$  sigue la distribución que deseábamos.

Esto es lo que hacen paquetes estadísticos, como **R**, para la generación de números pseudo-aleatorios. Solo necesitan tener implementada la función de distribución o probabilidad (y su inversa, la función cuantil) de la distribución que queramos generar y ser capaces de producir números aleatorios distribuídos uniformemente en el intervalo  $[0,1]$ .

Lo que hace intrínsecamente el programa es aplicarle la función cuantil (o la generación vista en el caso discreto) a los valores uniformes generados, y así crear una muestra de números aleatorios con la distribución elegida.

## Capítulo 3

# Pruebas para la uniformidad en el intervalo unidad

A la hora de analizar la uniformidad en una muestra se puede recurrir a diferentes pruebas que nos facilitan ver dicha distribución.

### 3.1. Pruebas gráficas

Para realizar dichos análisis, existen pruebas gráficas que pueden ilustrar la supuesta uniformidad desde una perspectiva visual. Algunos de los gráficos más útiles serán descritos a continuación.

#### 3.1.1. Histogramas

Gráficos que agrupan los datos en una serie de subintervalos del intervalo principal (en nuestro caso el intervalo  $[0,1]$ ) y muestran la frecuencia de cada subintervalo. Si nuestros datos siguen una distribución uniforme, las frecuencias de cada subintervalo deberían ser iguales o parecidas.

Así, si al realizar un histograma con nuestros datos obtenemos algo parecido a la Figura 3.1, podemos intuir, sin suficiente rigor, que nuestra muestra sigue una distribución uniforme.

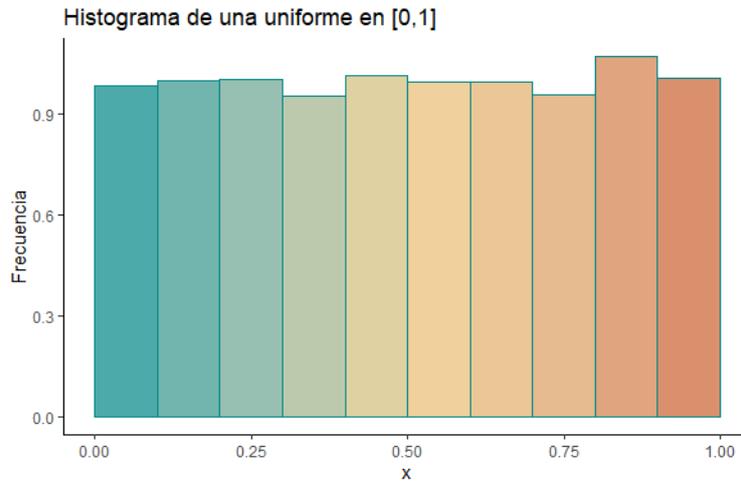


Figura 3.1: Histograma de una distribución uniforme en  $[0,1]$ .

### 3.1.2. Diagramas de cajas

Son representaciones visuales en forma de rectángulos y unos bigotes horizontales a los lados, que llegan a los valores máximo y mínimo de la muestra. En las cajas aparecen el primer cuartil, en el borde izquierdo, la mediana, dentro de la caja, y el tercer cuartil, en el borde derecho.

Si los datos provienen de una uniforme estándar el valor mínimo no será menor que 0, ni el máximo mayor que 1. El primer cuartil rondará el valor 0.25, la mediana coincide con la media en el 0.5 y el tercer cuartil será 0.75, tal y como se muestra en la Figura 3.2.

### 3.1.3. Gráficos Q-Q

Con estos gráficos podemos estudiar cuán se aproxima la distribución de un conjunto de datos a una distribución teórica, en este caso la uniforme estándar. Para ello, se comparan los cuantiles teóricos de la distribución frente a los cuantiles empíricos de los datos, graficando una línea recta si estos fueran completamente iguales. Esta línea aparece en el fondo del gráfico a modo de control, nuestros cuantiles deberían ajustarse a ella en caso de

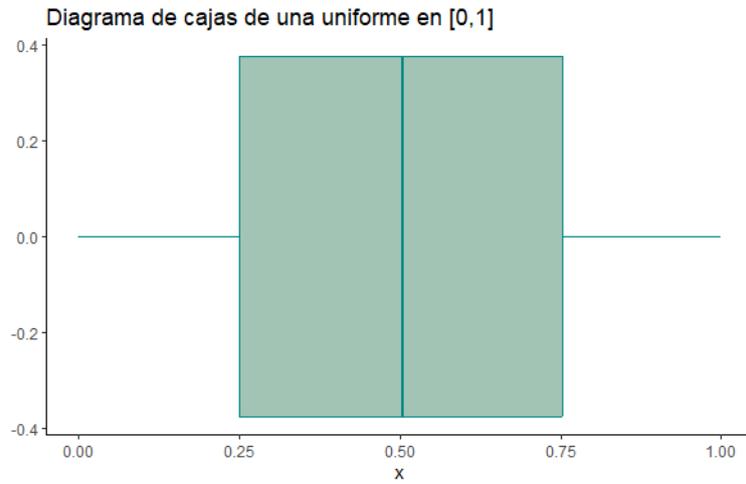


Figura 3.2: Diagrama de cajas de una distribución uniforme en  $[0,1]$ .

tener igualdad de distribuciones. En la Figura 3.3 aparece el gráfico Q-Q habiéndole dado al programa una muestra uniforme en  $[0,1]$ .

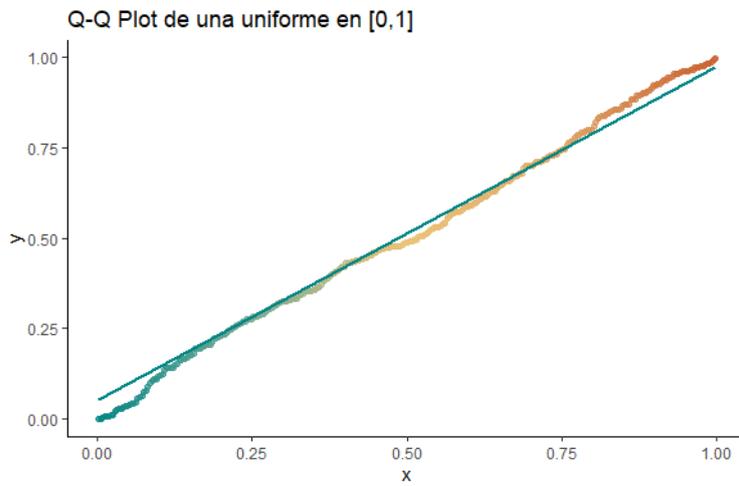


Figura 3.3: Gráfico Q-Q de una distribución uniforme en  $[0,1]$ .

## 3.2. Tests de uniformidad

Además de las pruebas gráficas, con el fin de proporcionar resultados más certeros, desde el siglo XX varios matemáticos como A. Ronald Fisher [7] o Egon Pearson [12] han creado tests específicos para este tipo de contrastes cuya hipótesis nula ( $H_0$ ) es comprobar si un conjunto de datos se distribuye uniformemente en el intervalo  $[0,1]$ .

Existe una variedad muy amplia de tests para la uniformidad debido a la aplicación tan frecuente de la distribución uniforme. Esto puede suponer un problema de indecisión a la hora de escoger el mejor test.

**Definición 3.1.** [Test de hipótesis. ] Los tests de hipótesis son una serie de reglas que nos permiten, a raíz de unas evidencias generadas, aceptar o rechazar una sentencia sobre una población. Dicha sentencia se llamará hipótesis, que se dirá *paramétrica* si hace referencia a un parámetro de la población (una proporción, una media, etc.). Además dichas hipótesis podrán ser *simples*, si determinan completamente una distribución o *compuestas*, en otro caso.

Un test o contraste de hipótesis consta de una *Hipótesis Nula* ( $H_0$ ), la sentencia a contrastar y que se rechazará si existen suficientes evidencias para hacerlo, y de una *Hipótesis Alternativa* ( $H_1$ ), que generalmente niega  $H_0$ . Formalmente, se define un test de hipótesis puro como cualquier partición medible del espacio muestral ( $\mathcal{X}$ ) que viene caracterizada por la función test  $\varphi : \mathcal{X}^n \rightarrow \mathbb{R}$  tal que

$$\varphi(\vec{x}) = \begin{cases} 1 & \text{si } \vec{x} \in \text{R.C.} \\ 0 & \text{si } \vec{x} \in \text{R.A.} \end{cases}$$

donde la *Región Crítica* (R.C.), es la región de la partición en la que se encuentran los resultados muestrales para los que se rechaza  $H_0$ , y la *Región Alternativa* (R.A.), donde no se rechaza  $H_0$ .

**Definición 3.2.** [Tests de bondad de ajuste. ] Son tests de hipótesis específicos, cuyo objetivo es determinar si unos datos se ajustan o no a una distribución concreta. Si la distribución viene completamente determinada se dirá un test de hipótesis nula simple, con hipótesis nula

$$H_0 : X \equiv F_0$$

O si la distribución depende de algún parámetro desconocido, será de hipótesis nula compuesta, es decir

$$H_0 : X \equiv F_\theta \quad \theta \in \Theta$$

**Definición 3.3** [Test de uniformidad. ] Son tests de bondad de ajuste donde la distribución de interés será una uniforme. En este trabajo daremos más importancia a los tests de hipótesis simples, es decir, donde se conozcan los extremos de los intervalos. De hecho, en este capítulo solo trataremos la uniforme en el intervalo unidad, a continuación entenderemos por qué.

De aquí en adelante, introduciremos los tests que se compararán según su potencia en el siguiente capítulo. Primeramente se definirán los tests más tradicionales de bondad de ajuste, aplicados a la uniformidad, y posteriormente se verán tests más innovadores que surgieron a raíz de la necesidad de detectar la uniformidad, es decir, específicos para esta distribución.

### **3.3. Tests tradicionales adaptados a la uniformidad en $[0,1]$**

#### **3.3.1. Test de la Razón de Verosimilitudes**

Es una prueba que se encarga de medir la razón entre la función de verosimilitud definida en la Ecuación (2.3) y la función de verosimilitud

cuando  $\theta \in \Theta_0$ , es decir, bajo la hipótesis nula. Su estadístico viene definido por el cociente entre ellas.

$$\Lambda = \frac{\sup_{\theta \in \Theta_0} L(\vec{x}; \theta)}{\sup_{\theta \in \Theta} L(\vec{x}; \theta)}$$

donde  $\Theta_0$  es el espacio paramétrico bajo  $H_0$ .

Si queremos aplicar este test en la familia uniforme  $\mathcal{U}(\theta_1, \theta_2)$  para la hipótesis

$$H_0 \equiv (\theta_1, \theta_2) = (0, 1)$$

Tomamos una m.a.s.  $(U_1, \dots, U_n)$ , y por lo concluído anteriormente, en la Ecuación (2.5), el *EMV* viene dado por los estadísticos de orden  $(U_{(1)}, U_{(n)})$ , así

$$\begin{aligned} \Lambda &= \frac{\sup_{\theta \in \Theta_0} L(\vec{u}; \theta)}{\sup_{\theta \in \Theta} L(\vec{u}; \theta)} = \frac{\mathbb{I}_{U_{(1)} \geq 0} \mathbb{I}_{U_{(n)} \leq 1}}{\frac{1}{(U_{(n)} - U_{(1)})^n} \mathbb{I}_{U_{(1)} \geq U_{(1)}} \mathbb{I}_{U_{(n)} \leq U_{(n)}}} \\ &= (U_{(n)} - U_{(1)})^n \mathbb{I}_{U_{(1)} \geq 0} \mathbb{I}_{U_{(n)} \leq 1} \end{aligned}$$

De esta manera podremos definir la región crítica del test como

$$RC_\alpha = \{\vec{u} \mid U_{(1)} < 0\} \cup \{\vec{u} \mid U_{(n)} > 1\} \cup \{U_{(n)} - U_{(1)} < k\}$$

Puesto que el cálculo de estas dos constantes involucra el cálculo de la distribución conjunta de  $U_{(1)}$  y  $U_{(n)}$ , optamos por una opción más sencilla en la que expresamos la región crítica del test directamente como función de  $\Lambda$ :

$$RC_\alpha = \{\vec{u} \mid \Lambda < c\}$$

Procedemos a calcular el valor de la constante  $c$  por simulación, y utilizando muestras de tamaños  $n \in \{25, 100\}$  para unos niveles de significación  $\alpha \in \{0.1, 0.05, 0.01\}$  se obtienen dichos cuantiles en la Tabla 3.1.

$\alpha \backslash n$	0.1	0.05	0.01
25	0.01897745	0.007933056	0.001157525
100	0.02005068	0.008486622	0.00127456

Tabla 3.1: Cuantiles para la región crítica del TRV.

### 3.3.2. Test de Kolmogorov-Smirnov

Propuesto por Andréi Kolmogorov y Nikolai Smirnov [9, 20], es un test cuyo estadístico mide la diferencia vertical entre la función de distribución empírica  $F_n$  de nuestros datos y la función de distribución teórica de una uniforme en  $[0,1]$ .

Sea  $U_1, \dots, U_n$  una muestra aleatoria simple, y sea  $U_{(1)}, \dots, U_{(n)}$  los estadísticos de orden tal que  $U_{(1)} < U_{(2)} < \dots < U_{(n)}$ . Se define el estadístico del test de Kolmogorov-Smirnov de esta manera:

$$D_n = \sup_{u \in \mathbb{R}} |F_n(u) - F(u)| = \max(D_n^+, D_n^-) \quad (3.1)$$

donde  $D_n^+$  mide la diferencia para las  $u \in \mathbb{R}$  tal que  $F_n(u) > F(u)$ , y  $D_n^-$  se centra en las  $u \in \mathbb{R}$  que hacen que  $F_n(u) < F(u)$ . Formalmente, se definen así:

$$D_n^+ = \max_{1 \leq i \leq n} \left\{ \frac{i}{n} - U_{(i)} \right\} \quad D_n^- = \max_{1 \leq i \leq n} \left\{ U_{(i)} - \frac{i-1}{n} \right\}$$

Es claro que este test rechaza  $H_0$  cuando  $D_n$  toma valores grandes (lejanos de 0), es decir,  $RC_\alpha = \{\bar{u} \mid D_n > c\}$ , y además para un tamaño de muestra grande el estadístico, bajo  $H_0$ , toma valores próximos a 0 (como resultado del Teorema de Glivenko-Canteli [21]).

Se ha simulado la distribución del estadístico por Monte Carlo [22] para encontrar las constantes que determinan la región crítica para distintos

tamaños muestrales y distintos niveles de significación, en concreto para  $n \in \{25, 100\}$  y  $\alpha \in \{0.1, 0.05, 0.01\}$ . Estos valores vienen reflejados en la Tabla 3.2.

$\alpha \backslash n$	0.1	0.05	0.01
25	0.2378846	0.2646243	0.3151940
100	0.1200320	0.1335213	0.1603922

Tabla 3.2: Cuantiles para la región crítica de Kolmogorov-Smirnov.

El estadístico  $D_n$  es de *libre distribución* bajo la hipótesis nula, es decir, la distribución bajo  $H_0$  de  $D_n$  no depende de la hipótesis nula. En efecto, para calcular la distribución de  $D_n$  bajo  $H_0$ , tomamos cualquier distribución continua, por ejemplo  $X \sim \mathcal{U}(0, 1)$ , y tomamos  $(X_1, \dots, X_n)$  un m.a.s. suya.

$$\begin{aligned}
 D_n &= \sup_{x \in (0,1)} |F_n(x) - F_0(x)| = \sup_{x \in (0,1)} \left| \frac{\#(X_i \leq x)}{n} - F_0(x) \right| \\
 &= \sup_{x \in (0,1)} \left| \frac{\#(F_0(X_i) \leq F_0(x))}{n} - F_0(x) \right| =_{H_0} \sup_{u \in (0,1)} \left| \frac{\#(U_i \leq u)}{n} - u \right|.
 \end{aligned}$$

### 3.3.3. Test Chi-Cuadrado de Pearson

Este test está diseñado para variables discretas, por lo que, en el caso de tratar con una continua, hemos de discretizarla. Esto es, agrupar los datos en un número finito de clases  $C_1, \dots, C_k$  con  $k \in \mathbb{N}$ .

El estadístico propuesto por Karl Pearson [13] mide la diferencia entre las frecuencias observadas  $n_i$  y las esperadas bajo  $H_0$ ,  $E_i = n \cdot p_i^0$ , donde  $p_i^0 = P_F(X \in C_i)$ , para cada  $i \in \{1, \dots, k\}$ , y siendo  $F$  la función de distribución que queremos contrastar, en este caso  $F \equiv U(0, 1)$ . Se define:

$$D = \sum_{i=1}^k \frac{(n_i - E_i)^2}{E_i} \tag{3.2}$$

Al igual que Komogorov se contruye la Región Crítica para valores grandes:

$$RC_\alpha = \{\vec{u} \mid D > c\}$$

El valor de  $c$  viene dado por el cuantil de una  $\chi_{k-1}^2$  [23].

### 3.3.4. Test de Cramer-von Mises

Harald Cramer [3] y, posteriormente Richard Von Mises [4], crearon un test de bondad de ajuste cuyo estadístico mide, nuevamente, la discrepancia entre funciones de distribución empíricas y teóricas.

Conjuntamente, con ayuda de Smirnov [19], se creó la siguiente familia de estadísticos:

$$W_n^2 = n \int_{\mathbb{R}} (F_n(u) - F_0(u))^2 \phi(u) dF_0(u) = E_{F_0}[\phi(u)(F_n(u) - F_0(u))^2]$$

donde  $\phi(u)$  será una función de pesos que pondera las diferencias entre las funciones de distribución empírica y teórica, y  $E_{F_0}$  denota la esperanza calculada bajo  $H_0$ . En general,  $W_n^2$  representa el valor esperado de la diferencia entre  $F_n$  y  $F_0$  ponderada por  $\phi(u)$ .

Si le damos la misma importancia a todas las diferencias entre  $F_n$  y  $F_0$ , (es decir,  $\phi(u) = 1$ ) obtenemos el estadístico de Cramer-von Mises

$$\begin{aligned} W^2 &= n \sum_{i=1}^n \int_{u_{(i)}}^{u_{(i+1)}} \left( \frac{i}{n} - F_0(u) \right)^2 dF_0(u) \\ &= \frac{1}{12n} + \sum_{i=1}^n \left[ \frac{2i-1}{2n} - F_0(u_{(i)}) \right]^2 \end{aligned} \quad (3.3)$$

Es un estadístico cuya distribución no está determinada de una forma analítica, es decir, necesitaremos simular su distribución bajo la hipótesis nula para obtener la Región Crítica.

$$RC_\alpha = \{\vec{u} \mid W^2 > c\}$$

A raíz de esta simulación podemos comprobar que, a pesar de depender del tamaño muestral  $n$ ,  $W^2$  es de libre distribución bajo cualquier hipótesis nula simple. Los cuantiles obtenidos en dicha simulación aparecen en la Tabla 3.3.

$\alpha \backslash n$	0.1	0.05	0.01
25	0.3434592	0.4551386	0.7370520
100	0.3427222	0.4510198	0.7337144

Tabla 3.3: Cuantiles para la región crítica de Cramer-von Mises.

### 3.3.5. Test de Anderson-Darling

El principal problema del test anterior es que, al ser  $\phi(u)$  constante, no daba suficiente peso a las colas. Lo que Anderson y Darling [1] propusieron fue usar  $\phi(u) = \frac{1}{F_0(u)(1-F_0(u))}$  como función de pesos y aplicarlo a la familia  $W_n^2$  vista anteriormente.

Así, se define entonces el estadístico de contraste:

$$A = -n - \frac{1}{n} \sum_{i=1}^n (2i-1) [\log(F_0(u_{(i)})) + \log(1 - F_0(u_{(n-i+1)}))] \quad (3.4)$$

Un estadístico cuya distribución también ha de ser aproximada por Monte Carlo [22] y cuyos valores grandes hacen construir la Región Crítica del test.

$$RC_\alpha = \{\vec{u} \mid A > c\}$$

Los cuantiles vienen recogidos en la Tabla 3.4 para  $n \in \{25, 100\}$  y  $\alpha \in \{0.1, 0.05, 0.01\}$ .

$\alpha \backslash n$	0.1	0.05	0.01
25	1.920127	2.546230	3.996155
100	1.921962	133.58567	3.881333

Tabla 3.4: Cuantiles para la región crítica de Anderson Darling.

### 3.3.6. Test de Kuiper

El matemático holandés Nicolaas Hendrik Kuiper fue capaz de demostrar el Teorema de Kuiper, en el terreno de la topología, pero también tuvo un hueco en la estadística inferencial en bondad de ajuste.

La prueba de Kuiper está inspirada en el test de Kolmogorov-Smirnov. Kuiper [10] define, análogamente, los estadísticos

$$D_n^+ = \max_{1 \leq i \leq n} \left\{ \frac{i}{n} - U_{(i)} \right\} \quad D_n^- = \max_{1 \leq i \leq n} \left\{ U_{(i)} - \frac{i-1}{n} \right\}$$

pero esta vez se considera la suma de ambos, es decir

$$V = D_n^+ + D_n^- = \max_{i=1, \dots, n} \left\{ \frac{i}{n} - U_{(i)} \right\} + \max_{i=1, \dots, n} \left\{ U_{(i)} - \frac{i-1}{n} \right\}$$

Esta forma de expresar el estadístico hace que la prueba sea tan sensible en la mediana como en las colas, tal y como ocurre con el test de Anderson-Darling. A diferencia de este,  $V$  es *invariante bajo transformaciones cíclicas*.<sup>1</sup>

---

<sup>1</sup>La investigación de Kuiper, de donde nace este test, es un estudio sobre la dirección migratoria de un número observado de pájaros. Para ello, Kuiper partía de una muestra de valores angulares  $(\phi_1, \dots, \phi_n)$ , con  $\phi_i \in [0, 2\pi]$ , la cual estudiamos si se ajusta a una distribución circular determinada. Estas muestras representan direcciones en el plano, tiempo o fenómeno cíclico. Es por esto, que este test debe ser invariante a transformaciones cíclicas de la variable.

Valores grandes del estadístico indican una falta de uniformidad, por lo cual su Región Crítica viene expresada como

$$RC_\alpha = \{\vec{u} \mid V > c\}$$

y los límites de la región crítica vienen dados, a raíz de simular la distribución de  $V$ , en la Tabla 3.5.

$\alpha \backslash n$	0.1	0.05	0.01
25	0.3107406	0.3339817	0.3829180
100	0.1586045	0.171793	0.1973797

Tabla 3.5: Cuantiles para la región crítica de Kuiper.

### 3.4. Tests específicos para la uniformidad

#### 3.4.1. Tests de Zhang

El matemático Jin Zhang [24] también creó una familia de funciones que sirven como estadísticos de bondad de ajuste para contrastar la uniformidad en una muestra. Se definen como

$$Z = \int_{-\infty}^{\infty} Z_t dw(t) \quad (3.5)$$

$$Z_{max} = \sup_{t \in \mathbb{R}} Z_t \cdot w(t) \quad (3.6)$$

donde la distribución de  $Z$  depende claramente de otro estadístico  $Z_t$  y de la función de pesos  $w(t)$ .

Dos candidatos para  $Z_t$  son (tras simplificarlos), el estadístico  $\chi^2$  de Pearson

$$X_t^2 = \frac{n(F_n(t) - F_0(t))^2}{F_0(t)(1 - F_0(t))}$$

y el estadístico del Test de la Razón de Verosimilitudes (*TRV*)

$$G_t^2 = 2n \left[ F_n(t) \log \left( \frac{F_n(t)}{F_0(t)} \right) + (1 - F_n(t)) \log \left( \frac{1 - F_n(t)}{1 - F_0(t)} \right) \right]$$

A la hora de definir los estadísticos, Zhang tuvo en cuenta las discontinuidades de la función de distribución empírica. Para ello modificó  $F_n(t) = \frac{i}{n}$  en sus puntos de discontinuidad  $U_{(i)}$  ( $i=1,2,\dots,n$ ). Definió entonces  $F_n(U_{(i)}) = \frac{(i-1/2)}{n}$ , una corrección de  $F_n$ , que usó para construir los estadísticos siguientes.

**Si escogemos**  $Z_t = X_t^2$ , las diferentes definiciones de  $w(t)$  nos proporcionan tests usuales como por ejemplo:

- $w(t) = \frac{1}{n} F_0(t)(1 - F_0(t))$

$$\begin{aligned} Z &= \sup_{t \in \mathbb{R}} \left\{ \frac{n(F_n(t) - F_0(t))^2}{F_0(t)(1 - F_0(t))} \cdot \frac{1}{n} F_0(t)(1 - F_0(t)) \right\} \\ &= \left( \sup_{t \in \mathbb{R}} |F_n(t) - F_0(t)| \right)^2 = K_S^2 \end{aligned} \quad (3.7)$$

obteniendo así el cuadrado del estadístico del test de Kolmogorov-Smirnov dado en la Ecuación (3.1).

- $w(t) = F_0(t)$

$$\begin{aligned} Z &= n \int_{-\infty}^{\infty} \frac{(F_n(t) - F_0(t))^2}{F_0(t)(1 - F_0(t))} dF_0(t) \\ &= -\frac{2}{n} \sum_{i=1}^n \left[ \left( i - \frac{1}{2} \right) \log(F_0(u_{(i)})) + \left( n - i + \frac{1}{2} \right) \log(1 - F_0(u_{(i)})) \right] - n = A \end{aligned} \quad (3.8)$$

obteniendo así el estadístico del test de Anderson-Darling definido en la Ecuación (3.4).

- $dw(t) = F_0(t)(1 - F_0(t))dF_0(t)$

$$\begin{aligned}
Z &= n \int_{-\infty}^{\infty} \frac{(F_n(t) - F_0(t))^2}{F_0(t)(1 - F_0(t))} F_0(t)(1 - F_0(t)) dF_0(t) \\
&= n \int_{-\infty}^{\infty} (F_n(t) - F_0(t))^2 dF_0(t) \\
&= \sum_{i=1}^n \left[ F_0(u_{(i)}) - \frac{i - \frac{1}{2}}{n} \right]^2 + \frac{1}{12n} = W^2
\end{aligned} \tag{3.9}$$

que coincide con el estadístico del test de Cramer-von Mises visto en la Ecuación (3.3).

**Posteriormente Zhang tomó**  $Z_t = G_t^2$ , y haciendo variar la función de pesos se obtienen nuevos tests para la uniformidad:

- $w(t) = 1$

$$\begin{aligned}
Z &= \sup_{t \in (-\infty, \infty)} (G_t^2) = \max_{1 \leq i \leq n} (G_{U_{(i)}}^2) \\
&= \max_{1 \leq i \leq n} \left( \left( i - \frac{1}{2} \right) \log \left\{ \frac{i - \frac{1}{2}}{n F_0(u_{(i)})} \right\} + \left( n - i + \frac{1}{2} \right) \log \left\{ \frac{n - i + \frac{1}{2}}{n(1 - F_0(u_{(i)}))} \right\} \right) =: Z_K
\end{aligned} \tag{3.10}$$

- $dw(t) = \frac{1}{F_n(t)(1 - F_n(t))} dF_n(t)$

$$\begin{aligned}
Z &= 2n \int_{-\infty}^{\infty} \left[ F_n(t) \log \left( \frac{F_n(t)}{F_0(t)} \right) + (1 - F_n(t)) \log \left( \frac{1 - F_n(t)}{1 - F_0(t)} \right) \right] \frac{1}{F_n(t)(1 - F_n(t))} dF_n(t) \\
&= 2 \sum_{i=1}^n \left( \frac{n}{n - i + \frac{1}{2}} \log \left\{ \frac{i - \frac{1}{2}}{n F_0(u_{(i)})} \right\} + \frac{n}{i - \frac{1}{2}} \log \left\{ \frac{n - i + \frac{1}{2}}{n(1 - F_0(u_{(i)}))} \right\} \right) \\
&= - \sum_{i=1}^n \left[ \frac{\log(F_0(u_{(i)}))}{n - i + \frac{1}{2}} + \frac{\log(1 - F_0(u_{(i)}))}{i - \frac{1}{2}} \right] =: Z_A
\end{aligned} \tag{3.11}$$

- $dw(t) = \frac{1}{F_0(t)(1 - F_0(t))} dF_0(t)$

$$\begin{aligned}
Z &= 2n \int_{-\infty}^{\infty} \left[ F_n(t) \log \left( \frac{F_n(t)}{F_0(t)} \right) + (1 - F_n(t)) \log \left( \frac{1 - F_n(t)}{1 - F_0(t)} \right) \right] \frac{1}{F_0(t)(1 - F_0(t))} dF_0(t) \\
&= \sum_{i=1}^n [\log(F_0(u_{(i)})^{-1} - 1) - b_{i-1} - b_i]^2 + K_n \\
&= \sum_{i=1}^n \left[ \log \left\{ \frac{F_0(u_{(i)})^{-1} - 1}{(n - \frac{1}{2}) / (i - \frac{3}{4}) - 1} \right\} \right]^2 =: Z_C
\end{aligned} \tag{3.12}$$

La distribución de estos nuevos estadísticos vienen simulada por Monte Carlo [22], y, al igual que para  $Z_t = G_t^2$ , las Regiones Críticas de  $Z_K, Z_A$  y  $Z_C$  vienen dadas por

$$RC_\alpha = \{\vec{u} \mid Z_K > c_K\}; \quad RC_\alpha = \{\vec{u} \mid Z_A > c_A\}; \quad RC_\alpha = \{\vec{u} \mid Z_C > c_C\}.$$

donde los valores de  $c_K, c_A$  y  $c_C$  vienen determinados, para ciertos  $n$  y ciertos  $\alpha$  en las Tablas 3.6, 3.7 y 3.8, respectivamente.

$n \backslash \alpha$	0.1	0.05	0.01
25	3.210303	3.867830	5.422330
100	3.837106	4.549315	6.247913

Tabla 3.6: Cuantiles para la región crítica de Zhang  $Z_k$ .

$n \backslash \alpha$	0.1	0.05	0.01
25	3.626382	3.712772	3.952378
100	3.395322	3.419980	3.481847

Tabla 3.7: Cuantiles para la región crítica de Zhang  $Z_A$ .

---

<sup>2</sup>Donde  $b_i = i \log\left(\frac{i}{n}\right) + (n-i) \log\left(1 - \frac{i}{n}\right)$  y  $K_n$  es una constante. Por lo que  $b_{i-1} - b_i \approx \log\left\{\left(n - \frac{1}{2}\right) / \left(i - \frac{3}{4}\right) - 1\right\}$ .

$\alpha$	0.1	0.05	0.01
$n$			
25	19.35709	23.69362	34.73662
100	23.26852	28.25562	40.55625

Tabla 3.8: Cuantiles para la región crítica de Zhang  $Z_C$ .

### 3.4.2. Test Greenwood

Major Greenwood fue un epidemiólogo, fisiólogo y médico estadístico británico, que, entre otras muchas cosas, fue director del Departamento de Estadística del Consejo de Investigación Médica (MRC) en Cambridge. En su último año de cátedra en la LSHTM (Escuela de Higiene y Medicina Tropical), 1946 [8], publicó su artículo más importante sobre el estudio estadístico de enfermedades infecciosas, donde presentó el estadístico  $G$ , que hoy en día recibe su nombre.

Lo definió como una medida de agrupación de eventos en el espacio o el tiempo. Este estadístico tiene como elemento principal la longitud de un intervalo entre dos valores consecutivos de la muestra.

$$G = \sum_{i=1}^{N+1} \left( U_{(i)} - U_{(i-1)} \right)^2$$

suponiendo  $U_0 = 0$  y  $U_{N+1} = 1$ . Parece un estadístico muy simple, pero en 2007 fue utilizado para demostrar que existe una importancia en el orden en el que los genes se colocan en un cromosoma [17].

En 1975, J. S. Rao y Sethuraman [16] demostraron que, bajo  $H_0$ ,  $G$  tiene una distribución asintóticamente normal, sin embargo, su convergencia es tan lenta que típicamente es necesario obtener sus percentiles mediante la simulación de Monte Carlo [22].

Apenas existen artículos que recojan estos percentiles, ya que el estadístico de Greenwood fue corregido, con la intención de aumentar su potencia. No obstante, se define la región crítica, para valores grandes del estadístico.

$$RC_\alpha = \{\vec{u} \mid G > c\}$$

Las constantes límites de la región, después de simular la distribución de  $G$ , constan en la Tabla 3.9.

$\alpha \backslash n$	0.1	0.05	0.01
25	0.09074716	0.09749356	0.11843118
100	0.02204608	0.02300411	0.02522304

Tabla 3.9: Cuantiles para la región crítica de Greenwood.

### 3.4.3. Test Q

Quesenberry y Miller [15] propusieron un nuevo test, basado en el de Greenwood, teniendo en cuenta cosas que este no tenía. Se define como

$$Q = \sum_{i=1}^{N+1} \left( U_{(i)} - U_{(i-1)} \right)^2 + \sum_{i=1}^N \left( U_{(i)} - U_{(i-1)} \right) \left( U_{(i+1)} - U_{(i)} \right)$$

El estadístico  $G$  solo tiene en cuenta las longitudes de los subintervalos que se forman en  $(0,1)$ , sin embargo la información adicional que el estadístico  $Q$  aporta en su definición es la distribución de estas longitudes.

El orden de estas amplitudes es importante en estos contrastes ya que pueden mostrar indicios de no uniformidad. Por ejemplo, no es lo mismo tener un intervalo cuyas longitudes se distribuyan así:

$$\left( \frac{1}{15}, \frac{1}{15}, \frac{1}{15}, \frac{3}{15}, \frac{4}{15}, \frac{5}{15} \right)$$

de donde es difícil deducir la uniformidad, que así

$$\left( \frac{1}{15}, \frac{3}{15}, \frac{1}{15}, \frac{4}{15}, \frac{1}{15}, \frac{5}{15} \right)$$

Una vez más, no se conoce una distribución asintótica para  $Q$ , pero gracias a la simulación de su distribución podemos conocer los cuantiles que limitan la Región Crítica (en la Tabla 3.10). Al igual que en Greenwood, esta viene dada por

$$RC_\alpha = \{\vec{u} \mid Q > c\}$$

$\alpha \backslash n$	0.1	0.05	0.01
25	0.12807910	0.1356961	0.15608156
100	0.03205735	0.0331604	0.03552491

Tabla 3.10: Cuantiles para la región crítica de Quesenberry y Miller.

#### 3.4.4. Test de Durbin-Kolmogorov

James Durbin fue un economista y estadístico británico muy influyente gracias a sus contribuciones sobre la correlación serial [6] y los tests de bondad de ajuste (entre otros muchos campos).

Asemejándose al test  $Q$ , un punto de partida para que Durbin creara este test es el estudio de las longitudes de los intervalos formados por los puntos adyacentes, a priori, aleatoriamente dispersos sobre el  $(0,1)$ .

Lo que Durbin [5] intenta hacer es un "reordenamiento" de los intervalos. Suponemos que tiene una muestra aleatoria simple  $x_1, \dots, x_n$ , sobre la que tratamos de ver si su función de distribución  $F(x)$  se corresponde con la

de una uniforme en  $(0,1)$ . Sabemos que la transformación  $u_i = F(x_i)$  se distribuye uniformemente en ese intervalo. Denotamos  $u_{(1)}, \dots, u_{(n)}$  la muestra ordenada, y definimos los elementos  $c_i$ ,

$$c_1 = u_{(1)}; \quad c_i = u_{(i)} - u_{(i-1)}; \quad c_{n+1} = 1 - u_{(n)}$$

cuya distribución viene dada simplemente por la probabilidad en cada subintervalo. De la misma forma hallamos la distribución de  $c_{(1)}, \dots, c_{(n+1)}$  (el conjunto de elementos ordenados) y podemos definir

$$g_i = (n + 2 - 1)(c_{(i)} - c_{(i-1)}) \quad (\text{con } c_{(0)} = 0; \quad i = 1, \dots, n + 1)$$

Encontramos entonces, unos elementos definidos por

$$w_r = \sum_{i=1}^r g_i \quad r = 1, \dots, n$$

para los que Durbin demostró que siguen la misma distribución que la muestra uniforme ordenada  $u_{(1)}, \dots, u_{(n)}$ .

El fin de transformar  $u_{(1)}, \dots, u_{(n)}$  en otra muestra igualmente distribuida era aumentar la potencia, así, Durbin define el test llamado Test Kolmogorov Modificado, que posteriormente pasaría a llamarse Test Durbin-Kolmogorov, como

$$DK = \max_{r=1, \dots, n} \left( \frac{r}{n} - w_r \right).$$

Un estadístico cuyo valor se espera que aumente a medida que nos alejamos de la hipótesis nula, por lo que se construye la Región Crítica

$$RC_\alpha = \{\vec{u} \mid DK > c\}$$

donde  $c$  viene tabulado para el contraste unilaterial de Kolmogorov-Smirnov, o análogamente, simulando la distribución de  $DK$ , por Monte Carlo [22], en la Tabla 3.11.

$\alpha \backslash n$	0.1	0.05	0.01
25	0.2067762	0.2368914	0.2931949
100	0.1060238	0.1207198	0.1516986

Tabla 3.11: Cuantiles para la región crítica de Durbin-Kolmogorov.

### 3.4.5. Test Kendall-Sherman

En 1946, Maurice George Kendall [8] propuso un test que, al igual que Greenwood, también basó en longitudes de intervalos de puntos adyacentes. El estadístico se define como

$$w_n = \frac{1}{2} \sum_{i=1}^{n+1} \left| c_i - \frac{1}{n+1} \right|$$

donde  $c_i = U_{(i)} - U_{(i-1)}$ .

Esta prueba lleva también el nombre de Sherman [18], ya que él fue quien, en 1950, se encargó de dar una expresión analítica tanto a la función de distribución como a su esperanza y varianza, obteniendo así

$$F_n(u) = \begin{cases} 0 & \text{si } u < 0 \\ b_n u^n + b_{n-1} u^{n-1} + \dots + b_1 u + b_0 & \text{si } 0 \leq u \leq \frac{n}{n+1} \\ 1 & \text{si } u > \frac{n}{n+1} \end{cases}$$

donde

$$b_n = \sum_{j=1}^r (-1)^{j+i+1} \binom{n+1}{j+1} \binom{j+i}{j} \binom{n}{i} \binom{n-j}{n+1}^{n-i}$$

Sin embargo en 1954, D. J. Bartholomew [2], demostró que para un tamaño de muestra superior a 21,  $w_n$  seguía aproximadamente una distribución nor-

mal de parámetros

$$\mu = \frac{1}{e} \left( 1 - \frac{1}{2n} + \frac{7}{24n^2} \right)$$

$$\sigma^2 = \frac{0.05908}{n} - \frac{0.07145}{n^2}$$

No obstante, en el mismo artículo, el propio Bartholomew propone una transformación al estadístico  $w_n$  para que los resultados, al usar su distribución asintótica, sean mucho más precisos. Esto es, si defino

$$v = \frac{w_n - 0.3679(1 - \frac{1}{2n})}{0.2431n^{-1}(1 - 0.605n^{-1})}$$

$$v' = v - \frac{0.0995}{\sqrt{n}}(u^2 - 1)$$

podemos usar que  $v' \equiv \mathcal{N}(0, 1)$ .

Al igual que  $w_n$ , la transformación  $v'$  aumenta su valor fuera de la hipótesis nula, así

$$RC_\alpha = \{\vec{u} \mid v' > c\}$$

donde  $c$  será el cuantil de una  $\mathcal{N}(0, 1)$ .

### 3.4.6. Test del Producto de Probabilidades de Pearson

Karl Pearson [14], en 1933, sugirió uno de los estadísticos más simples propuestos hasta el momento para contrastar uniformidad. Se construye, como su propio nombre indica, el estadístico

$$Q_1 = \prod_{i=1}^n u_i$$

que se define como la probabilidad conjunta de que una muestra aleatoria simple sea menor o igual que los correspondientes valores observados.

A su vez, en el mismo artículo, consideró un estadístico análogo

$$Q'_1 = \prod_{i=1}^n (1 - u_i)$$

que se define como la probabilidad de que la muestra sea mayor o igual que el valor observado.

Se considera el test que rechaza la hipótesis nula para valores pequeños de  $Q_1$  y  $Q'_1$  a nivel  $\alpha$ , cuya distribución no es trivialmente cognoscible. Sin embargo, haciendo una simple transformación

$$v_i = -2 \log u_i$$

es fácil comprobar que  $-2 \log(Q_1)$  y  $-2 \log(Q'_1)$  se distribuyen como una  $\chi^2_{2n}$ . Es por tanto que se tiene:

$$RC_\alpha = \{\vec{u} \mid Q_1 < c_\alpha\} = \{\vec{u} \mid -2 \log Q_1 > k_\alpha\}$$

con  $k_\alpha = F_{\chi^2_{2n}}^{-1}(1 - \alpha)$  y análogamente para  $Q'_1$ .

Karl Pearson estaba al tanto de la dificultad de elegir  $Q_1$  y  $Q'_1$ . En la experimentación posterior del Capítulo 4 se utilizará, por comodidad,  $Q_1$ . Además, bajo  $H_0$ , la Proposición 2.1 nos asegura que ambos estadísticos tienen la misma distribución, fuera de  $H_0$  no se comportarán exactamente igual.

### 3.4.7. Test de Sukhatme

Debido a la indecisión de cuál de los dos tests usar, seis años más tarde, Egon S. Pearson [12], hijo de Karl Pearson, propuso un tercer criterio que tuviera en cuenta si la distribución de la hipótesis alternativa pudiera concentrarse más en algún extremo concreto del intervalo (0,1). Para ello definió el estadístico

$$Q_2 = \prod_{i=1}^n u'_i$$

donde

$$u'_i = \begin{cases} 2u_i & \text{si } 0 < u_i \leq \frac{1}{2} \\ 2(1 - u_i) & \text{si } \frac{1}{2} \leq u_i < 1 \end{cases}$$

Se demostró también que,  $-2 \log Q_2$  sigue una distribución  $\chi_{2n}^2$ .

El criterio para rechazar  $H_0$  viene dado por intuir que, cuanto menor sea el valor de  $Q_2$ , significará que más valores se acumulan en las colas del intervalo  $(0,1)$ , luego definimos la Región Crítica, a un nivel  $\alpha$  como

$$RC_\alpha = \{\vec{u} \mid Q_2 < c_\alpha\} = \{\vec{u} \mid -2 \log Q_2 > k_\alpha\}$$

con  $k_\alpha = F_{\chi_{2n}^2}^{-1}(1 - \alpha)$ .

Pearson atribuyó este test al estadístico Pandurang V. Sukhatme, del que recibe su nombre.

Es fácil ver, gracias al lema Neymann-Pearson [11], que los tests  $Q_1$  y  $Q_2$  son los más potentes cuando en la hipótesis alternativa se contrastan densidades del tipo que más adelante asociaremos con densidades de la distribución  $\beta(p, p)$  con  $p < 1$ . La función  $h$  aparece dibujada en la Figura 3.4.

$$h(y, \lambda) = \begin{cases} (1 - \lambda)y^{-\lambda} & \text{si } 0 < y \leq \frac{1}{2} \\ (1 - \lambda)(1 - y)^{-\lambda} & \text{si } \frac{1}{2} < y < 1 \end{cases} \quad \text{con } 0 < \lambda < 1$$

Esto es porque la Región Crítica se define a partir de valores pequeños del estadístico, o análogamente, a partir de valores grandes de su transformación logarítmica. Como en este trabajo buscamos tests potentes para cualquier clase de alternativa, se definen nuevas Regiones Críticas

$$RC'_\alpha = \{\vec{u} \mid | -2 \log Q_1 | > k'_\alpha\} \quad RC''_\alpha = \{\vec{u} \mid | -2 \log Q_2 | > k''_\alpha\}$$

con  $k'_\alpha = F_{\chi_{2n}^2}^{-1}(1 - \frac{\alpha}{2})$  y  $k''_\alpha = F_{\chi_{2n}^2}^{-1}(1 - \frac{\alpha}{2})$ .



## Capítulo 4

# Experimentación

**Definición 4.1 [Potencia de un contraste. ]** Se define como la probabilidad de obtener una muestra que caiga dentro de la Región Crítica, es decir, es la probabilidad de rechazar  $H_0$ . Se espera que la potencia sea cercana a 0 cuando se trabaja bajo la hipótesis nula, y que sea cercana a 1 en caso contrario.

**Definición 4.2 [Nivel de significación. ]** Puede verse como una cota superior para la probabilidad de rechazar la hipótesis nula cuando esta es cierta. Se denota por  $\alpha$  y en este trabajo queda fijado en  $\alpha = 0.05$ .

En este capítulo se hará un análisis comparativo de los estadísticos presentados en el capítulo anterior basado en la potencia de los mismos. Proporcionaremos a los estadísticos muestras que procedan de distintas distribuciones con soporte en el intervalo  $(0,1)$ , debido al problema con las definiciones de algunos estadísticos, con la intención de que detecten o rechacen la uniformidad de dichas distribuciones. Como familias de alternativas se ha tomado la propia uniforme, distintas distribuciones betas, distribuciones triangulares y diferentes distribuciones trapezoidales.

Se ha fijado una semilla en **R Studio**, `set.seed(1)` y se han utilizado diez mil réplicas para simular las distribuciones. Además, para la distribución bajo la hipótesis nula,  $\mathcal{U}(0, 1)$ , se ha hecho variar el tamaño muestral  $n \in \{10, 25, 50, 100, 250, 500, 1000\}$  y, para el resto de familias de alternativas, se hacen variar los diferentes parámetros de cada distribución, restringiéndonos en algunos casos a  $n \in \{25, 100\}$ .

#### 4.1. La distribución bajo $H_0$ , la uniforme estándar

Primero de todo, para hacer una breve comprobación del programa implementado, se usarán muestras provenientes de una  $\mathcal{U}(0, 1)$ , cuya densidad aparece en la Figura 4.1.

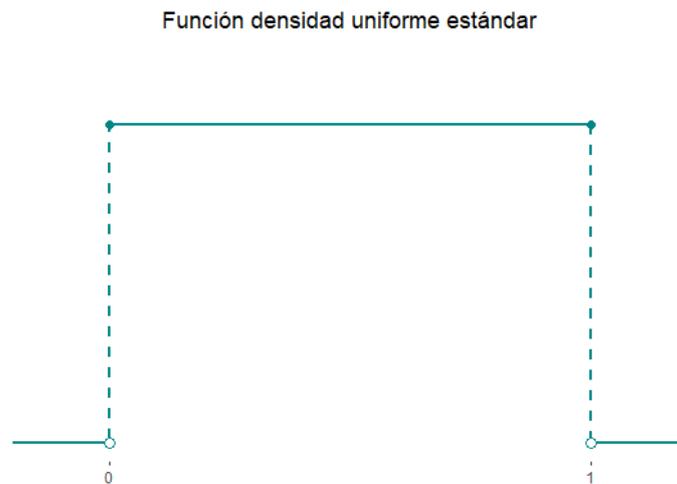


Figura 4.1: Función de densidad uniforme.

##### 4.1.1. La potencia en función del tamaño muestral

Para ver cuánto de buenos son nuestros tests, antes de ver si detectan rápidamente la uniformidad, debemos comprobar, dándole muestras uniformes de distintos tamaños, que la potencia se mantiene.

Efectivamente, en la Tabla 4.1 se aprecia que todas las potencias se ajustan bastante bien al nivel de significación fijado  $\alpha = 0.05$ , sin embargo, en la última columna de las dos últimas filas aparecen potencias de 1, debido a un problema con la tolerancia de R `Studio`. Esto es, al haber un número muy grande de elementos en la muestra (procedente de una  $\mathcal{U}(0, 1)$ ), se generan algunos números tan pequeños que son detectados como 0, por lo que  $Q_1 = Q_2 = 0$ . Esto provoca que, al calcular los cuantiles a partir de su distribución asintótica (para la cual es necesaria la transformación logarítmica  $-2\log(Q_1)$  y  $-2\log(Q_2)$ ), salgan ambos Inf. De esta manera rechazaríamos la hipótesis nula un 100% de las veces, obteniendo una potencia igual a 1.

$\mathcal{U}(0, 1)$	n=10	n=25	n=50	n=100	n=250	n=500	n=1000
TRV	0.0455	0.0508	0.0463	0.0495	0.0520	0.0532	0.0554
Kolmogorov-Smirnov	0.0503	0.0485	0.0486	0.0480	0.0458	0.0484	0.0501
Pearson Chi Cuadrado	0.0580	0.0490	0.0479	0.0554	0.0512	0.0506	0.0499
Cramer-von Mises	0.0494	0.0490	0.0477	0.0521	0.0518	0.0519	0.0496
Anderson Darling	0.0500	0.0471	0.0492	0.0510	0.0531	0.0513	0.0504
Kuiper	0.0505	0.0553	0.0415	0.0508	0.0458	0.0484	0.0501
Zhang ZK	0.0481	0.0542	0.0472	0.0439	0.0504	0.0525	0.0466
Zhang ZC	0.0485	0.0512	0.0536	0.0490	0.0502	0.0498	0.0506
Zhang ZA	0.0488	0.0542	0.0496	0.0463	0.0507	0.0520	0.0500
Greenwood	0.0492	0.0452	0.0503	0.0511	0.0528	0.0531	0.0483
Q	0.0485	0.0486	0.0472	0.0503	0.0489	0.0468	0.0471
Drubin Kolmogorov	0.0541	0.0480	0.0473	0.0476	0.0549	0.0461	0.0500
Kendall-Sherman	0.0491	0.0504	0.0517	0.0504	0.0477	0.0499	0.0508
Pearson Q1	0.0514	0.0474	0.0471	0.0488	0.0487	0.0530	1.0000
Sukhatme	0.0490	0.0515	0.0497	0.0480	0.0492	0.0492	1.0000

Tabla 4.1: Familia alternativa:  $\mathcal{U}(0, 1)$  para diferentes tamaños muestrales

## 4.2. La familia uniforme

Si tomamos como familia alternativa una  $\mathcal{U}(a, b)$  con  $a \leq 0$  y  $b \geq 1$  cualesquiera pero fijos, no obtendríamos resultados útiles en muchos tests, y es debido simplemente a la definición de su estadístico, sobre todo en aquellos estadísticos en los que intervenga una expresión logarítmica, como pueden ser Anderson-Darling o los propuestos por Jin Zhang. No obstante, se realizará una pequeña simulación para valores de  $a$  y  $b$  tales que  $0 < a < b < 1$ , en concreto tomaremos  $a \in \{0, 0.05, 0.1, 0.2, 0.3\}$  y  $b = 1 - a$  para no perder la simetría. Obtenemos así las densidades que aparecen en la Figura 4.2 y, tomando muestras de tamaño 25, las potencias para cada estadístico las vemos en la Tabla 4.2.

Función densidad uniformes

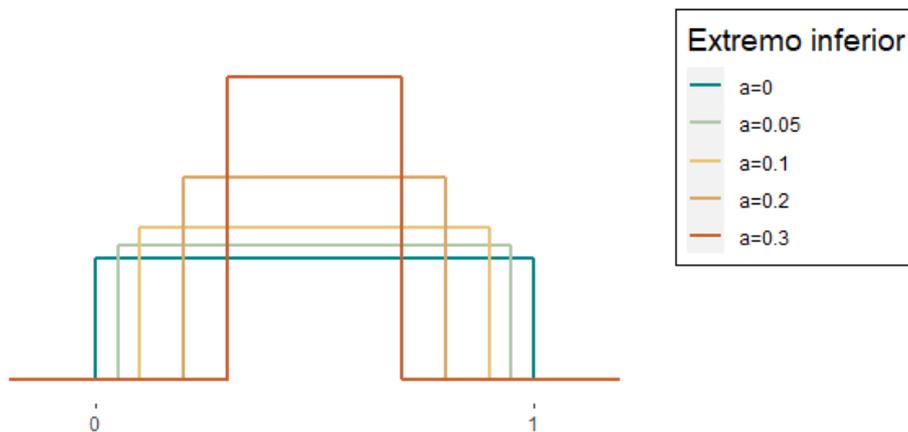


Figura 4.2: Función de densidad  $\mathcal{U}(a, b)$  con  $a \in \{0, 0.05, 0.1, 0.2, 0.3\}$  y  $b = 1 - a$ .

Al igual que para el resto de alternativas, se hará un estudio similar al de la Tabla 4.2, pero con muestras de tamaño 100, con el objetivo de observar si el tamaño muestral influye y en qué estadísticos es más notable

$\mathcal{U}(a, 1 - a)$	$a = 0$	$a = 0.05$	$a = 0.1$	$a = 0.2$	$a = 0.3$
TRV	0.0543	0.3637	1.0000	1.0000	1.0000
Kolmogorov-Smirnov	0.0485	0.0431	0.0510	0.2201	1.0000
Pearson Chi Cuadrado	0.0490	0.0557	0.1106	0.5198	0.9995
Cramer-von Mises	0.0490	0.0395	0.0494	0.3579	1.0000
Anderson Darling	0.0471	0.0315	0.0486	0.5414	1.0000
Kuiper	0.0553	0.1087	0.3131	1.0000	1.0000
Zhang ZK	0.0542	0.0470	0.0962	0.9893	1.0000
Zhang ZC	0.0512	0.0745	0.3798	1.0000	1.0000
Zhang ZA	0.0542	0.1249	0.6116	1.0000	1.0000
Greenwood	0.0452	0.0289	0.0695	1.0000	1.0000
Q	0.0486	0.0202	0.0305	0.7937	1.0000
Drubin Kolmogorov	0.0480	0.0367	0.1118	0.9645	1.0000
Kendall-Sherman	0.0504	0.0431	0.1302	0.7680	1.0000
Pearson Q1	0.0474	0.0351	0.0460	0.0632	0.0576
Sukhatme	0.0515	0.2061	0.7157	1.0000	1.0000

Tabla 4.2: Familia alternativa:  $\mathcal{U}(a, 1 - a)$  con  $a \in \{0, 0.05, 0.1, 0.2, 0.3\}$  y tamaño muestral 25.

esta influencia. Dicho estudio se corresponde con la Tabla 4.3

Podemos observar un gran cambio en las potencias dependiendo del tamaño muestral. En la Tabla 4.2, donde las muestras proporcionadas tenían un tamaño razonablemente pequeño ( $n = 25$ ), solo unos pocos tests, como el TRV, Kuiper, ZA de Zhang y Sukhatme tienen potencias altas contra la alternativa de  $\mathcal{U}(0.05, 0.95)$ , la más parecida a nuestra distribución en  $H_0$ . Sin embargo en la Tabla 4.3, para muestras de tamaño 100, solo el test de Kolmogorov-Smirnov junto con el de Cramer-von Mises mantienen una potencia baja para la distribución  $\mathcal{U}(0.05, 0.95)$ . De todos modos, los tests mencionados anteriormente siguen siendo los más potentes contra estas alternativas.

$\mathcal{U}(a, 1 - a)$	$a = 0$	$a = 0.05$	$a = 0.1$	$a = 0.2$	$a = 0.3$
TRV	0.0449	1.0000	1.0000	1.0000	1.0000
Kolmogorov-Smirnov	0.0480	0.0657	0.2673	1.0000	1.0000
Pearson Chi Cuadrado	0.0554	0.1064	0.3444	0.9887	1.0000
Cramer-von Mises	0.0521	0.0777	0.4516	1.0000	1.0000
Anderson Darling	0.0510	0.1380	0.9256	1.0000	1.0000
Kuiper	0.0508	0.3399	1.0000	1.0000	1.0000
Zhang ZK	0.0439	0.5052	1.0000	1.0000	1.0000
Zhang ZC	0.0490	0.9997	1.0000	1.0000	1.0000
Zhang ZA	0.0463	1.0000	1.0000	1.0000	1.0000
Greenwood	0.0511	0.3826	1.0000	1.0000	1.0000
Q	0.0503	0.1855	1.0000	1.0000	1.0000
Drubin Kolmogorov	0.0476	0.2308	1.0000	1.0000	1.0000
Kendall-Sherman	0.0504	0.2046	0.7169	1.0000	1.0000
Pearson Q1	0.0488	0.1455	0.3877	0.9172	0.9999
Sukhatme	0.0480	0.8797	1.0000	1.0000	1.0000

Tabla 4.3: Familia alternativa:  $\mathcal{U}(a, 1 - a)$  con  $a \in \{0, 0.05, 0.1, 0.2, 0.3\}$  y tamaño muestral 100.

Cabe destacar que, en la Tabla 4.2, el estadístico  $Q_1$  propuesto por Pearson no parece detectar diferencias entre una  $\mathcal{U}(0, 1)$  o cualquier uniforme con intervalo contenido en  $[0, 1]$ , esto se debe al bajo tamaño muestral, puesto que en la Tabla 4.3 las potencias aumentan para estas alternativas.

### 4.3. La familia $\beta(p, q)$

Una variable aleatoria  $X$  que sigue una distribución  $\beta(p, q)$ , tiene una función de densidad

$$f(x) = \frac{1}{B(p, q)} x^{p-1} (1-x)^{q-1} \quad \mathbb{I}_{(0,1)} \quad (4.1)$$

con  $p, q \in \mathbb{R}^+$  y  $B(p, q) = \int_0^1 x^{p-1}(1-x)^{q-1}dx = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$ .

Teniendo en cuenta que  $\mathcal{U}(0, 1) \equiv \beta(1, 1)$ , en la Tabla 4.4 se hace un examen análogo al de la Tabla 4.1, por puro control de mi programa.

#### 4.3.1. La potencia en función del tamaño muestral

Aunque las distribuciones son idénticas, las potencias deberían salir cercanas a 0.05 pero no iguales que en la Tabla 4.1, simplemente por la manera que tiene R Studio de generar números aleatorios.

$\beta(1, 1)$	n=10	n=25	n=50	n=100	n=250	n=500	n=1000
TRV	0.0492	0.0525	0.0436	0.0501	0.0577	0.0537	0.0516
Kolmogorov-Smirnov	0.0469	0.0463	0.0535	0.0474	0.0476	0.0521	0.0487
Pearson Chi Cuadrado	0.0593	0.0481	0.0527	0.0551	0.0521	0.0514	0.0486
Cramer-von Mises	0.0473	0.0501	0.0504	0.0508	0.0497	0.0522	0.0505
Anderson Darling	0.0480	0.0511	0.0492	0.0510	0.0503	0.0528	0.0500
Kuiper	0.0562	0.0536	0.0489	0.0509	0.0503	0.0514	0.0526
Zhang ZK	0.0495	0.0501	0.0477	0.0433	0.0494	0.0562	0.0449
Zhang ZC	0.0522	0.0550	0.0522	0.0492	0.0520	0.0525	0.0511
Zhang ZA	0.0535	0.0551	0.0507	0.0480	0.0515	0.0538	0.0480
Greenwood	0.0508	0.0452	0.0534	0.0513	0.0491	0.0498	0.0498
Q	0.0528	0.0509	0.0522	0.0475	0.0487	0.0461	0.0474
Drubin Kolmogorov	0.0553	0.0515	0.0513	0.0474	0.0536	0.0499	0.0491
Kendall-Sherman	0.0514	0.0495	0.0515	0.0501	0.0505	0.0477	0.0458
Pearson Q1	0.0468	0.0484	0.0519	0.0507	0.0519	0.0530	1.0000
Sukhatme	0.0521	0.0498	0.0498	0.0488	0.0464	0.0518	1.0000

Tabla 4.4: Familia alternativa:  $\beta(1, 1)$  con tamaño variable.

Podemos observar, las mismas conclusiones en la Tabla 4.4, con potencias distintas en cada caso, pero muy próximas a  $\alpha = 0.05$ , con la excepción de los últimos dos tests para n=1000, por el problema explicado anteriormente.

### 4.3.2. La potencia en función de los parámetros de forma

Según los valores que tome  $p$  y  $q$  dentro de la familia de betas, podemos obtener alternativas muy distintas, las cuales estudiaremos en las próximas tablas. Es importante analizar si, haciendo variar los parámetros, existen dificultades en los tests para detectar la uniformidad cuando el tamaño muestral no es suficientemente grande. Es por eso por lo que estudiaremos cada caso para  $n = 25$ , y un tamaño genéricamente grande  $n = 100$ .

#### 1. Betas simétricas

Siempre que  $p$  y  $q$  sean iguales, la función de densidad será simétrica. Tal y como se indicó anteriormente, si ambos son la unidad, estamos frente a una uniforme en  $(0,1)$ . Si ambos son menores que 1, la función de densidad es una especie de parábola cóncava hacia arriba, sin embargo si ambos parámetros son mayores que 1 obtendremos una parábola cóncava hacia abajo. Se han usado, como valores más grandes que 1  $p, q \in \{2, 10\}$  y como valores menores que 1, sus inversos  $p, q \in \{\frac{1}{10}, \frac{1}{2}\}$ . Además, la columna central tomará los valores  $p = q = 1$ , por control del estudio.

En la Figura 4.3 pueden verse representadas las densidades de estas alternativas.

Considerando una muestra de tamaño 25, podemos observar en la Tabla 4.5, que nuestros tests tienden a rechazar mejor la hipótesis de uniformidad con parámetros menores que 1 que con parámetros más altos. Por ejemplo, tests tradicionales como Kolmogorov-Smirnov, Cramer-von Mises o Anderson Darling, mantienen una potencia baja para las  $\beta(2, 2)$ . Para valores de  $p$  y  $q$  muy alejados de 1, les basta con muestras de tamaño 25 para mostrar potencias altas.

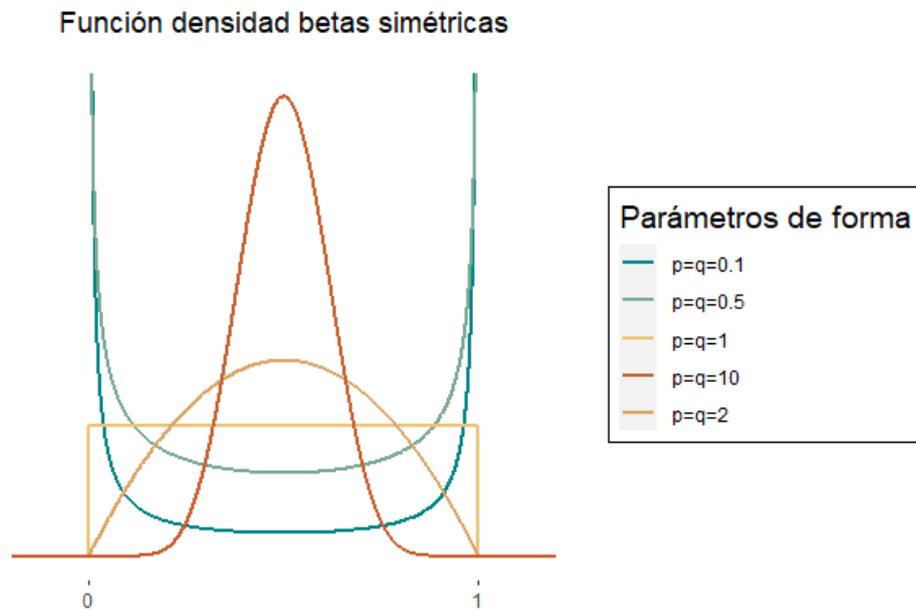


Figura 4.3: Función de densidad  $\beta(p, q)$  con  $p = q \in \{0.1, 0.5, 1, 2, 10\}$ .

Destaca el TRV, que, para  $\beta(0.1, 0.1)$  proporciona una potencia nula y muy baja para  $\beta(0.5, 0.5)$ . Esto se debe a que, para parámetros de la beta menores que 1, la mayoría de valores se acumulan en los extremos del intervalo, es decir, el máximo y el mínimo muestrales son cercanos a 1 y a 0 respectivamente. Por esto, caemos con menos frecuencia en la  $RC_\alpha$ . Además, el test del Producto de Pearson, con estadístico  $Q_1$ , arrastra una potencia bastante próxima a 0.05 incluso para  $\beta(10, 10)$ . Esto puede deberse a que sus cuantiles fueron calculados a partir de la distribución asintótica de  $-2\log(Q_1)$ , por lo que estos problemas deberían desaparecer una vez aumentemos el tamaño de muestra.

Realizamos el mismo estudio en la Tabla 4.6 y, efectivamente, el problema con el estadístico  $Q_1$  de Pearson se ha resuelto al aumentar el tamaño muestra.

$\beta(p, q)$	$p = q = 0.1$	$p = q = 0.5$	$p = q = 1$	$p = q = 2$	$p = q = 10$
TRV	0.0000	0.0013	0.0524	0.6487	1.0000
Kolmogorov-Smirnov	0.9991	0.2391	0.0463	0.0853	0.9917
Pearson Chi Cuadrado	0.8835	0.1801	0.0481	0.2561	0.9998
Cramer-von Mises	0.9997	0.2399	0.0501	0.0714	1.0000
Anderson Darling	1.0000	0.6119	0.0511	0.0708	1.0000
Kuiper	1.0000	0.5080	0.0536	0.4500	1.0000
Zhang ZK	1.0000	0.6735	0.0501	0.1742	1.0000
Zhang ZC	1.0000	0.7900	0.0550	0.3426	1.0000
Zhang ZA	1.0000	0.6933	0.0551	0.4834	1.0000
Greenwood	0.9995	0.3314	0.0452	0.1031	0.9999
Q	1.0000	0.4696	0.0509	0.0729	1.0000
Drubin Kolmogorov	1.0000	0.3998	0.0515	0.1062	0.9997
Kendall-Sherman	1.0000	0.4312	0.0495	0.1035	0.9988
Pearson Q1	0.9961	0.4391	0.0484	0.0481	0.0583
Sukhatme	1.0000	0.8761	0.0498	0.7311	1.0000

Tabla 4.5: Familia alternativa:  $\beta(p, q)$  con  $p = q \in \{0.1, 0.5, 1, 2, 10\}$  con  $n=25$ .

Seguimos observando la tendencia de rechazar con más facilidad alternativas con función de densidad cóncavas hacia abajo ( $p, q > 1$ ), pero en general todos los tests son bastante potentes para estas alternativas, sobretodo el test de Kuiper, los tres propuestos por Jin Zhang y el test de Sukhatme.

Al igual que pasaba en la Tabla 4.5, el Test de la Razón de Verosimilitudes da potencias nulas para parámetros de la beta menores que 1. Este problema es debido a la definición dada para nuestra Región Crítica.

Se observan varios tests que no rechazan un 100 % de las veces cuando los parámetros son más cercanos a 1, lo que parece en cierta parte lógico, la uniformidad aparece en una  $\beta(p, p)$  cuando  $p$  tiende a 1.

$\beta(p, q)$	$p = q = 0.1$	$p = q = 0.5$	$p = q = 1$	$p = q = 2$	$p = q = 10$
TRV	0.0000	0.0000	0.0508	0.9468	1.0000
Kolmogorov-Smirnov	1.0000	0.7592	0.0474	0.5783	1.0000
Pearson Chi Cuadrado	1.0000	0.6314	0.0551	0.7927	1.0000
Cramer-von Mises	1.0000	0.8470	0.0508	0.7361	1.0000
Anderson Darling	1.0000	0.9922	0.0510	0.9134	1.0000
Kuiper	1.0000	0.9844	0.0509	0.9860	1.0000
Zhang ZK	1.0000	0.9960	0.0433	0.9345	1.0000
Zhang ZC	1.0000	0.9991	0.0492	0.9969	1.0000
Zhang ZA	1.0000	0.9981	0.0480	0.9978	1.0000
Greenwood	1.0000	0.6313	0.0513	0.6519	1.0000
Q	1.0000	0.8202	0.0475	0.7143	1.0000
Drubin Kolmogorov	1.0000	0.7480	0.0474	0.5009	1.0000
Kendall-Sherman	1.0000	0.7789	0.0501	0.4277	1.0000
Pearson Q1	1.0000	0.8474	0.0507	0.3812	1.0000
Sukhatme	1.0000	1.0000	0.0488	1.0000	1.0000

Tabla 4.6: Familia alternativa:  $\beta(p, q)$  con  $p = q \in \{0.1, 0.5, 1, 2, 10\}$  con  $n=100$ .

## 2. Betas asimétricas

Si los parámetros de la beta son distintos obtenemos funciones de densidad asimétricas, pero dependiendo el valor que tomen con respecto al 1 veremos diferentes tipos de asimetrías.

- Para la primera clase de alternativas fijamos  $p = 1$  y hacemos variar  $q \in \{0.1, 0.5, 1, 2, 10\}$ . En la Figura 4.4 se dibujan sus densidades.

Por ser distribuciones asimétricas, no se espera que los tests encuentren dificultades para detectar la no uniformidad, es decir, se espera tener una tabla donde todas las potencias sean altas.

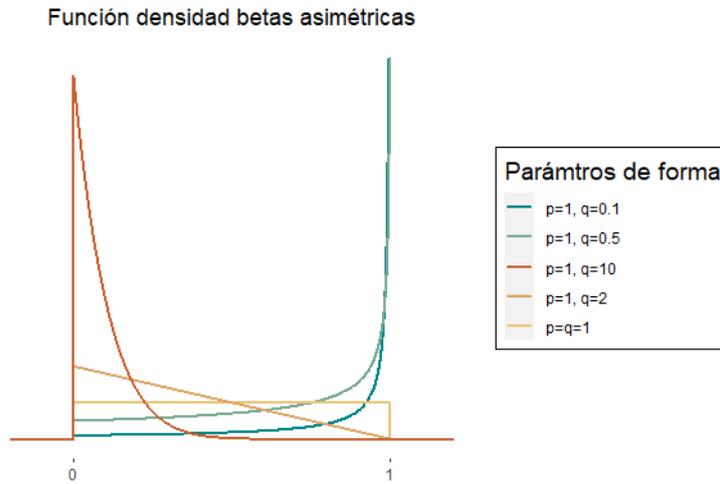


Figura 4.4: Función de densidad  $\beta(p, q)$  con  $p = 1$ ,  $q \in \{0.1, 0.5, 1, 2, 10\}$ .

En efecto, en la Tabla 4.7, incluso para un tamaño muestral pequeño ( $n = 25$ ), la primera y quinta columna, donde los parámetros más se alejan de 1, rechazan la hipótesis de uniformidad casi el 100 % de las veces.

Podemos ver cómo el TRV es de los menos potentes en esta tabla para alternativas donde  $q < 1$ . Esto, al igual que en la Tabla 4.6 se debe a que el máximo de la muestra es muy cercano a 1, y se puede ver que, para alternativas como  $\beta(1, 0.5)$ , donde la función tiene más densidad en 0 que la  $\beta(1, 0.1)$ , la potencia es incluso más baja, es decir, el test tiene una mayor duda acerca de si eso es una uniforme estándar.

Además apreciamos que el test de Sukhatme es poco potente para la alternativa  $\beta(1, 2)$ , cuya densidad se asemeja más a una distribución triangular que a una  $\mathcal{U}(0, 1)$ . En la última sección de este capítulo veremos qué pasa con esta alternativa y este estadístico.

$\beta(1, q)$	$q = 0.1$	$q = 0.5$	$q = 1$	$q = 2$	$q = 10$
TRV	0.6202	0.1017	0.0524	0.5630	1.0000
Kolmogorov-Smirnov	1.0000	0.7632	0.0463	0.8037	1.0000
Pearson Chi Cuadrado	1.0000	0.6219	0.0481	0.7376	1.0000
Cramer-von Mises	1.0000	0.8114	0.0501	0.8667	1.0000
Anderson Darling	1.0000	0.8941	0.0511	0.8609	1.0000
Kuiper	1.0000	0.5570	0.0536	0.5941	1.0000
Zhang ZK	1.0000	0.8746	0.0501	0.7956	1.0000
Zhang ZC	1.0000	0.9080	0.0550	0.8476	1.0000
Zhang ZA	1.0000	0.8820	0.0551	0.8851	1.0000
Greenwood	0.9997	0.3797	0.0452	0.4352	1.0000
Q	1.0000	0.5254	0.0509	0.4897	1.0000
Drubin Kolmogorov	1.0000	0.4649	0.0515	0.3529	1.0000
Kendall-Sherman	1.0000	0.4739	0.0495	0.3160	1.0000
Pearson Q1	0.9996	0.6107	0.0484	0.6111	1.0000
Sukhatme	1.0000	0.7522	0.0498	0.0495	0.9998

Tabla 4.7: Familia alternativa:  $\beta(1, q)$  y  $q \in \{0.1, 0.5, 1, 2, 10\}$   $n=25$ .

Como siempre, se hace un estudio equivalente aumentando el tamaño muestral.

Para  $n = 100$  obtenemos la Tabla 4.8, donde se aprecia una mayor cantidad de tests que rechazan el 100 % de las veces la uniformidad para valores de  $q$  más cercanos a 1. Para valores de  $q < 1$ , el TRV sigue siendo el test menos potente, por lo ya mencionado anteriormente, y para  $q > 1$ , en concreto  $q = 2$ , el test de Sukhatme sigue presentando una potencia próxima al nivel de significación.

$\beta(1, q)$	$q = 0.1$	$q = 0.5$	$q = 1$	$q = 2$	$q = 10$
TRV	0.6231	0.0894	0.0508	0.8406	1.0000
Kolmogorov-Smirnov	1.0000	0.9994	0.0474	1.0000	1.0000
Pearson Chi Cuadrado	1.0000	0.9975	0.0551	0.9999	1.0000
Cramer-von Mises	1.0000	0.9998	0.0508	1.0000	1.0000
Anderson Darling	1.0000	1.0000	0.0510	1.0000	1.0000
Kuiper	1.0000	0.9970	0.0509	0.9988	1.0000
Zhang ZK	1.0000	1.0000	0.0433	1.0000	1.0000
Zhang ZC	1.0000	1.0000	0.0492	1.0000	1.0000
Zhang ZA	1.0000	1.0000	0.0480	1.0000	1.0000
Greenwood	1.0000	0.7988	0.0513	0.9567	1.0000
Q	1.0000	0.9290	0.0475	0.9859	1.0000
Drubin Kolmogorov	1.0000	0.8674	0.0474	0.8793	1.0000
Kendall-Sherman	1.0000	0.8930	0.0501	0.7999	1.0000
Pearson Q1	1.0000	0.9866	0.0507	0.9974	1.0000
Sukhatme	1.0000	0.9982	0.0488	0.0519	1.0000

Tabla 4.8: Familia alternativa:  $\beta(1, q)$  y  $q \in \{0.1, 0.5, 1, 2, 10\}$   $n=100$ .

- Fijando ahora  $p = 2$ , y haciendo variar de la misma forma anterior el segundo parámetro de forma, obtenemos las densidades que aparecen en la Figura 4.5.

Se puede apreciar que, en este caso, no existe ninguna distribución equivalente a una  $\mathcal{U}(0, 1)$ , por lo que esperamos que las potencias de todos nuestros estadísticos salgan altas.

Para muestras de tamaño 25, podemos observar potencias altas en prácticamente todas las columnas de la Tabla 4.9, sin embargo, para la alternativa  $\beta(2, 2)$ , cuya densidad es una parábola cóncava hacia abajo centrada en mitad del intervalo, es decir, cuya densidad es la más parecida a una uniforme estándar que hay en

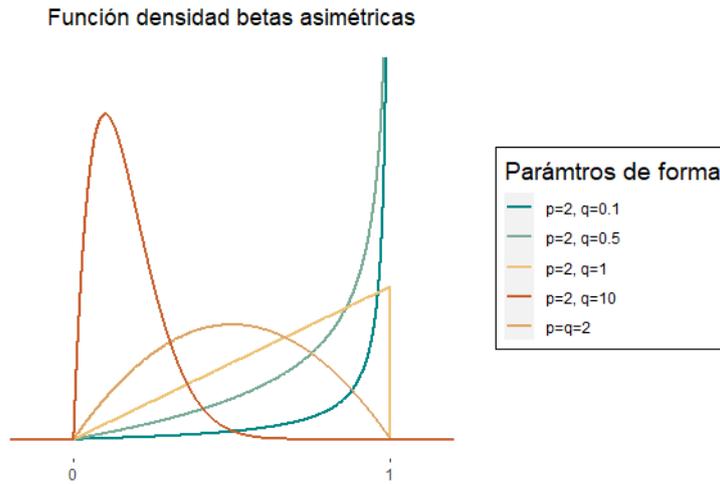


Figura 4.5: Función de densidad  $\beta(p, q)$  con  $p = 2, q \in \{0.1, 0.5, 1, 2, 10\}$ .

la Tabla 4.9, los tests parecen ser menos potentes en general.

Cabe destacar la proximidad al nivel  $\alpha$  en los tests de Kolmogorov-Smirnov, Cramer-von Mises, Anderson Darling, el test de Quesenberry y Miller o el test de la Probabilidad Producto de Pearson, cuya potencia es la que más se ajusta al nivel de significación.

No pasa lo mismo en la Tabla 4.10, donde se ha aumentado el tamaño de las muestras a 100. Aún siendo la cuarta columna (donde se considera una  $\beta(2, 2)$ ) la columna con potencias más bajas, nuestros tests ya no parecen dudar de la uniformidad de nuestras alternativas.

$\beta(2, q)$	q=0.1	q=0.5	q=1	q=2	q=10
TRV	0.9561	0.7414	0.5630	0.6487	1.0000
Kolmogorov-Smirnov	1.0000	0.9993	0.8037	0.0853	1.0000
Pearson Chi Cuadrado	1.0000	0.9974	0.7376	0.2561	1.0000
Cramer-von Mises	1.0000	0.9999	0.8667	0.0714	1.0000
Anderson Darling	1.0000	1.0000	0.8609	0.0708	1.0000
Kuiper	1.0000	0.9880	0.5941	0.4500	1.0000
Zhang ZK	1.0000	0.9999	0.7956	0.1742	1.0000
Zhang ZC	1.0000	0.9998	0.8476	0.3426	1.0000
Zhang ZA	1.0000	0.9998	0.8851	0.4834	1.0000
Greenwood	1.0000	0.9428	0.4352	0.1031	1.0000
Q	1.0000	0.9795	0.4897	0.0729	1.0000
Drubin Kolmogorov	1.0000	0.9507	0.3529	0.1062	1.0000
Kendall-Sherman	1.0000	0.9590	0.3160	0.1035	1.0000
Pearson Q1	1.0000	1.0000	0.9220	0.0481	1.0000
Sukhatme	1.0000	0.9549	0.0495	0.7311	0.2398

Tabla 4.9: Familia alternativa:  $\beta(2, q)$  con  $q \in \{0.1, 0.5, 1, 2, 10\}$   $n=25$ .

- Para finalizar con las betas asimétricas, fijamos un valor de  $p$  menor que 1, por ejemplo  $p = 0.5$ , y hacemos variable la  $q \in \{0.1, 0.5, 1, 2, 10\}$ . Para tener una noción de cuánto de uniformes son nuestras alternativas, podemos ver en la Figura 4.6 sus funciones de densidad.

En la Tabla 4.11 aparecen las potencias de nuestros estadísticos bajo esas distribuciones alternativas, que como veíamos en la Figura 4.6 nada tienen que ver con una densidad uniforme. La columna con potencias más bajas es para la distribución  $\beta(\frac{1}{2}, \frac{1}{2})$ , ya que, al ser simétrica, es la que más se parece a una uniforme estándar.

$\beta(2, q)$	q=0.1	q=0.5	q=1	q=2	q=10
TRV	0.9878	0.9168	0.8406	0.9468	1.0000
Kolmogorov-Smirnov	1.0000	1.0000	1.0000	0.5783	1.0000
Pearson Chi Cuadrado	1.0000	1.0000	0.9999	0.7927	1.0000
Cramer-von Mises	1.0000	1.0000	1.0000	0.7361	1.0000
Anderson Darling	1.0000	1.0000	1.0000	0.9134	1.0000
Kuiper	1.0000	1.0000	0.9988	0.9860	1.0000
Zhang ZK	1.0000	1.0000	1.0000	0.9345	1.0000
Zhang ZC	1.0000	1.0000	1.0000	0.9969	1.0000
Zhang ZA	1.0000	1.0000	1.0000	0.9978	1.0000
Greenwood	1.0000	1.0000	0.9567	0.6519	1.0000
Q	1.0000	1.0000	0.9859	0.7143	1.0000
Drubin Kolmogorov	1.0000	1.0000	0.8793	0.5009	1.0000
Kendall-Sherman	1.0000	1.0000	0.7999	0.4277	1.0000
Pearson Q1	1.0000	1.0000	1.0000	0.3812	1.0000
Sukhatme	1.0000	1.0000	0.0519	1.0000	0.9527

Tabla 4.10: Familia alternativa:  $\beta(2, q)$  con  $q \in \{0.1, 0.5, 1, 2, 10\}$   $n=100$ .

A pesar de ser un tamaño muestral pequeño ( $n = 25$ ), todos los tests rechazan con seguridad todas las clases de alternativas. Es por eso que no se ha realizado el mismo estudio con un número mayor de muestras, ya que los resultados iban a ser prácticamente iguales.

No obstante, seguimos encontrando dificultades en el TRV con parámetros menores que 1, es un problema que no se soluciona aumentando el tamaño muestral, ya que la distribución de su estadístico no es asintótica, y es debido a la definición de la Región Crítica.

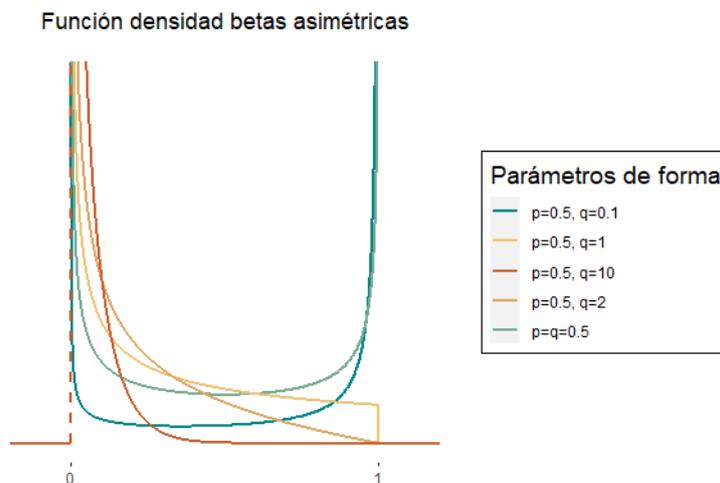


Figura 4.6: Función de densidad  $\beta(p, q)$  con  $p = 0.5, q \in \{0.1, 0.5, 1, 2, 10\}$ .

- Para concluir esta sección, donde se han considerado varias alternativas asimétricas distintas de la familia  $\beta(p, q)$ , notamos que los tests tienen más problemas en detectar la uniformidad si uno de nuestros parámetros es igual a 1. Cuando ambos son distintos a la unidad, el mayor problema lo tienen cuando dichos parámetros coinciden, debido a su simetría.

Además, en casos muy concretos, hay tests que fallan. Por ejemplo, el TRV, cuando el máximo y mínimo de la muestra son cercanos a 0 y 1 respectivamente, el test de Sukhatme para alternativas que se asemejan a triangulares, o algún test tradicional como Kolmogorov-Smirnov o Cramer-von Mises contra una alternativa  $\beta(2, 2)$ .

$\beta(0.5, q)$	q=0.1	q=0.5	q=1	q=2	q=10
TRV	0.1320	0.0013	0.1017	0.7414	1.0000
Kolmogorov-Smirnov	0.9999	0.2391	0.7632	0.9993	1.0000
Pearson Chi Cuadrado	0.9974	0.1801	0.6219	0.9974	1.0000
Cramer-von Mises	0.9999	0.2399	0.8114	0.9999	1.0000
Anderson Darling	1.0000	0.6119	0.8941	1.0000	1.0000
Kuiper	1.0000	0.5080	0.5570	0.9880	1.0000
Zhang ZK	1.0000	0.6735	0.8746	0.9999	1.0000
Zhang ZC	1.0000	0.7900	0.9080	0.9998	1.0000
Zhang ZA	1.0000	0.6933	0.8820	0.9998	1.0000
Greenwood	0.9972	0.3314	0.3797	0.9428	1.0000
Q	0.9999	0.4696	0.5254	0.9795	1.0000
Drubin Kolmogorov	1.0000	0.3998	0.4649	0.9507	1.0000
Kendall-Sherman	0.9997	0.4312	0.4739	0.9590	1.0000
Pearson Q1	0.8483	0.4391	0.9401	0.9998	1.0000
Sukhatme	1.0000	0.8761	0.7522	0.9549	1.0000

Tabla 4.11: Familia alternativa:  $\beta(0.5, q)$  con  $q \in \{0.1, 0.5, 1, 2, 10\}$   $n=25$ .

#### 4.4. Familia Trapezoidal

La distribución trapezoidal es una distribución de probabilidad continua, cuya función de densidad toma formas geométricas como el trapecio, triángulo o rectángulo dependiendo de sus parámetros  $a, m_1, m_2$  y  $b$ . Estos valores corresponden al valor mínimo, moda inferior, moda superior y valor máximo, respectivamente.

Su función de densidad viene dada por

$$f(x) = \begin{cases} 0 & \text{si } x < a \\ \mu \left( \frac{x-a}{m_1-a} \right) & \text{si } a \leq x < m_1 \\ \mu & \text{si } m_1 \leq x \leq m_2 \\ \mu \left( \frac{b-x}{b-m_2} \right) & \text{si } m_2 < x \leq b \\ 0 & \text{si } b < x \end{cases}$$

donde  $\mu = \frac{2}{b-a+m_2-m_1}$  es una constante normalizadora.

**Función densidad trapezoidal**

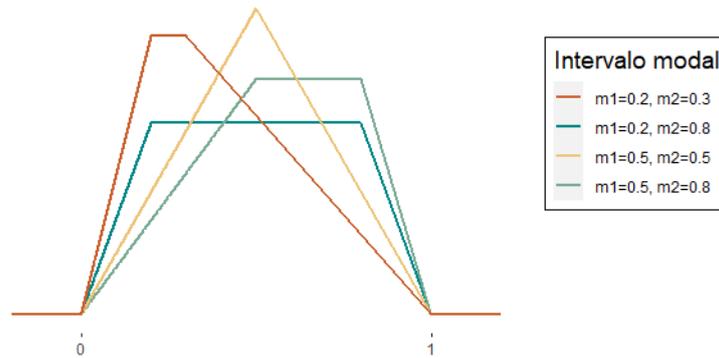


Figura 4.7: Función de densidad trapezoidal con distintos intervalos modales.

En nuestro caso fijaremos  $a = 0$ ,  $b = 1$  y tomaremos  $m_2 = 1 - m_1$  para tener una distribución simétrica y que sea más difícil distinguir la uniformidad, así, nuestro intervalo modal  $(m_1, m_2)$  tendrá longitud  $1 - 2m_1$ . De esta manera, el único parámetro variable será  $m_1 \in \{0, 0.1, 0.2, 0.3, 0.4\}$ , sin considerar  $m_1 = m_2 = 0.5$  ya que pasaríamos a tener una distribución triangular, de la cual hablaremos más adelante.

Podemos fijarnos que para  $m_1 = 0$  obtenemos idénticamente una distribución  $\mathcal{U}(0, 1)$ , por lo que la potencia para esa muestra debería rondar el valor del nivel de significación 0.05.

#### 4.4.1. La potencia en función del intervalo modal

En la Figura 4.8 podemos observar las densidades de las distribuciones alternativas que proporcionaremos a nuestros tests. Se puede intuir que la uniformidad desaparecerá a medida que la moda inferior crezca.

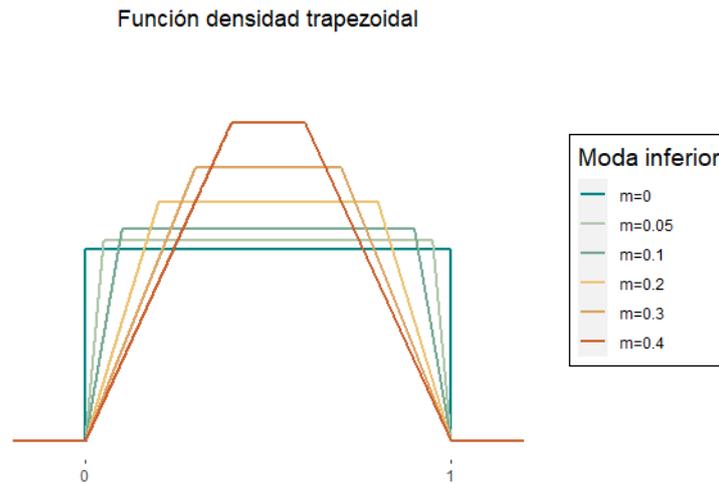


Figura 4.8: Función de densidad trapezoidal con moda inferior  $m_1 \in \{0, 0.05, 0.1, 0.2, 0.3, 0.4\}$ .

Tomando los intervalos modales que reflejan en la Figura 4.8, y proporcionando muestras de tamaño 25 a nuestros estadísticos, obtenemos la Tabla 4.12. En ella podemos observar que la primera columna, que hace referencia a una uniforme estándar, se comporta bien, dando valores cercanos a  $\alpha = 0.05$ .

Para  $m_1 = 0.05$ , ninguno de los tests, salvo el TRV, son capaces de distinguir la distribución planteada de una uniforme estándar. Parece lógico, ya que son distribuciones con densidades muy parecidas y teniendo en cuenta el tamaño muestral escogido.

Si nos alejamos de la uniformidad, para  $m_1 = 0.1$  estamos en una situa-

ción parecida, aunque tests como TRV, Kuiper y ZA de Zhang nos presentan potencias superiores a 0.1, son tests bastante potentes para estas alternativas.

En la última columna, la distribución más dispar a una uniforme, aparecen potencias más altas, pero no lo suficiente como esperamos. De hecho, el test del Producto de Pearson sigue devolviendo una potencia cercana al nivel  $\alpha$ .

$Tp(0, m_1, 1 - m_1, 1)$	$m_1 = 0$	$m_1 = 0.05$	$m_1 = 0.1$	$m_1 = 0.2$	$m_1 = 0.3$	$m_1 = 0.4$
TRV	0.0543	0.1363	0.2982	0.5924	0.7153	0.7644
Kolmogorov-Smirnov	0.0485	0.0444	0.0431	0.0510	0.0907	0.1509
Pearson Chi Cuadrado	0.0490	0.0485	0.0557	0.1106	0.2414	0.4476
Cramer-von Mises	0.0490	0.0421	0.0391	0.0456	0.0739	0.1325
Anderson Darling	0.0471	0.0335	0.0310	0.0410	0.0806	0.1492
Kuiper	0.0553	0.0677	0.1030	0.2584	0.4961	0.6615
Zhang ZK	0.0542	0.0395	0.0471	0.0962	0.2063	0.2949
Zhang ZC	0.0512	0.0376	0.0653	0.2110	0.4143	0.5423
Zhang ZA	0.0542	0.0653	0.1180	0.3476	0.5563	0.6656
Greenwood	0.0452	0.0320	0.0306	0.0578	0.1293	0.1986
Q	0.0486	0.0285	0.0219	0.0335	0.0916	0.1635
Drubin Kolmogorov	0.0480	0.0321	0.0361	0.0719	0.1292	0.1881
Kendall-Sherman	0.0504	0.0304	0.0387	0.0712	0.1294	0.1855
Pearson Q1	0.0474	0.0327	0.0349	0.0427	0.0498	0.0539
Sukhatme	0.0515	0.0776	0.1828	0.5246	0.7922	0.8972

Tabla 4.12: Familia alternativa: Trapezoidal de tamaño 25 con  $m_1 \in \{0, 0.05, 0.1, 0.2, 0.3, 0.4\}$ .

Podemos observar qué pasa si aumentamos dicho tamaño muestral. En la Tabla 4.13, se realiza el mismo estudio para muestras de tamaño 100, donde son más los tests que detectan la no uniformidad en la  $Tp(0, 0.05, 0.95, 1)$ . Entre ellos está el TRV, que ya era potente con  $n = 25$ , y se le suman

el test de Kuiper, ZC y ZA de Zhang y el test de Sukhatme, que ofrecen una potencia mayor a 0.1 para dicha distribución. Son solo los tests de Kolmogorov-Smirnov y Cramer-von Mises que mantienen una potencia baja para  $m_1 = 0.1$ .

$Tp(0, m_1, 1 - m_1, 1)$	$m_1 = 0$	$m_1 = 0.05$	$m_1 = 0.1$	$m_1 = 0.2$	$m_1 = 0.3$	$m_1 = 0.4$
TRV	0.0449	0.6326	0.8529	0.9438	0.9642	0.9726
Kolmogorov-Smirnov	0.0480	0.0490	0.0657	0.2468	0.6872	0.8923
Pearson Chi Cuadrado	0.0554	0.0647	0.1064	0.3444	0.7781	0.9740
Cramer-von Mises	0.0521	0.0529	0.0742	0.3208	0.8220	0.9712
Anderson Darling	0.0510	0.0573	0.1178	0.6376	0.9612	0.9960
Kuiper	0.0508	0.1051	0.3027	0.8991	0.9968	1.0000
Zhang ZK	0.0439	0.0861	0.3443	0.8460	0.9696	0.9918
Zhang ZC	0.0490	0.2501	0.7066	0.9821	0.9991	1.0000
Zhang ZA	0.0463	0.3848	0.8212	0.9908	0.9995	1.0000
Greenwood	0.0511	0.0523	0.1604	0.5305	0.7732	0.8840
Q	0.0503	0.0379	0.1209	0.5531	0.8302	0.9320
Drubin Kolmogorov	0.0476	0.0571	0.1128	0.3418	0.6025	0.7725
Kendall-Sherman	0.0504	0.0613	0.1082	0.2818	0.5138	0.6907
Pearson Q1	0.0488	0.0658	0.1269	0.2771	0.4231	0.5239
Sukhatme	0.0480	0.3081	0.7775	0.9983	0.9999	1.0000

Tabla 4.13: Familia alternativa: Trapezoidal de tamaño 100 con  $m_1 \in \{0, 0.05, 0.1, 0.2, 0.3, 0.4\}$ .

A la vista de la Tabla 4.12 y la Tabla 4.13 parece claro que, obviando el caso donde  $m_1 = 0$ , el caso más problemático es en el que el intervalo modal es  $(0.05, 0.95)$ . Es donde, a la gran mayoría de nuestros tests, se le complica la detección de la uniformidad. Por ello, se realizará un estudio para la alternativa  $Tp(0, 0.05, 0.95, 1)$  haciendo variar el tamaño muestral.

#### 4.4.2. La potencia en función del tamaño de muestra

En el estudio anterior, las Tablas 4.12 y 4.13 reflejaban conclusiones algo dudosas para la distribución trapezoidal con  $m_1 = 0.05$  y  $m_2 = 0.95$ . Sien-

do una distribución que se asemeja bastante a la uniforme estándar, parece normal que los tests tengan mayor dificultad para detectar la uniformidad. Es por eso que haremos un estudio para la distribución  $Tp(0, 0.05, 0.95, 1)$  haciendo variar el tamaño muestral. Estos resultados se muestran en la Tabla 4.14.

$Tp(0, 0.05, 0.95, 1)$	$n = 10$	$n = 25$	$n = 50$	$n = 100$	$n = 250$	$n = 500$	$n = 1000$
TRV	0.0806	0.1363	0.3174	0.6326	0.8993	0.9675	0.9891
Kolmogorov-Smirnov	0.0451	0.0444	0.0469	0.0490	0.0588	0.0805	0.1577
Pearson Chi Cuadrado	0.0588	0.0485	0.0542	0.0647	0.0822	0.1102	0.1760
Cramer-von Mises	0.0420	0.0421	0.0439	0.0529	0.0633	0.0986	0.2256
Anderson Darling	0.0329	0.0335	0.0395	0.0573	0.1121	0.3172	0.8377
Kuiper	0.0546	0.0677	0.0680	0.1051	0.1956	0.4040	0.8354
Zhang ZK	0.0344	0.0395	0.0511	0.0861	0.5259	0.9544	0.9999
Zhang ZC	0.0218	0.0376	0.0825	0.2501	0.8580	0.9990	1.0000
Zhang ZA	0.0383	0.0653	0.1326	0.3848	0.9196	0.9997	1.0000
Greenwood	0.0343	0.0320	0.0379	0.0523	0.1705	0.4199	0.7983
Q	0.0289	0.0285	0.0284	0.0379	0.1463	0.4532	0.8689
Drubin Kolmogorov	0.0330	0.0321	0.0392	0.0571	0.1157	0.1682	0.3170
Kendall-Sherman	0.0274	0.0304	0.0419	0.0613	0.0975	0.1529	0.2445
Pearson Q1	0.0276	0.0327	0.0395	0.0658	0.1471	0.2847	1.0000
Sukhatme	0.0396	0.0776	0.1441	0.3081	0.7305	0.9661	1.0000

Tabla 4.14: Familia alternativa: Trapezoidal con intervalo modal (0.05,0.95) y tamaño variable

Se puede observar en la Tabla 4.14 la importancia de dar una muestra suficientemente grande no solo a tests basados en resultados asintóticos. Ningún test alcanza una potencia mayor a 0.1 para muestras de tamaño 10, y solo el TRV lo hace cuando la  $n = 25$ .

Podemos considerar los tests más potentes a los que, para tamaño muestral 50, superan una potencia de 0.1, estos son el TRV, ZA de Zhang y Sukhatme. Se le suman, para  $n = 100$ , ZC de Zhang y el test de Kuiper.

Sin embargo, podremos decir que no son test potentes bajo esta alternativa el test de Kolmogorov-Smirnov o el de Cramer-von Mises, ya que para muestras de tamaño 500 proporcionan una potencia que no alcanza el 0.1

## 4.5. Familia Triangular

La distribución triangular es una distribución de probabilidad continua que viene definida por 3 valores que coinciden con los vértices de un triángulo, denotados por  $a, b, c$ . Su función de densidad viene dada por

$$f(x) = \begin{cases} 0 & \text{si } x < a \\ \frac{2(x-a)}{(b-a)(c-a)} & \text{si } a \leq x < c \\ \frac{2}{b-a} & \text{si } x = c \\ \frac{2(b-x)}{(b-a)(b-c)} & \text{si } c < x \leq b \\ 0 & \text{si } b < x \end{cases}$$

Donde  $a$  sería el valor mínimo que toma la variable,  $b$  el valor máximo, y  $c$  sería la altura del triángulo, o en términos estadísticos, la moda. Puede verse como un caso particular de la distribución trapezoidal, donde  $m_1 = m_2 = c$ , es decir el intervalo modal sería el punto  $c$ .

En nuestro caso fijaremos  $a = 0$  y  $b = 1$  para asegurar que el soporte de la variable es el intervalo unidad, al igual que la distribución  $\mathcal{U}(0, 1)$ , y haremos variar la moda. En la Figura 4.9 quedan representadas todas las densidades que daremos a nuestros estadísticos. Se intuye que son distribuciones que, a simple vista, carecen de uniformidad, aunque nuestros estadísticos podrían confundirse para valores centrales de la  $c$ .

Esto se ve reflejado en la Tabla 4.15, donde presento la potencia para esas 5 distribuciones con  $c \in \{0, 0.25, 0.5, 0.75, 1\}$  y un tamaño muestral de  $n = 25$ .

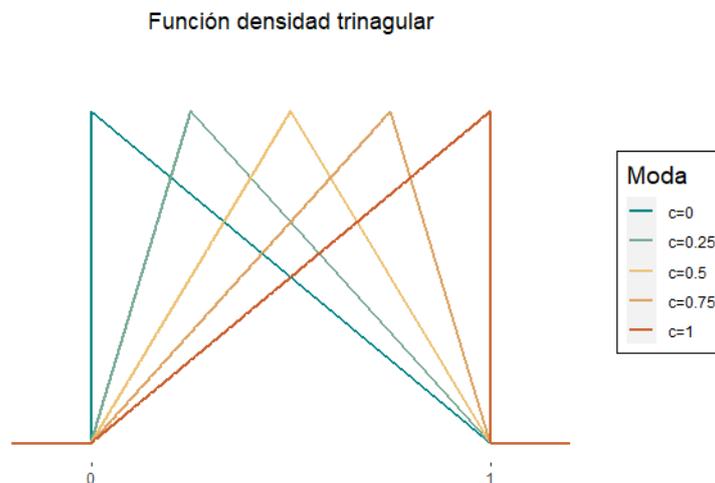


Figura 4.9: Función de densidad triangular con moda  $c \in \{0, 0.25, 0.5, 0.75, 1\}$ .

#### 4.5.1. La potencia en función de la moda

En la Tabla 4.15 se aprecian resultados muy interesantes. Se pueden observar, en casi toda la tabla, potencias altas, pero que en ningún caso llegan a 1, eso quiere decir que ningún test, logra distinguir el 100 % de las veces una triangular de una uniforme estándar. En un estudio posterior veremos si esto se debe al tamaño muestral.

Por otro lado, el test del Producto de Probabilidades de Pearson da potencias bajas para valores centrales de la moda, al igual que en la Tabla 4.12. Puede deberse a la definición de la Región Crítica a dos colas y por ser una distribución simétrica con más densidad en el centro del intervalo.

Lo que más puede llamar la atención de esta Tabla 4.15 es que, para distribuciones que nada tienen que ver con una  $\mathcal{U}(0, 1)$ , cuando la moda vale 0 o 1, el test de Sukhatme devuelve potencias bajas, muy próximas a  $\alpha = 0.05$ . Podemos pensar que es debido al bajo tamaño muestral.

$Triang(0, c, 1)$	$c = 0$	$c = 0.25$	$c = 0.5$	$c = 0.75$	$c = 1$
TRV	0.5435	0.7431	0.7791	0.7433	0.5389
Kolmogorov-Smirnov	0.7999	0.4515	0.1811	0.4480	0.8000
Pearson Chi Cuadrado	0.7383	0.4282	0.5198	0.4230	0.7375
Cramer-von Mises	0.8630	0.4618	0.1669	0.4593	0.8671
Anderson Darling	0.8582	0.4389	0.1871	0.4344	0.8628
Kuiper	0.5860	0.6902	0.7177	0.6911	0.5862
Zhang ZK	0.7913	0.4883	0.3282	0.4874	0.7836
Zhang ZC	0.8373	0.6468	0.5854	0.6408	0.8314
Zhang ZA	0.8795	0.7383	0.7000	0.7382	0.8750
Greenwood	0.4202	0.2661	0.2259	0.2593	0.4177
Q	0.4790	0.2543	0.1943	0.2503	0.4739
Drubin Kolmogorov	0.3472	0.2277	0.2209	0.2234	0.3442
Kendall-Sherman	0.3116	0.2175	0.2147	0.2186	0.3171
Pearson Q1	0.6135	0.0024	0.0553	0.5674	0.9250
Sukhatme	0.0500	0.6816	0.9202	0.6734	0.0496

Tabla 4.15: Familia alternativa: Triangular de tamaño 25 con modas distintas.

Para solventar las dudas en cuanto al tamaño muestral, la Tabla 4.16 recoge las potencias para las mismas alternativas con muestras de tamaño 100. Podemos apreciar que muchos de los tests alcanzan potencia 1 para casi todas las alternativas. Sin embargo, la potencia de Sukhatme se mantiene en torno al nivel de significación.

Es curioso porque esto se debe a que, la distribución del estadístico  $Q_2$  bajo  $H_0$  coincide, o se aproxima, a la distribución bajo la hipótesis de  $Triang(0, 0, 1)$  o  $Triang(0, 1, 1)$ . Simulando ambas distribuciones, obtenemos las densidades que aparecen en la Figura 4.10.

Es por esto por lo que el test no logra distinguir muestras uniformes de muestras cuyas densidades dibujan un triángulo rectángulo y así, acaba siendo un test poco potente bajo estas alternativas.

Triang(0,c,1)	c=0	c=0.25	c=0.5	c=0.75	c=1
TRV	0.8409	0.9676	0.9750	0.9665	0.8444
Kolmogorov-Smirnov	1.0000	0.9947	0.9318	0.9931	1.0000
Pearson Chi Cuadrado	0.9999	0.9835	0.9887	0.9810	0.9999
Cramer-von Mises	1.0000	0.9982	0.9865	0.9973	1.0000
Anderson Darling	1.0000	0.9994	0.9989	0.9995	1.0000
Kuiper	0.9996	0.9999	1.0000	0.9999	0.9994
Zhang ZK	1.0000	0.9985	0.9946	0.9989	1.0000
Zhang ZC	1.0000	0.9999	1.0000	1.0000	1.0000
Zhang ZA	1.0000	0.9999	1.0000	1.0000	1.0000
Greenwood	0.9579	0.9162	0.9119	0.9182	0.9564
Q	0.9863	0.9596	0.9546	0.9594	0.9869
Drubin Kolmogorov	0.8797	0.8277	0.8244	0.8194	0.8720
Kendall-Sherman	0.7981	0.7559	0.7554	0.7511	0.8023
Pearson Q1	0.9978	0.0057	0.5560	0.9993	1.0000
Sukhatme	0.0480	0.9999	1.0000	1.0000	0.0489

Tabla 4.16: Familia alternativa: Triangular de tamaño 100 con modas distintas.

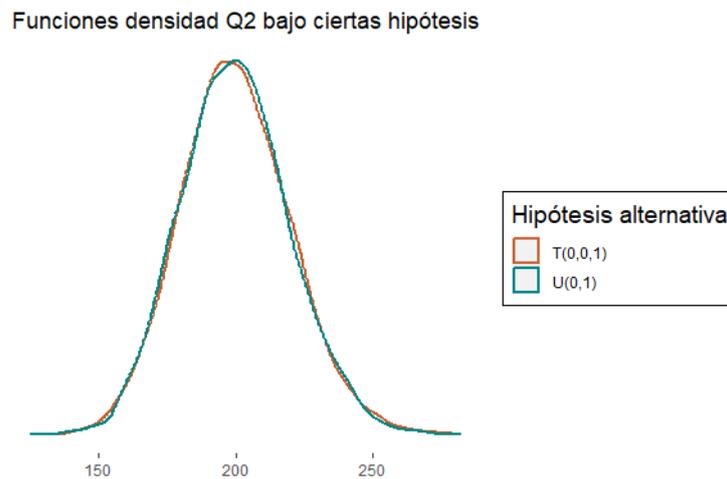


Figura 4.10: Función de densidad estimada del estadístico  $Q_2$  bajo la hipótesis de uniformidad y bajo la alternativa  $Triang(0, 0, 1)$ .

## Capítulo 5

# Conclusiones

Conocida la importancia de saber si una muestra proviene de una distribución uniforme estándar, se han considerado en este trabajo catorce de los estadísticos aplicados a test de bondad de ajuste más importantes y conocidos. Seis de ellos eran tests tradicionales de bondad de ajuste, sin embargo los ocho restantes se han construido a lo largo de la historia específicamente para testear uniformidad. Para establecer una comparativa entre los tests, se ha realizado un análisis de la potencia de cada uno, ofreciéndoles muestras con tamaños variables de diferentes familias de distribuciones, como por ejemplo: la propia familia uniforme, la familia beta, la trapezoidal y la triangular, haciendo variar en todas ellas los parámetros, pero siempre considerando como soporte el intervalo  $[0,1]$ . Nuestro objetivo era llegar a conocer qué tests eran mejores en función de la potencia. Es decir, qué tests presentaban una potencia en torno al nivel de significación si la distribución subyacente era una uniforme estándar y qué tests daban una potencia alta cuando la distribución subyacente no era uniforme.

Dependiendo de la alternativa que les ofrecíamos, si compartían alguna característica con nuestra distribución bajo  $H_0$  (cierta simetría, cierta uniformidad...) dar una potencia correcta dependía en varias ocasiones del

tamaño muestral. Incluso, debido al tamaño muestral, cuando la muestra subyacente era una uniforme estándar, estadísticos como  $Q_1$  o  $Q_2$  rechazaban dicha uniformidad por problemas con la tolerancia del programa.

Otros problemas que nos hemos encontrado a lo largo del Capítulo 4 fueron la definición de nuestras Regiones Críticas ya que, por ejemplo, era complicado conocer la distribución conjunta del estadístico de contraste en el Test de la Razón de Verosimilitudes. Además, en el caso de la  $\beta(2, 1)$  o de las triangulares con moda 0 o 1, la distribución de nuestro estadístico  $Q_2$  era la misma si consideramos como distribución de referencia una  $\mathcal{U}(0, 1)$ , una  $\beta(2, 1)$ , una  $Triang(0, 0, 1)$  o una  $Triang(0, 1, 1)$ .

También se ha notado que, en general, los tests tradicionales no proporcionan potencias muy altas fuera de la hipótesis nula, incluso para tamaños muestrales suficientemente altos, en comparación con los tests específicamente diseñados para uniformidad. En concreto, el test de Kolmogorov-Smirnov junto con el de Cramer-von Mises son los menos potentes de los que se han estudiado.

No obstante, se han encontrado tests que se podrían considerar más potentes para las alternativas consideradas en este trabajo. Tests como el TRV, a la excepción de la familia de betas, el test de Kuiper, los tests ZC y ZA de Zhang, y el test de Sukhatme, a excepción de la familia triangular, han mostrado potencias altas mucho antes que los demás tests, es decir, con tamaños muestrales más pequeños son capaces de detectar la no uniformidad antes que el resto.

En este trabajo se ha restringido el estudio a tests de uniformidad en el intervalo  $[0,1]$ , pero existen ocasiones donde será necesario testear la unifor-

midad de forma general, con parámetros cualesquiera. En este caso, estadísticos como los propuestos por Zhang no podrían ser utilizados, ya que, por su definición, fallan bajo ciertas hipótesis. Sin embargo, el test de Kolmogorov-Smirnov puede usarse aplicándole la corrección de Lilliefors. De una manera sencilla, el estadístico de Cramer-von Mises también puede adaptarse para que funcione con muestras uniformes fuera del intervalo unidad. Por otro lado, haciendo uso de la Proposición 2.2, se concluye que cualquier test de uniformidad en el intervalo unidad se puede adaptar para realizar bondad de ajuste a cualquier otra distribución totalmente especificada.

# Bibliografía

- [1] Anderson, T & Darling, D. Asymptotic theory of certain “goodness of fit” criteria based on stochastic processes, *The Annals of Mathematical Statistics*, 23(2) (1952), 193–212.
- [2] Bartholomew, D. J. Note on the use of Sherman’s statistic as a test for randomness, *Biometrika*, 41 (1954), 556–568.
- [3] Cramér, H. On the composition of elementary errors. *Scandinavian Actuarial Journal*, (1928), 13–74.
- [4] Csorgo, S. & Faraday, J. The exact and asymptotic distributions of Cramér-von Mises statistics. *Journal of the Royal Statistical Society Series B*, 58(1) (1996), 221–234.
- [5] Durbin, J. Some methods of constructing exact tests. *Biometrika*, 48 (1961), 41–55.
- [6] Durbin, J. Testing for Serial Correlation in Systems of Simultaneous Regression Equations. *Biometrika*, 44(3/4) (1957), 370.
- [7] Fisher, R. A. The goodness of fit of regression formulae, and the distribution of regression coefficients. *Journal of the Royal Statistical Society*, 85(4) (1922), 597–612.
- [8] Greenwood, M. The statistical study of infectious diseases, *Journal of Royal Statistical Society Series A* (1946), 85–110.

- [9] Kolmogorov, A. Sulla determinazione empirica di una legge di distribuzione, *Giornale dell'Istituto Italiano degli Attuari*, 4 (1933), 83–91.
- [10] Kuiper, N. H. Tests concerning random points on a circle, *Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen, Series A* (1960), 38–47.
- [11] Neyman, J. & Pearson, E. On the problem of the most efficient tests of statistical hypotheses. *Philosophical Transactions of the Royal Society of London. Series A* (1933), 289–337.
- [12] Pearson, E. S. The probability integral transformation for testing goodness of fit and combining independent tests of significance, *Biometrika*, 30 (1939), 134–148.
- [13] Pearson, K. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 50 (1900), 157–175.
- [14] Pearson, K. On a method of determining whether a sample of size  $n$  supposed to have been drawn from a parent population having a known probability integral has probably been drawn at random, *Biometrika*, 25 (1933), 379–410.
- [15] Quesenberry, C. P. & Miller, F. L. Jr. Power studies of some tests for uniformity. *Journal of Statistical Computation and Simulation*, 5 (1977), 169–191.
- [16] Rao J. S. & Sethuraman, J. Weak convergence of empirical distribution functions of random variables subject to perturbations and scale factors, *The Annals of Statistics*, 3 (1975), 299–313.

- [17] Riley M.C., Clare , A., King R. D. Locational distribution of gene functional classes in *Arabidopsis thaliana* .BMC Bioinformatics.(2007), 8–112.
- [18] Sherman, B. A random variable related to spacings of sample values, The Annals of Mathematical Statistics, 21 (1950), 339–61.
- [19] Smirnov, N. Sur la distribution de  $\omega_2$  critérium de von Mises. Rec. Math., 2(44) (1937), 973–993.
- [20] Smirnov, N. Table for estimating the goodness of fit of empirical distributions, The Annals of Mathematical Statistics, 19 (1948) 279–281.
- [21] Tucker H. G. A Generalization of the Glivenko–Cantelli Theorem, The Annals of Mathematical Statistics, 30 (3) (1959), 828—830.
- [22] Von Neumann, J. & Ulam, S. Monte Carlo method. National Bureau of Standards Applied Mathematics Series, 12 (1951), 36.
- [23] Wilks, S. S. The large-sample distribution of the likelihood ratio for testing composite hypotheses. Annals of Mathematical Statistics. 9 (1) (1938), 60–62.
- [24] Zhang, J. Powerful goodness-of-fit tests based on the likelihood ratio, Journal of Royal Statistical Society Series B, 64 (2002), 281–294.