



UNIVERSIDAD DISTRITAL

UNIVERSIDAD DISTRITAL FRANCISCO JOSÉ DE
CALDAS

FACULTAD DE CIENCIAS Y EDUCACIÓN

SOBRE LA DISTRIBUCIÓN DE RAYLEIGH
GENERALIZADA Y TRANSMUTADA

QUE PARA OBTENER EL TÍTULO DE

MATEMÁTICO

P R E S E N T A

FREDY ALEXANDER VELÁSQUEZ BOHÓRQUEZ
COD: 20102167005

DIRECTOR

LUIS FERNANDO VILLARRAGA POVEDA



UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS

BOGOTÁ, 2019

Dedicado a mis padres Isabel y Néstor y mi hermana Angélica

Agradecimientos

Ante todo quiero dar gracias a mi familia, ellos no solo constituyen una gran motivación para mis acciones, además de su incondicional apoyo y amor son la frescura que me reconforta ante las dificultades, mi padre con su ejemplo, mi madre con su cariño y mi hermana con su apoyo son quienes hacen que para mí sea posible lo imposible. De igual manera agradezco de forma muy especial a todo el equipo humano de la Universidad Distrital Francisco José de Caldas quienes con dedicación y esfuerzo permiten que se lleven a cabo las aspiraciones tanto mías como de mis compañeros, al personal docente, en particular al profesor Fernando Villarraga quien con su paciente instrucción y enseñanza me motivó a continuar en el camino del aprendizaje. Asimismo expreso mi gratitud a mis demás familiares, compañeros, amigos y allegados que hicieron de la experiencia del aprendizaje fuese no solo académica sino una huella imborrable de vida.

Resumen

En el presente trabajo se pretende extender la distribución de Rayleigh generalizada usando el mapa de transmutación de rango cuadrático estudiado por Shaw y Buckley en [1] de donde sale un modelo del cual se estudiarán las principales propiedades y características para posteriormente realizar inferencia estadística y mostrar una aplicación con datos reales. Finalmente, se presentan las principales conclusiones

Abstract

This work try to extend the generalized Rayleigh distribution using the quadratic range transmutation map studied by Shaw and Buckley in [1] from where a model is left out of which the main properties and characteristics will be studied to later perform statistical inference and show an application with real data. Finally, the main conclusions are presented

Índice general

Agradecimientos	II
Resumen	III
Abstract	IV
1. Justificación	1
2. Objetivos	2
2.1. Objetivo General	2
2.2. Objetivos Específicos	2
3. Marco conceptual	3
3.1. Funciones Aditivas de Conjunto	3
3.2. Espacio Medible	5
3.3. Espacio probabilístico	6
3.4. Variable aleatoria	6
3.4.1. Tipos de variables aleatorias	6
3.4.2. Variable aleatoria continua	6
3.5. Función de distribución de probabilidad	7
4. Análisis de la distribución	8
4.1. Contexto	8
4.2. Rango de transmuciones, asimetría y curtosis	9
4.2.1. Transmutación cuadrática	10

<i>ÍNDICE GENERAL</i>	VI
4.2.2. Asimetría	12
4.2.3. Curtosis	12
5. Distribución Rayleigh generalizada transmutada	14
5.1. Algunas propiedades	15
5.2. Distribuciones relacionadas	17
5.3. Momentos	18
6. Inferencia	20
6.1. Estimadores de máxima verosimilitud	20
6.1.1. Verosimilitud	21
6.1.2. Estimador de máxima verosimilitud	21
6.1.3. Método para obtener el E.M.V	22
6.1.4. E.M.V para la distribución RGT	24
6.2. Matriz de información observada	25
6.3. Análisis de confiabilidad	26
6.3.1. Pruebas de bondad de ajuste	26
6.3.2. Función de confiabilidad y tasa de riesgo	28
7. Estimación de parámetros	30
7.0.1. Método de inversión de la función distribución	30
7.0.2. Ilustración con datos reales	33
7.0.3. Uso en análisis de software	34
8. Conclusiones	36
9. Referencias	38

Índice de figuras

3.1. Caso particular medida de Dirac para conjunto de 3 elementos	5
4.1. Gráfica del coeficiente de asimetría y de curtosis RGT para diferentes valores del parámetro λ [4]	11
5.1. Gráficas que ilustran la función de la distribución de Rayleigh generalizada y transmutada para algunas elecciones de parámetros α, β, λ [4]	15
6.1. Función de confiabilidad de la distribución RGT	29
7.1. Interpretación gráfica del método de inversión	31
7.2. Modelado de datos empíricos con GR y RGT	33
7.3. Comparación gráfica de ajuste	35

Capítulo 1

Justificación

La elaboración de modelos y en particular la modelación probabilística constituyen una herramienta fundamental en la praxis de la matemática aplicada al permitir unir la observación de hechos empíricos con elementos matemáticos aparentemente desligados, lo que ha permitido el surgimiento de ramas teórico-prácticas de la matemática como lo son la investigación de operaciones, los procesos estocásticos y la teoría de la información entre otros cuyo desarrollo se ha enfocado en mayor medida en aumentar la cercanía entre el modelo matemático y la realidad a modelar. La precisión y consecuente calidad de los procedimientos utilizados en un análisis estadístico depende en gran medida del modelo o las distribuciones de probabilidad que se emplean. Debido a esto, diversos autores han realizado un esfuerzo considerable en generalizar, extender o puntualizar distribuciones de probabilidad presentes en la literatura estadística. Es así que, Voda en [2] introduce la distribución de probabilidad Rayleigh generalizada; esta distribución es bastante utilizada en el análisis estadístico de confiabilidad, pero en algunos casos en particular se dificulta ajustar algunos datos a la distribución en cuestión, por lo que aquí se quiere plantear una distribución que conserve una estructura tipo Rayleigh generalizada pero con mayor o menor asimetría y curtosis (según sea el objetivo) y de esta manera proporcionar un mejor ajuste que la distribución de Rayleigh generalizada

Capítulo 2

Objetivos

2.1. Objetivo General

Estudiar la distribución Generalizada y transmutada de Rayleigh (RGT) a partir de la distribución generalizada de Rayleigh (GR) con mayor ajuste y vista en algunas aplicaciones.

2.2. Objetivos Específicos

1. Analizar algunas propiedades de la distribución encontrada (RGT) y su relación con otras distribuciones
2. Identificar los momentos de la distribución RGT
3. Establecer los parámetros para la distribución RGT por medio del método de máxima verosimilitud
4. Contrastar las informaciones obtenidas mediante la distribución RGT con datos obtenidos en ensayos reales.

Capítulo 3

Marco conceptual

A continuación se presentan definiciones y conceptos básicos de espacio medible y de medida tomados de [1] y [2], posteriormente se exhiben los conceptos de espacio probabilístico, variable aleatoria y función de probabilidad tomados de [1] y [3] para finalmente particularizar en la función de interés. Se presentan también algunos de los resultados básicos que habrán de ser útiles en el resto de este trabajo.

3.1. Funciones Aditivas de Conjunto

Definición 3.1.1 Sea $\mu : \mathcal{C} \rightarrow \overline{\mathbb{R}} \cup \{-\infty, \infty\}$ una función definida sobre una colección de conjuntos \mathcal{C} . Decimos que μ es finitamente aditiva si:

i) $\mu(\emptyset) = 0$

ii) Para toda colección finita $(A_n)_{1 \leq n \leq m}$ de conjuntos de \mathcal{C} , disjuntos dos a dos tales que $\bigcup_{i=1}^m A_i \in \mathcal{C}$ se tiene que

$$\mu \bigcup_{i=1}^m A_i = \sum_{n=1}^m \mu A_n \tag{3.1}$$

Observación 3.1.1 Todas las sumas en (1) deben estar bien definidas, de modo que no puede ocurrir que $\mu(A_i)$ se indetermina para algún i

Ejemplo 3.1.1 Sea Ω un conjunto cualquiera y sea $C = P(\Omega)$. Definimos $\mu(A)$ como el número de puntos de A . Si A tiene n elementos, $n = 0, 1, 2, \dots$, entonces $\mu(A) = n$; si A es un conjunto infinito, $\mu(A) = \infty$.

Teorema 3.1.1 Sea μ una función aditiva sobre el álgebra \mathcal{A}

- a) $\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B)$ para todo $A, B \in \mathcal{A}$
- b) Si $A, B \in \mathcal{F}$ y $B \subset A$, entonces $\mu(A) = \mu(B) + \mu(A - B)$, entonces, si $\mu(B)$ es finita, $\mu(A - B) = \mu(A) - \mu(B)$ y si $\mu(A - B) \geq 0$, $\mu(B) \leq \mu(A)$
- c) Si μ es no negativa, $\mu(\cup_{i=1}^n A_i) \leq \sum_{i=1}^n \mu(A_i)$.

Demostración 3.1.1 (a) Por aditividad tenemos

$$\begin{aligned}\mu(A) &= \mu(A \cap B) + \mu(A - B) \\ \mu(B) &= \mu(A \cap B) + \mu(B - A)\end{aligned}$$

Sumando las dos ecuaciones anteriores tenemos:

$$\begin{aligned}\mu(A) + \mu(B) &= \mu(A \cap B) + [\mu(A - B) + \mu(B - A) + \mu(A \cap B)] \\ &= \mu(A \cap B) + \mu(A \cup B)\end{aligned}$$

(b) Podemos escribir $A = B \cup (A - B)$, y entonces $\mu(A) = \mu(B) + \mu(A - B)$.

(c) Tenemos que

$$\bigcup_{i=1}^n A_i = A_1 \cup (A_1^c \cap A_2) \cup (A_1^c \cap A_2^c \cap A_3) \cup \dots \cup (A_1^c \cap \dots \cap A_{n-1}^c \cap A_n)$$

Los conjuntos que aparecen al lado derecho de esta relación son disjuntos y por lo tanto

$$\mu \left(\bigcup_{i=1}^n A_i \right) = \mu A_1 + \mu(A_1^c \cap A_2) + \mu(A_1^c \cap A_2^c \cap A_3) + \dots + \mu(A_1^c \cap \dots \cap A_{n-1}^c \cap A_n)$$

por otro lado tenemos que $A_1^c \cap \dots \cap A_{n-1}^c \cap A_n \subset A_n$ y como μ es no negativa entonces por (b) se deduce el resultado.

3.2. Espacio Medible

Definición 3.2.1 si \mathcal{F} es una σ -álgebra de subconjuntos del conjunto Ω , decimos que (Ω, \mathcal{F}) es un **espacio medible**. Una medida μ sobre \mathcal{F} es una función σ -aditiva que toma únicamente valores no negativos $\mu: \mathcal{F} \rightarrow [0, \infty]$

Ejemplo 3.2.1 (Medida de Dirac) Si (Ω, \mathcal{A}) es un espacio medible y $\omega \in \Omega$, la función de conjuntos ϵ_ω definida para $A \in \mathcal{A}$ por $\epsilon_\omega(A) = 1$ si $\omega \in A$ o 0 si no, es una probabilidad que llamaremos probabilidad de Dirac (o distribución de probabilidad degenerada) en el punto ω . En la figura 3.1 se muestra un diagrama con todos los subconjuntos posibles de un conjunto Ω de 3 puntos $\{x, y, z\}$. La medida de Dirac δ_x asigna un tamaño de 1 a todos los conjuntos en la mitad superior izquierda del diagrama y 0 a todos los conjuntos en la mitad inferior derecha.

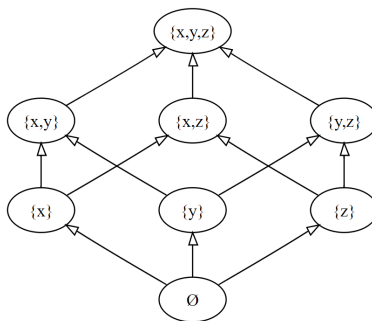


Figura 3.1: Caso particular medida de Dirac para conjunto de 3 elementos

3.3. Espacio probabilístico

Un espacio probabilístico es un caso particular de espacio de medida, donde todo conjunto medible tiene una medida o "tamaño" finito, dado por su probabilidad.

3.4. Variable aleatoria

Una variable aleatoria (v.a) X es una función real definida en el espacio de probabilidad (X, Σ, P) asociado a un experimento aleatorio.

3.4.1. Tipos de variables aleatorias

Para comprender de una manera más amplia y rigurosa los tipos de variables, es necesario conocer la definición de conjunto discreto. Un conjunto es discreto si está formado por un número finito de elementos, o si sus elementos se pueden enumerar en secuencia de modo que haya un primer elemento, un segundo elemento, un tercer elemento, y así sucesivamente (es decir, un conjunto infinito numerable sin puntos de acumulación). Las variables aleatorias reales se clasifican usualmente en v.a discretas y v.a continuas; Una v.a es discreta si su recorrido es un conjunto discreto.

3.4.2. Variable aleatoria continua

Una v.a. es continua si su recorrido es un conjunto no numerable. Intuitivamente esto significa que el conjunto de posibles valores de la variable abarca todo un intervalo de números reales. Por ejemplo, la estatura a una persona extraída de una determinada población es una variable continua en un intervalo de números reales. Las definiciones anteriores pueden generalizarse fácilmente a variables aleatorias con valores sobre \mathbb{R}^n o \mathbb{C}^n . Esto no agota el tipo de variables aleatorias ya que el valor de una variable aleatoria puede ser también una partición, como sucede en el proceso estocástico del restaurante chino o el conjunto de valores de una variable aleatoria puede ser un conjunto de funciones como el proceso estocástico de Dirichlet.

3.5. Función de distribución de probabilidad

La distribución de probabilidad (fda) de una variable aleatoria es una función que asigna a cada evento definido sobre la correspondiente variable aleatoria la probabilidad de que dicho evento ocurra. La distribución de probabilidad está definida sobre el conjunto de todos los eventos y cada uno de los eventos es el rango de la variable aleatoria. Se intuye que la variable aleatoria tiene una estrecha relación con las distribuciones de frecuencia. De hecho, se puede observar que una distribución de probabilidades sería una frecuencia teórica, ya que ésta última es aquella que describe cómo se modelan los posibles resultados.

La distribución de probabilidad está completamente determinada por la función de distribución, cuyo valor en cada x real es la probabilidad de que la variable aleatoria sea menor o igual que x .

Capítulo 4

Análisis de la distribución

4.1. Contexto

La distribución de Rayleigh debe su nombre al físico británico lord Rayleigh, también conocido como Barón Jhon William Strutt Rayleigh, premio nobel de física en el año 1904. En la teoría de la probabilidad y las estadísticas, la distribución de Rayleigh es una distribución de probabilidad continua para las variables aleatorias de valores positivos. Es una distribución χ en dos grados de libertad.

A menudo se observa una distribución de Rayleigh cuando la magnitud general de un vector está relacionada con sus componentes direccionales. Un ejemplo donde surge naturalmente la distribución de Rayleigh es cuando la velocidad del viento se analiza en dos dimensiones. Suponiendo que las componentes no estén correlacionadas, distribuidas normalmente con la misma varianza y media cero, entonces la velocidad global del viento (magnitud del vector) se caracterizará por una distribución de Rayleigh. Un segundo ejemplo de la distribución surge en el caso de números complejos aleatorios cuyos componentes reales e imaginarios tienen distribución normal de forma independiente e idéntica con igual varianza y media cero. En ese caso, la norma del número complejo se corresponde con la distribución de Rayleigh.

Definición 4.1.1 Una variable aleatoria X sigue una distribución de Rayleigh si su

función de probabilidad está dada por:

$$f(x, \sigma) = \frac{x}{\sigma^2} \exp\left(-\frac{x^2}{2\sigma^2}\right), \quad x > 0, \sigma > 0$$

Voda en [2] propone la distribución Rayleigh generalizada. Una variable aleatoria X sigue una distribución Rayleigh generalizada, denotada como $X \sim GR(\theta, \alpha)$, si su función de densidad está dada por

$$f_X(x; \theta, \alpha) = \frac{2\theta^{\alpha+1}}{\Gamma(\alpha+1)} x^{2\alpha+1} e^{-\theta x^2},$$

donde $x > 0$, $\theta > 0$ y $\alpha > -1$ y $\Gamma(\alpha) = \int_0^\infty u^{\alpha-1} e^{-u} du$ es la función Gamma. La función de distribución acumulada GR está dada por

$$F_X(x; \theta, \alpha) = \frac{\gamma(\alpha+1, x^2)}{\Gamma(\alpha+1)} x^{2\alpha+1},$$

donde $\gamma(\alpha, x) = \int_0^x u^{\alpha-1} e^{-u} du$ es la función Gamma incompleta. Para $\alpha = 0$ y $\theta = (2\lambda^2)^{-1}$

se tiene la distribución Rayleigh. Para $\alpha = \frac{1}{2}$ y $\theta = (2\lambda^2)^{-1}$ tenemos la distribución de Maxwell. Si $\alpha = -\frac{1}{2}$, $\theta = (2\sigma^2)^{-1}$ se tiene la distribución Half-Normal. Si $\alpha = \frac{v}{2} - 1$ y $\theta = \frac{1}{2}$ se tiene la distribución Chi cuadrado.

4.2. Rango de transmutaciones, asimetría y curtosis

Shaw y Buckley en [1] dan paso a una clase de distribuciones llamada la clase de distribuciones transmutadas. A continuación se definirán y tratarán algunos conceptos necesarios para puntualizar la idea de distribución transmutada.

En este trabajo se consideran funciones conformadas de la siguiente manera: Los mapas de transmutación de rango (MTR) $v = G[F^{-1}(u)]$, donde F y G son funciones de distribución acumulada (FDA). Dichas funciones han atraído la atención de los estadísticos en el pasado, en particular de Gilchrist [10], que se refiere a los MTR como transformaciones P . Para el presente trabajo, se usa el MTR como una herramienta para modular una distribución base dada con la finalidad de modificar sus momentos, en particular la asimetría y la curtosis. Un ejemplo importante para precisar los

conceptos será tomar como distribución base la normal, pero el foco del trabajo estará puesto en tomar como base la distribución Rayleigh generalizada (que de ahora en adelante se abreviará como RG).

Dadas dos distribuciones con un espacio muestral común con funciones de distribución acumulada F_1 y F_2 podemos definir un par de mapas de transmutación de rango (MTR) de la siguiente manera

$$\begin{aligned} G_{R_{12}}(u) &= F_2(F_1^{-1}(u)) , & . \\ G_{R_{21}}(u) &= F_1(F_2^{-1}(u)) & (2) \end{aligned}$$

Las funciones $G_{R_{12}}(u)$ y $G_{R_{21}}(u)$ llevan el intervalo unitario $I = [0, 1]$ en si mismo, y bajo hipótesis adecuadas son mutuamente inversas. De manera natural satisfacen $G_{R_{ij}}(0) = 0$ y $G_{R_{ij}}(1) = 1$. Para garantizar que las densidades transmutadas sean continuas, se agrega la hipótesis de que los MTR son continuamente derivables y, en general, se exige que las funciones sean monótonas. En lo que sigue la estructura será tomar F_1 como la distribución base, es decir aquella de la que se parte, y F_2 como la distribución modulada o aquella a la que se llega después de aplicar las correspondientes funciones .

4.2.1. Transmutación cuadrática

Un MTR en el que se llega a pensar naturalmente, llamado MTR cuadrático (MTRC), tiene la siguiente forma cuadrática simple, para $|\lambda| \leq 1$:

$$G_{R_{12}}(u) = u + \lambda u(1 - u),$$

de lo cual se sigue que las FDAS cumplen la relación

$$F_2(x) = (1 - \lambda)F_1(x) - \lambda F_1(x)^2$$

Puesto que el MTR inverso está disponible en forma cerrada, el algoritmo de muestreo sigue siendo manejable:

$$F_2^{-1} = F_1^{-1}(G_{R_{12}}(u)), G_{R_{12}}(u) = \frac{1 + \lambda - \sqrt{(\lambda + 1)^2 - 4\lambda u}}{2\lambda} \quad (3)$$

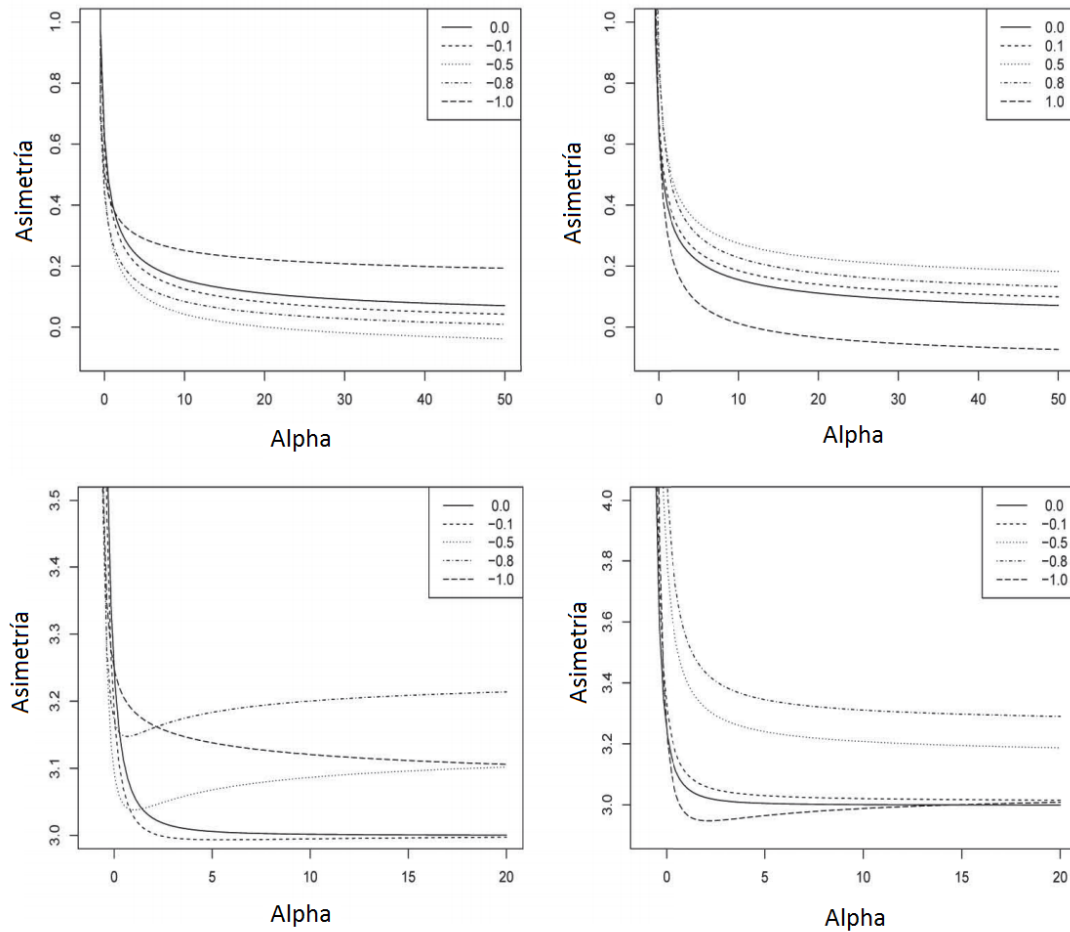


Figura 4.1: Gráfica del coeficiente de asimetría y de curtosis RGT para diferentes valores del parámetro λ [4]

El efecto del MTRC es introducir un sesgo a una distribución de base simétrica. No hay un requisito específico de que la distribución base F_1 sea simétrica. Sin embargo, si la distribución F_1 es simétrica sobre el origen, en el sentido de que $F_1(x) = 1 - F_1(-x)$, se tiene el resultado de que la distribución del cuadrado de la variable aleatoria transmutada es idéntico al de la distribución del cuadrado de la variable aleatoria original. Una consecuencia de esto es que si la distribución base es simétrica, entonces el MTRC conserva todos los momentos pares. La estructura dentro de los cuales la distribución del cuadrado se conserva bajo la transformación oblicua está bien documentada. Ver por ejemplo, Roberts y Gesser [9]. Se procede aplicando el MTRC a los casos en que la distribución base es uniforme, exponencial y normal.

Específicamente, una variable aleatoria X sigue una distribución transmutada si su función de distribución acumulada está dada por

$$F(x) = (1 + \lambda)G(x) - \lambda G^2(x), \quad |\lambda| < 1$$

donde $G(x)$ es la FDA de la distribución base. Para $\lambda = 0$ se obtiene la distribución de la variable aleatoria base. Aryal et al en [5] introducen la distribución Gumbel transmutada al considerar como base la función de distribución acumulada de la distribución Gumbel. Aryal et al. en [6] establecen la distribución Weibull transmutada al considerar como función base la FDA de la distribución Weibull de dos parámetros. Merovci en [7] introduce la distribución Rayleigh transmutada al considerar como función base la función de distribución acumulada de la distribución Rayleigh. Merovci en [8] introduce la distribución Rayleigh generalizada transmutada al considerar como función base la función de distribución acumulada de la distribución Burr tipo X de dos parámetros.

4.2.2. Asimetría

Los parámetros de asimetría son indicadores que permiten establecer la correspondencia de forma y posición respecto a una recta particular (en este caso) como eje de simetría se considera una recta paralela al eje de ordenadas que pasa por la media de la distribución sin hacer uso de su representación gráfica. Una distribución $F(x)$ es simétrica si dado un eje de simetría $x = \alpha$ entonces se tiene que $f(\alpha + \epsilon) = f(\alpha - \epsilon)$ para $\epsilon > 0$, por tanto, existe el mismo número de desviaciones con signo positivo que con signo negativo. Se dice que hay asimetría positiva (o a la derecha) si la cola.^a la derecha de la media es más larga que la de la izquierda, es decir, si $\int_{\alpha}^{\infty} f(x)dx > \int_{-\infty}^{\alpha} f(x)dx$. Se dice que hay asimetría negativa (o a la izquierda) si la cola.^a la izquierda de la media es más larga que la de la derecha, es decir, $\int_{\alpha}^{\infty} f(x)dx < \int_{-\infty}^{\alpha} f(x)dx$.

4.2.3. Curtosis

Un coeficiente de apuntamiento o de curtosis es el cuarto momento con respecto a la media definida como:

$$\beta_2 = \frac{\mu_4}{\sigma^4}$$

donde μ_4 es el cuarto momento centrado con respecto a la media y σ es la desviación estándar. En la distribución normal se verifica que $\mu_4 = 3\sigma^4$, donde μ_4 es el momento de orden 4 respecto a la media y σ la desviación típica. Por eso, está más extendida la siguiente definición del coeficiente de curtosis:

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

donde se resta 3 (debido a que es la curtosis de la distribución normal) con el objetivo de generar un coeficiente con valor 0 para la Normal y tomar a ésta como referencia de curtosis. El coeficiente de curtosis tiene una manifestación gráfica en la media de la distribución, dicho coeficiente de curtosis se clasifica en: leptocurtica si $\beta_2 > 3$ y $g_2 > 0$, platicurticas si $\beta_2 < 3$ y $g_2 < 0$ y mesocurticas si $\beta_2 = 3$ y $g_2 = 0$.

Según su concepción clásica, una mayor curtosis implica una mayor concentración de valores de la variable muy cerca de la media de la distribución (pico) y muy lejos de la misma (colas), al tiempo que existe una relativamente menor frecuencia de valores intermedios (hombros). Esto explica una forma de la distribución de frecuencias/probabilidad con colas más gruesas, con un centro más apuntado y una menor proporción de valores intermedios entre pico y colas.

Una mayor curtosis no implica una mayor varianza, ni viceversa.

Capítulo 5

Distribución Rayleigh generalizada transmutada

La distribución Burr tipo X de dos parámetros puede ser entendida como una generalización de la distribución Rayleigh clásica y, debido a esto, diversos autores prefieren utilizar el término de distribución Rayleigh generalizada para referirse a ella.

Aquí se utiliza el mapa de transmutación de rango cuadrático para introducir una extensión de la distribución Rayleigh generalizada mostrada anteriormente. Específicamente, se utiliza las expresiones anteriores para generar la función de distribución acumulada de la extensión. El resultado es una nueva distribución de 3 parámetros que flexibiliza el modelo Rayleigh generalizado en términos de asimetría y curtosis. Se presenta entonces la función de densidad y la función acumulada Rayleigh generalizada y transmutada, y se calculan sus momentos distribucionales, esperanza, varianza y coeficientes de asimetría y curtosis.

Sea X una variable aleatoria que sigue una distribución Rayleigh generalizada transmutada, denotada como $X \sim TGR(\theta, \alpha, \lambda)$. Entonces la función de densidad de X está dada por

$$f_X(x; \theta, \alpha, \lambda) = \frac{2\theta^{\alpha+1}}{\Gamma(\alpha+1)} x^{2\alpha+1} e^{-\alpha x^2} (1 + \lambda - 2\lambda H(\theta x^2; \alpha + 1, 1)) \quad (5)$$

y su función de distribución acumulada está dada por

$$F(x; \alpha, \theta, \lambda) = H(\theta x^2; \alpha + 1, 1)(1 + \lambda - \lambda H(\theta x^2; \alpha + 1, 1)) \quad (6)$$

donde $x > 0, \theta > 0, \alpha > -1, |\lambda| < 1$ y

$$H(x; \alpha, \beta) = \int_0^x \frac{\beta^\alpha}{\Gamma(\alpha)} u^{\alpha-1} e^{\beta u} du$$

A continuación (figura 5.1) se presentan algunas gráficas que ilustran la función de la distribución de Rayleigh generalizada y transmutada para algunas elecciones de parámetros α, β, λ

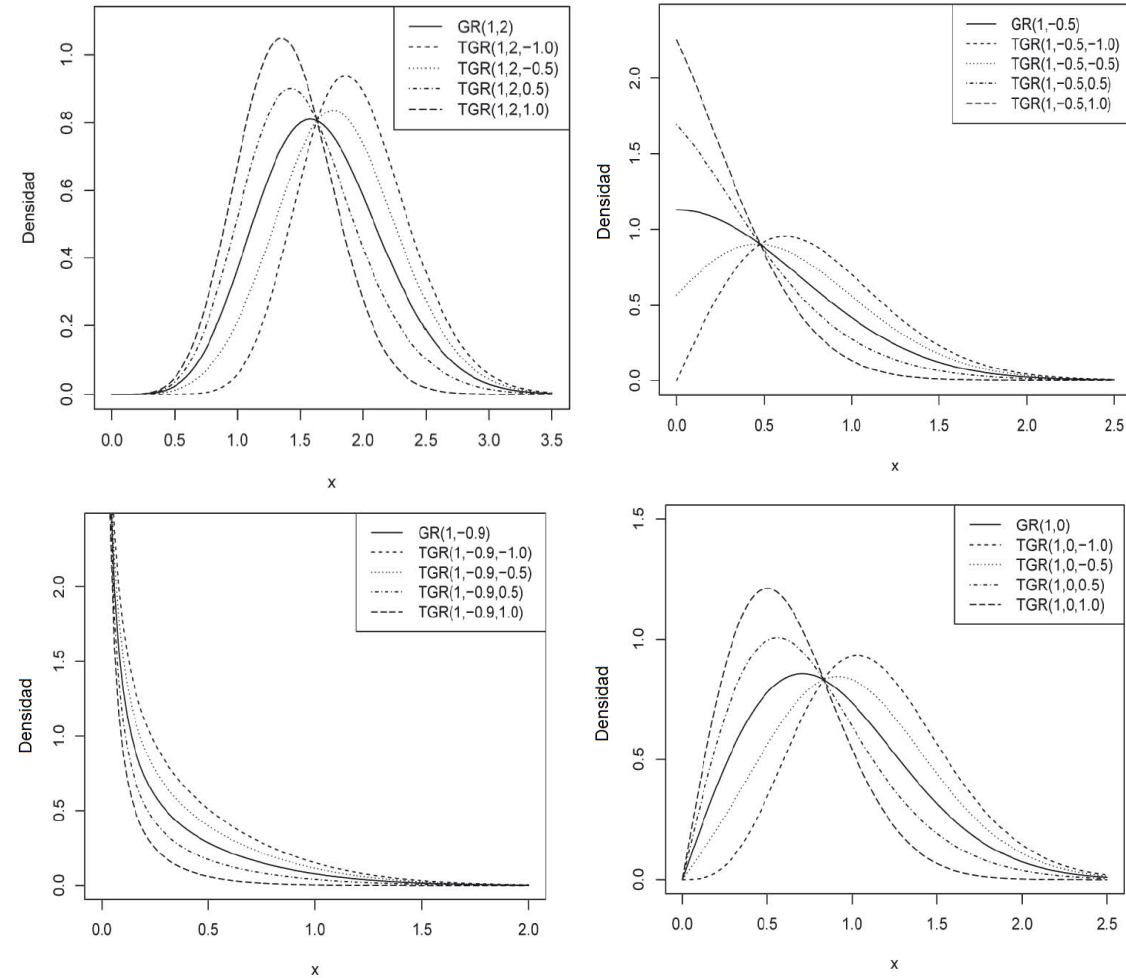


Figura 5.1: Gráficas que ilustran la función de la distribución de Rayleigh generalizada y transmutada para algunas elecciones de parámetros α, β, λ [4]

5.1. Algunas propiedades

Aquí se muestran algunas propiedades básicas de la distribución de Rayleigh generalizada transmutada.

Sea $X \sim RGT(\theta, \alpha, \lambda)$; entonces se tiene que:

1 Para $\lambda = 0$ la expresión (5) se reduce a

$$f_X(x; \theta, \alpha) = \frac{2\theta^{\alpha+1}}{\Gamma(\alpha + 1)} x^{2\alpha+1} e^{-\alpha x^2}$$

que es la distribución de Rayleigh generalizada.

2 Para $\theta = \frac{1}{2\sigma^2}$, $\alpha = 0$ y $\lambda = 0$, la expresión (5) se reduce a

$$f_X(x; \sigma^2) = \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}}$$

que es la función de densidad de la distribución Rayleigh [8]

3 Para $\theta = \frac{1}{2\sigma^2}$, $\alpha = \frac{-1}{2}$ y $\lambda = 0$, la expresión (5) se reduce a

$$f_X(x; \sigma) = \frac{\sqrt{2}}{\sigma\sqrt{2}} \exp\left(-\frac{x^2}{2\sigma^2}\right); \quad x > 0$$

que es la función de densidad de la distribución Half-Normal [9]

4 Para $\theta = \frac{\sigma}{2}$, $\alpha = \frac{1}{2}$ y $\lambda = 0$, la expresión (5) se reduce a

$$f_X(x; \sigma) = \sqrt{\frac{2}{\pi}} \sigma^{\frac{3}{2}} x^2 e^{-\frac{\sigma}{2} x^2},$$

que es la función de densidad de la distribución de Maxwell [10]

5 Para $\theta = \frac{1}{2}$, $\alpha = \frac{\sigma}{2} - 1$ y $\lambda = 0$, la expresión (5) se reduce a

$$f_X(x; \sigma) = \frac{1}{2^{\sigma/2-1} \Gamma(\sigma/2)} x^{\sigma-1} e^{-\frac{x^2}{2}},$$

que es la función de densidad de la distribución Chi [13]

5.2. Distribuciones relacionadas

Los siguientes corolarios son consecuencias directas de (5), y se obtienen como extensiones de los casos particulares derivados de la distribución Rayleigh generalizada. En los corolarios 6.4, 2.5 y 2.6, $\Gamma(\alpha) = \int_{x=0}^{\infty} u^{\alpha-1} e^{-u} du$ corresponde a la función Gamma y $F(x, \alpha, \beta) = \int_0^x \frac{\beta^\alpha}{\Gamma(\alpha)} u^{\alpha-1} e^{-\beta u} du$ es la función de la distribución acumulada Gamma.

Corolario 5.2.1 Sea $X \sim RGT(\frac{1}{2\sigma^2}, 0, \lambda)$ entonces X se distribuye de acuerdo con la distribución de Rayleigh transmutada, denotada por $X \sim RT(\sigma, \lambda)$. La función de densidad de X está dada por

$$f(x, \sigma, \lambda) = \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}} \left(1 - \lambda + 2\lambda e^{-\frac{x^2}{2\sigma^2}} \right),$$

donde $x > 0, \sigma > 0, |\lambda| \leq 1$.

Corolario 5.2.2 Sea $T \sim RGT(\frac{1}{2\sigma^2}, -\frac{1}{2}, \lambda)$. Entonces X se distribuye de acuerdo con la distribución Half-normal transmutada, denotada por $X \sim HNT(\sigma, \lambda)$. La función de densidad de X está dada por

$$f(x, \sigma, \lambda) = \sqrt{\frac{2}{\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \left(1 + \lambda - 2\lambda H\left(\frac{x^2}{2\sigma^2}, \frac{1}{2}, 1\right) \right)$$

donde $x > 0, \sigma > 0, |\lambda| \leq 1$.

Corolario 5.2.3 Sea $X \sim RGT(\frac{\sigma}{2}, \frac{1}{2}, q)$. Entonces X se distribuye de acuerdo con la distribución Maxwell transmutada, denotada por $X \sim MT(\sigma, \lambda)$. La función de densidad de X está dada por

$$f(x, \sigma, \lambda) = \sqrt{\frac{2}{\pi}} \sigma^{2/3} x^2 e^{-\frac{\sigma}{2} x^2} \left(1 + \lambda - 2\lambda H\left(\frac{\sigma}{2} x^2, \frac{3}{2}, 1\right) \right),$$

donde $x > 0, \sigma > 0, |\lambda| \leq 1$.

Corolario 5.2.4 Sea $X \sim RGT(\frac{1}{2}, \frac{\sigma}{2}, 1, \lambda)$. Entonces X se distribuye de acuerdo con la distribución Chi transmutada, denotada por $X \sim T_{\chi_{\sigma,q}}$. La función de densidad de X está dada por

$$f(x, \sigma, \lambda) = \frac{1}{2^{\frac{\sigma}{2}-1} \Gamma\left(\frac{\sigma}{2}\right)} x^{\sigma-1} e^{-\frac{x^2}{2}} \left(1 + \lambda - 2\lambda H\left(\frac{x^2}{2}, \frac{\sigma}{2}, 1\right) \right),$$

donde $x > 0, \sigma > 0, |\lambda| \leq 1$.

5.3. Momentos

Proposición 5.3.1 Sea $X \sim RGT(\theta, \alpha, \lambda)$. Entonces, para $r = 1, 2, \dots$ se tiene que el r -ésimo momento es

$$\mu_r = E(X^r) = \frac{\theta^{-\frac{r}{2}}}{\Gamma(\alpha + 1)} \left[(1 + \lambda)\Gamma\left(\frac{r}{2} + \alpha + 1\right) - 2\lambda I(r, \alpha) \right]$$

donde $\Gamma = \int_0^\infty u^{\alpha-1} e^{-u} du$ es la función Gamma, mientras que

$$I(r, \alpha) = \int_0^\infty u^{\frac{r}{2} + \alpha} e^{-u} H(u, \alpha + 1, 1) du$$

es una función que depende de r y α con

$$H(u, \alpha + 1, 1) = \int_0^u \frac{1}{\Gamma(\alpha + 1)} y^\alpha e^{-y} dy$$

Demostración 5.3.1 Por definición se tiene que los momentos de la distribución Rayleigh generalizada transmutada están dados por

$$\mu_r = \frac{2\theta^{\alpha+1}}{\Gamma(\alpha + 1)} \int_0^\infty x^{r+2\alpha+1} e^{-\theta x^2} [1 + \lambda - 2\lambda H(\theta x^2, \alpha + 1, 1)] dx$$

considerando el cambio de variable $u = \theta x^2$ se obtiene el resultado.

Q.E.D.

Observación 5.3.1 sea $X \sim RGT(\theta, \alpha, \lambda)$ entonces que tiene que:

- $\mu_1 = E(X) = \frac{\theta^{-1/2}}{\Gamma(\alpha + 1)} a_1$, donde $a_1 = (1 + \lambda)\Gamma\left(\frac{3}{2} + \alpha\right) - 2\lambda I(1, \alpha)$,
- $\mu_2 = E(X^2) = \frac{\theta^{-1}}{\Gamma(\alpha + 1)} a_2$, donde $a_2 = (1 + \lambda)\Gamma(2 + \alpha) - 2\lambda I(2, \alpha)$,
- $\mu_3 = E(X^3) = \frac{\theta^{-3/2}}{\Gamma(\alpha + 1)} a_3$, donde $a_3 = (1 + \lambda)\Gamma\left(\frac{5}{2} + \alpha\right) - 2\lambda I(3, \alpha)$,
- $\mu_4 = E(X^4) = \frac{\theta^{-2}}{\Gamma(\alpha + 1)} a_4$, donde $a_4 = (1 + \lambda)\Gamma(3 + \alpha) - 2\lambda I(4, \alpha)$,

Corolario 5.3.1 Sea $X \sim RGT(\theta, \alpha, \lambda)$ entonces que tiene:

$$E(X) = \frac{\theta^{-1/2}}{\Gamma(\alpha + 1)} a_1 \quad y \quad Var(X) = \frac{\theta^{-1}}{\Gamma(\alpha + 1)} \left(a_2 - \frac{a_1^2}{\Gamma(\alpha + 1)} \right)$$

Corolario 5.3.2 Sea $X \sim RGT(\theta, \alpha, \lambda)$. Entonces los coeficientes de asimetría ($\sqrt{\beta_1}$) y curtosis (β_2) son respectivamente

$$\sqrt{\beta_1} = \frac{\Gamma^{1/2}(\alpha + 1) \left(a_3 - \frac{3a_1a_2}{\Gamma(\alpha + 1)} + \frac{2a_1^3}{\Gamma^2(\alpha + 1)} \right)}{\left(a_2 - \frac{a_1^2}{\Gamma(\alpha + 1)} \right)^{3/2}}$$

y

$$\beta_2 = \frac{\Gamma(\alpha + 1) \left(a_4 - \frac{4a_1a_3}{\Gamma(\alpha + 1)} + \frac{6a_2a_1^2}{\Gamma^2(\alpha + 1)} - \frac{3a_1^4}{\Gamma^3(\alpha + 1)} \right)}{\left(a_2 - \frac{a_1^2}{\Gamma(\alpha + 1)} \right)^{3/2}}$$

Observación 5.3.2 Cuando $\lambda = 0$ los coeficientes $\sqrt{\beta_1}$ y β_2 de la distribución RGT toman respectivamente los valores

$$\sqrt{\beta_{1RG}} = \frac{\Gamma^2(\alpha + 1)\Gamma(\alpha + \frac{5}{2}) - 3\Gamma(\alpha + \frac{3}{2})\Gamma(\alpha + 2)\Gamma(\alpha + 1) + 2\Gamma^3(\alpha + \frac{3}{2})}{\left[\Gamma(\alpha + 1)\Gamma(\alpha + 2) - \Gamma^2(\alpha + \frac{3}{2}) \right]^{3/2}}$$

y

$$\beta_{2RG} = \frac{\Gamma^3(\alpha + 1)\Gamma(\alpha + 3) - 4\Gamma(\alpha + \frac{3}{2})\Gamma(\alpha + \frac{5}{2})\Gamma^2(\alpha + 1)}{\left[\Gamma(\alpha + 1)\Gamma(\alpha + 2) - \Gamma^2(\alpha + \frac{3}{2}) \right]^2} + \frac{6\Gamma^2(\alpha + \frac{3}{2})\Gamma(\alpha + 1)\Gamma(\alpha + 2) - 3\Gamma^4(\alpha + \frac{3}{2})}{\left[\Gamma(\alpha + 1)\Gamma(\alpha + 2) - \Gamma^2(\alpha + \frac{3}{2}) \right]^2}$$

que son los respectivos coeficientes de asimetría y curtosis de la distribución Rayleigh generalizada. La Figura 4.1 muestra las gráficas de los coeficientes de asimetría y curtosis de la distribución Rayleigh generalizada transmutada para diferentes valores de λ .

Capítulo 6

Inferencia

En este apartado se presentan los estimadores de máxima verosimilitud para los parámetros θ , α y λ de la distribución Rayleigh generalizada transmutada, se realiza la matriz de información observada, además de un estudio de simulación y una aplicación con datos reales.

6.1. Estimadores de máxima verosimilitud

Supóngase una urna con 4 bolas que pueden ser rojas o azules. La proporción, θ , de bolas rojas en la urna es indeterminada. Entonces el objetivo es estimar el valor de θ . Para ello, se extraen de la urna 2 bolas con reemplazo. Suponiendo que la primera bola extraída es roja (R) y la segunda es Azul (A).

Se buscará el valor (posible) de θ que haga más verosímil (que se acerque mas a la realidad) el resultado obtenido en el experimento.

Para ello se calcula

$$P(R, A|\theta)$$

y se elegirá el valor de θ que de mayor probabilidad; Es decir, se busca el valor de θ que maximiza la función de probabilidad conjunta $P(R, A|\theta)$, también llamada **verosimilitud**.

La estimación obtenida se conoce como **estimación de máxima verosimilitud**

(EMV): Se usa la información disponible en la muestra para elegir el valor del parámetro para el cuál es más probable haber observado ese resultado muestral.

6.1.1. Verosimilitud

Definición 6.1.1 Sea X_1, \dots, X_n una muestra aleatoria de una población X con función de probabilidad P_θ . Para cada muestra particular (x_1, \dots, x_n) , la función verosimilitud se define como la función de probabilidad conjunta de (X_1, \dots, X_n) evaluada en (x_1, \dots, x_n)

$$\begin{aligned}\mathcal{L}(\theta) &= \mathcal{L}(x_1, \dots, x_n; \theta) = P_\theta(X_1 = x_1, \dots, X_n = x_n) \text{ si } X \text{ es discreta} \\ \mathcal{L}(\theta) &= \mathcal{L}(x_1, \dots, x_n; \theta) = f_\theta(x_1, \dots, x_n) \text{ si } X \text{ es continua.}\end{aligned}$$

Observación 6.1.1 La notación $\mathcal{L}(\theta)$ indica que \mathcal{L} es una función de θ y no de (x_1, \dots, x_n) .

Observación 6.1.2 θ puede ser un escalar o un vector ($\theta = (\theta_1, \dots, \theta_n)$).

Observación 6.1.3 El subíndice θ en la función de probabilidad o de densidad indica que dicha función depende del valor del parámetro.

6.1.2. Estimador de máxima verosimilitud

Definición 6.1.2 Sea X_1, \dots, X_n una muestra aleatoria de una población X con función de verosimilitud $\mathcal{L}(\theta)$. Para cada muestra particular (x_1, \dots, x_n) , la estimación de máxima verosimilitud de θ es el valor $\hat{\theta}_{MV}$ que maximiza la verosimilitud. Es decir:

$$\mathcal{L}(x_1, \dots, x_n; \hat{\theta}_{MV}) = \max_{\theta} \mathcal{L}(x_1, \dots, x_n; \theta)$$

El estimador de máxima verosimilitud, $\hat{\theta}_{MV}(X_1, \dots, X_n)$, es aquel que evaluado en cada muestra particular nos da la estimación de máxima verosimilitud ($\hat{\theta}_{MV}(x_1, \dots, x_n)$).

Observación 6.1.4 Muchas veces se utiliza el término estimador de máxima verosimilitud para denotar tanto el estimador como cualquier estimación particular.

Observación 6.1.5 Ambos se denotan por $\hat{\theta}_{MV}$ y se abrevian por E.M.V.

6.1.3. Método para obtener el E.M.V

El método de la máxima verosimilitud puede emplearse en situaciones donde existen varios parámetros desconocidos $\theta_1, \theta_2, \dots, \theta_k$ que es necesario estimar. En tales casos, la función de verosimilitud es una función multivariada en los parámetros. En tales casos, la función de verosimilitud es una función multivariada en los parámetros desconocidos $\theta_1, \theta_2, \dots, \theta_k$ y los estimadores de máxima verosimilitud $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_k$ se pueden obtener de la siguiente manera:

Si $\theta = (\theta_1, \dots, \theta_k)$, para obtener el E.M.V. de θ_j dada una muestra particular (x_1, \dots, x_n) se utiliza el siguiente procedimiento:

- Escribir la Verosimilitud : $\mathcal{L}(\theta) = \mathcal{L}(x_1, \dots, x_n; \theta)$
- Escribir el logaritmo de la verosimilitud : $\mathcal{L}\mathcal{L}(\theta) = \ln\mathcal{L}(\theta)$, $\mathcal{L}\mathcal{L}(\theta)$ se llama log-verosimilitud.
- Obtener el θ_j tal que

$$\frac{\partial}{\partial \theta_j} \mathcal{L}\mathcal{L}(\theta) = 0$$

y denotarlo por $\hat{\theta}_j$

Ejemplo 6.1.1 Continuando con el ejemplo con el que se abre la subsección se escribe la función de verosimilitud en el caso de extraer dos bolas de la urna con reemplazamiento.

Al extraer una bola al azar

$$X = \begin{cases} 1 & \text{si es roja} \\ 0 & \text{si es azul} \end{cases}$$

Se tiene entonces un muestreo aleatorio simple (m.a.s) de tamaño 2, es decir, $X_1, X_2 \sim$

$B(\theta)$ Para una muestra particular cualquiera (x_1, x_2) :

$$\begin{aligned} \mathcal{L}(x_1, x_2; \theta) &= P(X_1 = x_1, X_2 = x_2; \theta) \\ &= P(X_1 = x_1; \theta)P(X_2 = x_2; \theta) \\ &= [\theta^{x_1}(1 - \theta)^{1-x_1}] [\theta^{x_2}(1 - \theta)^{1-x_2}] \end{aligned}$$

por tanto, derivando la función de verosimilitud se tiene:

$$\frac{\partial \mathcal{L}}{\partial \theta} = \frac{x_1 + x_2 - 2\theta}{\theta(1 - \theta)}$$

así el EMV correspondiente es:

$$\hat{\theta} = \frac{x_1 + x_2}{2}$$

Ejemplo 6.1.2 Sea X_1, \dots, X_n un muestreo aleatorio de $X \sim N(\mu, \sigma)$. Obtener el E.M.V de μ .

La función de verosimilitud es:

$$\begin{aligned} \mathcal{L} &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(x_1 - \mu)^2}{2\sigma^2}\right) \cdots \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(x_n - \mu)^2}{2\sigma^2}\right) \\ &= \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \exp\left(\frac{\sum_{i=1}^n (x_i - \mu)^2}{-2\sigma^2}\right) \end{aligned} \quad (6.1)$$

entonces se tiene que la función de log-verosimilitud es:

$$\begin{aligned} \mathcal{L}\mathcal{L} &= \log(\mathcal{L}) \\ &= -\frac{n}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \end{aligned} \quad (6.2)$$

con el objetivo de encontrar los valores críticos de μ y σ se toman las derivadas parciales de $\mathcal{L}\mathcal{L}$ con respecto a μ y σ^2 , se igualan a cero y se resuelven las ecuaciones resultantes

$$\begin{aligned} \frac{\partial \mathcal{L}\mathcal{L}}{\partial \mu} &= \frac{\partial}{\partial \mu} \left(-\frac{n}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \right) \\ &= 0 - \left(-2 \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu) \right) \\ &= \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \mu) \end{aligned} \quad (6.3)$$

por otro lado se tiene:

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \sigma^2} &= \frac{\partial}{\partial \sigma^2} \left(-\frac{n}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \right) \\ &= \frac{1}{2\sigma^4} \sum_{i=1}^n (x_i - \mu)^2 - \frac{n}{2\sigma^2}\end{aligned}\tag{6.4}$$

igualando (4) y (5) a cero, finalmente se tiene:

$$\begin{aligned}\hat{\mu} &= \frac{1}{n} \sum_{i=1}^n x_i \\ \hat{\sigma}^2 &= \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2\end{aligned}$$

6.1.4. E.M.V para la distribución RGT

Dada una variable aleatoria X_1, \dots, X_n que con distribución $RGT(\theta, \alpha, \lambda)$ se calcula la función de verosimilitud así:

$$\begin{aligned}\mathcal{L} &= 2\alpha\beta^2 x_1 e^{-(\beta x_1)^2} (1 - e^{-(\beta x_1)^2})^{\alpha-1} (1 + \lambda - 2\lambda - 2\lambda(1 - e^{-(\beta x_1)^2})^\alpha) \bullet \dots \bullet \\ &\quad 2\alpha\beta^2 x_n e^{-(\beta x_n)^2} (1 - e^{-(\beta x_n)^2})^{\alpha-1} (1 + \lambda - 2\lambda - 2\lambda(1 - e^{-(\beta x_n)^2})^\alpha) \\ &= (2\alpha\beta^2)^n \exp\left(-\beta^2 \sum_{i=1}^n x_i^2\right) \prod_{i=1}^n x_i (1 - \exp(-(\beta x_i)^2))^{\alpha-1} [1 + \lambda - 2\lambda(1 - \exp(-\beta x_i)^2)^\alpha]\end{aligned}$$

así la función de log-verosimilitud es:

$$\begin{aligned}\mathcal{L}\mathcal{L} = \ln(\mathcal{L}) &= n(\ln(2) + \ln(\alpha) + 2\ln(\beta)) - \beta^2 \sum_{i=1}^n x_i^2 \\ &\quad + \sum_{i=1}^n \ln(x_i) + (\alpha - 1) \sum_{i=1}^n \ln(1 - e^{(\beta x_i)^2}) \\ &\quad + \sum_{i=1}^n \ln \left[1 + \lambda - 2\lambda(1 - e^{-(\beta x_i)^2})^\alpha \right]\end{aligned}\tag{6.5}$$

Las ecuaciones de máxima verosimilitud están dadas por

Observación 6.1.6 En adelante se usará la notación f_x, f_y para referirse a las primeras derivadas parciales y f_{xx}, f_{yy}, f_{xy} y f_{yx} para las segundas derivadas parciales

Ahora, haciendo

$$\mathcal{L}\mathcal{L}_\alpha = 0, \quad \mathcal{L}\mathcal{L}_\beta = 0, \quad \mathcal{L}\mathcal{L}_\lambda = 0$$

entonces se tiene:

$$\frac{n(\alpha + 1)}{\theta} - \sum_{i=1}^n x_i^2 - 2\lambda \sum_{i=1}^n \frac{H_1(x_i)}{1 + \lambda - 2\lambda H(x_i)} = 0 \quad (6.6)$$

$$n \log(\theta) - n\Psi(\alpha + 1) + 2 \sum_{i=1}^n \log(x_i) - 2\lambda \sum_{i=1}^n \frac{H_2(x_i)}{1 + \lambda - 2\lambda H(x_i)} = 0 \quad (6.7)$$

$$\sum_{i=1}^n \frac{1 - 2H(x_i)}{1 + \lambda - 2\lambda H(x_i)} = 0 \quad (6.8)$$

donde $H(x_i) = H(\theta x_i^2; \alpha + 1, 1)$, $H_1(x_i) = \frac{d}{d\theta} H(x_i)$, $H_2(x_i) = \frac{d}{d\alpha} H(x_i)$. Sin embargo las ecuaciones (6.6), (6.7), (6.8) no conducen a soluciones analíticas explícitas para las estimaciones de los parámetros del modelo. Por lo general, es más conveniente utilizar algoritmos de optimización no lineales, tales como el algoritmo de cuasi-Newton, para maximizar numéricamente la función de log-verosimilitud dada en (6.5). Por ejemplo, el software R cuenta con una librería de optimización no lineal (optim) para resolver este tipo de problemas.

6.2. Matriz de información observada

Aquí se presenta la matriz de información observada sobre las ecuaciones de máxima verosimilitud (jacobiano) que consiste en la matriz de segundas derivadas parciales con respecto a cada uno de los parámetros en cuestión. Si $X \sim RGT(\theta, \alpha, \lambda)$;

entonces la matriz de información observada está dada por:[4]

$$\begin{pmatrix} -\frac{n(\alpha+1)}{\theta^2} - 2\lambda G_{\theta,1}(x_i) & \frac{n}{\theta} - 2\lambda G_{\alpha,1}(x_i) & -2 \sum_{i=1}^n \frac{H_i(x_i)}{1+\lambda-2\lambda H(x_i)} - 2\lambda G_{\lambda,1}(x_i) \\ \frac{n}{\theta} - 2\lambda G_{\alpha,2}(x_i) & -n\psi_1(\alpha+1) - 2\lambda G_{\alpha,2}(x_i) & -2 \sum_{i=1}^n \frac{H_i(x_i)}{1+\lambda-2\lambda H(x_i)} - 2\lambda G_{\lambda,2}(x_i) \\ \sum_{i=1}^n \frac{d}{d\theta} \frac{1-2H(x_i)}{1+\lambda-2\lambda H(x_i)} & \sum_{i=1}^n \frac{d}{d\alpha} \frac{1-2H(x_i)}{1+\lambda-2\lambda H(x_i)} & - \sum_{i=1}^n \left(\frac{1-2H(x_i)}{1+\lambda-2\lambda H(x_i)} \right)^2 \end{pmatrix}$$

Donde

$$G_{v,k}(x_i) = \sum_{i=1}^n \frac{d}{dv} \left(\frac{H_k(x_i)}{1+\lambda-2\lambda H(x_i)} \right),$$

$v = \theta, \alpha, \lambda$ y ψ_1 es la función trigamma, con $H_k(x_i)$, $k = 1, 2, \dots$ dadas anteriormente en las ecuaciones de máxima verosimilitud

6.3. Análisis de confiabilidad

La hipótesis fundamental del presente trabajo es que la distribución RGT es un mejor modelo para distintos procesos, fenómenos y eventos que las distribuciones comúnmente usados para tal fin como la Raileygh, la RG y la RT. En la presente sección se hace uso de un análisis de confiabilidad para sustentar dicha hipótesis.

6.3.1. Pruebas de bondad de ajuste

La bondad de ajuste (GOF por sus siglas en inglés) de un modelo estadístico representa qué tan bien se ajusta a un conjunto de observaciones. Los índices GOF resumen la diferencia entre los valores observados y los valores esperados bajo un modelo estadístico. Los estadísticos de GOF son indicadores de bondad de ajuste con distribuciones de muestreo conocidas, usualmente se obtienen mediante métodos asintóticos, que se utilizan en la prueba de hipótesis estadística. En el presente trabajo se hará uso de varios criterios de bondad de ajuste para validar el uso de la distribución RGT sobre la distribución GR.

Para comparar las distribuciones, se consideran algunos criterios como lo son K-S

(Kolmogorov Smirnov), AIC (Akaike Information Criterion), AICC (Akaike Information Criterion Corrected) y BIC (Bayesian information criterion) para el conjunto de datos reales; A continuación se enuncian las ecuaciones de los referidos criterios.

La prueba de Kolmogorov-Smirnov es un estadístico que prueba la máxima diferencia entre la muestra extraída de la población y la distribución teórica que se le ha estimado, dicha distancia se define por:

$$D = \max |F_n(x) - F_0(x)|$$

donde $F_n(x)$ es la función de distribución muestral y $F_0(x)$ es la función teórica especificada en la hipótesis nula.

En el caso de AIC se tiene la ecuación

$$AIC = 2k - 2 \ln(L)$$

Donde k es el número de parámetros en el modelo estadístico, y L es el máximo valor de la función de verosimilitud para el modelo estimado.

Cuando el tamaño de la muestra es pequeño, existe una gran probabilidad de que AIC seleccione modelos que tengan demasiados parámetros, es decir, que AIC se adapte en exceso, generando problemas de cálculo. Para abordar este posible exceso, el AICC, (AICC es AIC con una corrección para tamaños de muestra pequeños) provee un buen ajuste con una mínima cantidad de parámetros.

La Ecuación para AICC depende del modelo estadístico. Suponiendo que el modelo es univariado, es lineal en sus parámetros y tiene residuos distribuidos normalmente (condicionales a los regresores), entonces la fórmula para AICC es la siguiente

$$AICC = AIC + \frac{2k^2 + 2k}{nk - 1} \quad AICC = AIC + \frac{2k^2 + 2k}{nk - 1}$$

El BIC fue desarrollado por Gideon E. Schwarz y publicado en un artículo de 1978, [14] donde presentó un argumento bayesiano para adoptarlo. El BIC se define formalmente

como

$$BIC = 2LL + k \log(n)$$

donde LL es la probabilidad de registro del modelo , k es el número de parámetros independientes y n es el tamaño de la muestra.

6.3.2. Función de confiabilidad y tasa de riesgo

Aunque la fdp $f(t)$ describe el tiempo hasta que un sistema falla completamente (en el caso que así esté definida) no indica directamente la posibilidad de que parte del sistema o el sistema continúe trabajando por un determinado lapso o como la probabilidad de que falle depende del tiempo de funcionamiento de la parte del sistema en cuestión, por ello se define la función de confiabilidad.

Definición 6.3.1 Sea $f(t)$ una fdp, la **función de confiabilidad** notada por $R(t)$ está dada por.

$$\begin{aligned} R(t) &= P(T > t) \\ &= \int_t^{\infty} f(x) dx \\ &= 1 - F(t) \end{aligned}$$

donde $F(t)$ es la fda y $1 - F(t)$ es probabilidad de que el sistema funcione hasta el momento t

Así definida la función de confiabilidad y realizando un cambio de parámetros para facilitar los cálculos se tiene que en el caso particular de la distribución RGT la función de confiabilidad está dada por:

$$R(t) = 1 - H(\theta t^2, \alpha + 1, 1)(1 + \lambda - \lambda H(\theta t^2, \alpha + 1, 1))$$

La otra característica de interés de una variable aleatoria es la función de tasa de riesgo definida por

$$h(t) = \frac{f(t)}{1 - F(t)}$$

Que informalmente hablando es una variable que caracteriza el fenómeno de la vida útil de un elemento. Puede interpretarse libremente como la probabilidad condicional de fracaso, dado que ha sobrevivido al tiempo t . La función de riesgo para una variable aleatoria de Rayleigh generalizada transmutada es dado por

$$h(x) = \frac{(2\theta^{\alpha+1}x^{2\alpha+1}e^{-\theta x^2})(1 + \lambda - 2\lambda H(\theta x^2, \alpha + 1, 1))}{\Gamma(\alpha + 1)(1 - H(\theta x^2, \alpha + 1, 1))(1 + \lambda - \lambda H(\theta x^2, \alpha + 1, 1))}$$

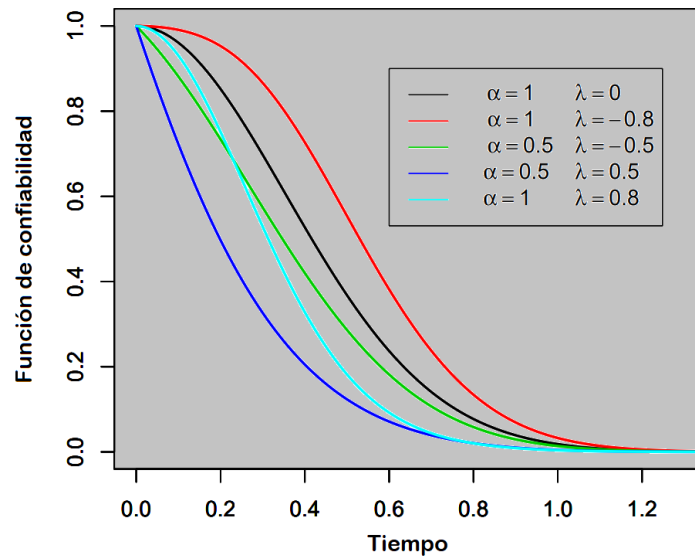


Figura 6.1: Función de confiabilidad de la distribución RGT

Capítulo 7

Estimación de parámetros

Con la finalidad de realizar un estudio de simulación para ilustrar el comportamiento de los estimadores de máxima verosimilitud se hace uso del software R para generar muestras de números pseudoaleatorios distribuidos RGT. Los números aleatorios son aquellos que pueden ser generados a partir de fuentes de aleatoriedad las cuales generalmente son de naturaleza física (dados, ruletas, mecanismos eléctricos o mecánicos), y son gobernados por las leyes del azar, éstos exhiben verdadera aleatoriedad en la realización de experimentos, los números pseudo aleatorios son aquellos que tienen un comportamiento similar a la naturaleza aleatoria, pero están ceñidos a un patrón, generalmente de naturaleza matemática que hace que su comportamiento sea determinístico; en el actual trabajo se ejemplificara sobre la utilización de algunos métodos de generación de números pseudoaleatorios.

7.0.1. Método de inversión de la función distribución

Para generar valores de X a partir de números pseudoaleatorios, utilizando el método de inversión, se divide el intervalo $(0, 1)$ en tantas partes como valores tome la variable X , de modo que el i -ésimo intervalo tenga probabilidad p_i .

Se asigna a X el valor x_i si el valor u generado de la distribución $\mathcal{U}(0, 1)$ verifica:

$$\sum_{k=1}^{i-1} p_k < u \leq \sum_{k=1}^i p_k,$$

es decir, si $F_X(x_{i-1}) < u < F_X(x_i)$. Que este método genera efectivamente valores que siguen la distribución de X se sigue de la expresión siguiente

$$P(X = x_i) = P\left(\sum_{k=1}^{i-1} p_k < U \leq \sum_{k=1}^i p_k\right) = \sum_{k=1}^i p_k - \sum_{k=1}^{i-1} p_k = p_i;$$

Obsérvese que una vez generado el número aleatorio u , encontrar el valor generado de X consiste en encontrar el intervalo $(F_X(x_{i-1}), F_X(x_i))$ al que pertenece u lo que equivale a encontrar la inversa de F_X . La interpretación gráfica sería la siguiente:

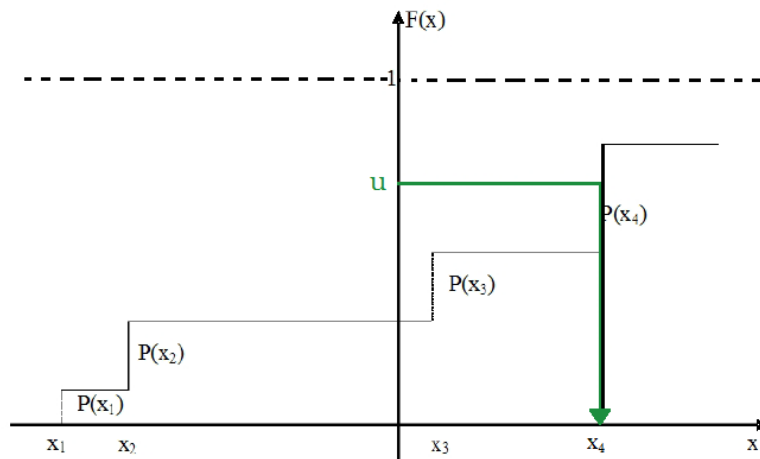


Figura 7.1: Interpretación gráfica del método de inversión

Con la finalidad de ejemplificar el comportamiento de los parámetros y sus respectivas estimaciones bajo el modelo RGT se hace uso de la librería *optimx* del software *r* se optimiza la log-verosimilitud .

Con el método antes descrito, se realiza una simulación generando 500 números pseudo aleatorios con muestras de $n = 50$, $n = 150$ y $n = 300$ que siguen la distribución $RGT(\theta, \alpha, \lambda)$ para valores fijos de los parámetros, en la tabla anexa se presentan las medias y desviaciones estándar resultado de la simulación , sus respectivos estimadores y la desviación estándar de los mismos.

$n = 50$					
θ	α	λ	$\hat{\theta}$ (SD)	$\hat{\alpha}$ (SD)	$\hat{\lambda}$ (SD)
1	0	-1	1,143 (0,281)	-0,391(0,179)	-0,874 (0,229)
		0,1	1,040(0,273)	0,005 (0,218)	0,057 (0,436)
		1	1,259(0,402)	0,081 (0,179)	0,887(0,203)
2	-0,5	-0,5	2,064 (0,549)	-0,421 (0,143)	-0,274 (0,442)
		0,1	2,063 (0,658)	-0,497 (0,111)	0,088 (0,447)
		0,5	2,353(0,798)	-0,509 (0,099)	0,334 (0,451)
3	1	0,1	3,103 (0,750)	1,031 (0,449)	0,072 (0,439)
		0,5	3,174 (0,694)	1,001 (0,377)	0,383 (0,466)
		1	3,669 (1,106)	1,191 (0,420)	0,865 (0,248)
$n = 100$					
θ	α	λ	$\hat{\theta}$	$\hat{\alpha}$	$\hat{\lambda}$
1	0	-1	1,097 (0,193)	0,152 (0,274)	-0,899 (0,201)
		0,1	0,989 (0,204)	-0,031 (0,174)	0,056 (0,440)
		1	1,168 (0,317)	0,044 (0,121)	0,916 (0,171)
2	-0,5	-0,5	1,986 (0,393)	-0,442 (0,126)	-0,314 (0,425)
		0,1	1,993 (0,499)	-0,512 (0,088)	0,064 (0,430)
		0,5	2,232 (0,635)	-0,513 (0,077)	0,345 (0,415)
3	1	0,1	2,976 (0,537)	0,965 (0,337)	0,065 (0,429)
		0,5	3,119 (0,626)	0,965 (0,295)	0,374 (0,422)
		1	3,478 (0,875)	1,125 (0,291)	0,885(0,223)
$n = 300$					
θ	α	λ	$\hat{\theta}$	$\hat{\alpha}$	$\hat{\lambda}$
1	0	-1	1,053 (0,124)	0,100 (0,217)	-0,921 (0,181)
		0,1	0,965 (0,147)	-0,033 (0,121)	0,089 (0,385)
		1	1,103(0,226)	0,023 (0,070)	0,935 (0,157)
2	-0,5	-0,5	1,959 (0,215)	-0,480 (0,098)	-0,425 (0,323)
		0,1	1,949 (0,327)	-0,509 (0,062)	0,093 (0,343)
		0,5	2,117 (0,524)	-0,511 (0,045)	0,411 (0,317)
3	1	0,1	2,906 (0,366)	0,926 (0,237)	0,072 (0,392)
		0,5	3,044 (0,507)	0,952 (0,188)	0,415 (0,366)
		1	3,279 (0,614)	1,071 (0,175)	0,928 (0,175)

Tabla 1 : Algunos valores de los parámetros θ , α , λ con sus estimaciones[4]

7.0.2. Ilustración con datos reales

En esta sección, se usa un conjunto de datos reales para mostrar que la distribución RGT puede ser un mejor modelo que la Rayleigh y la Rayleigh transmutada, se hace uso de un conjunto de datos recopilados por la comisión federal de comercio que es una agencia independiente del gobierno de los Estados Unidos, cuya misión principal es la promoción de protección al consumidor. El informe titulado "tar, nicotine, and carbon monoxide of the smoke of 1206 varieties of domestic cigarettes for the year of 1998", el estudio en cuestión consiste en los conjuntos de datos y alguna información sobre la fuente de los datos, Comportamiento y creencias de los fumadores sobre los contenidos de nicotina, alquitrán y monóxido de carbono en los cigarrillos, el estudio contiene $n = 384$ observaciones, se analizan los datos sobre la nicotina, medido en miligramos por cigarrillo, de varias marcas de cigarrillos. Algunas estadísticas de resumen para los datos de nicotina son como sigue: $media = 0,852$, $mediana = 0,9$, $mnimo = 0,1$ y $mximo = 2$

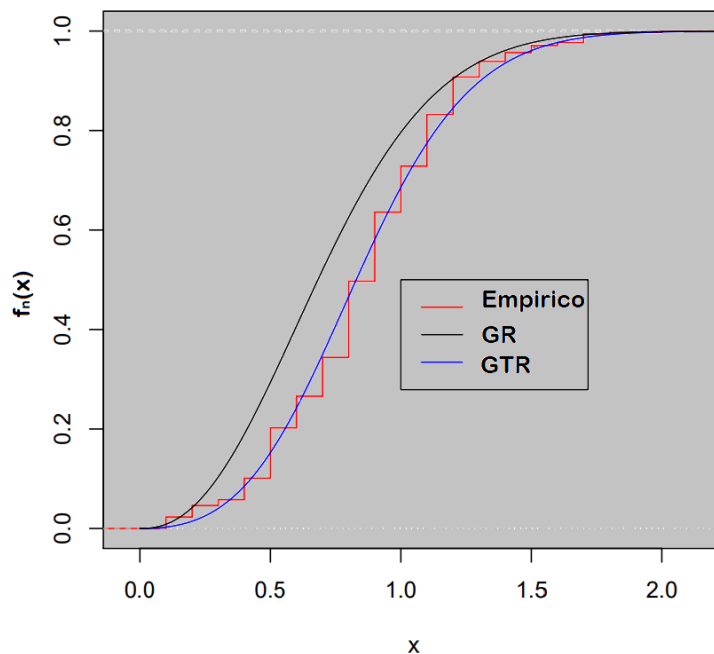


Figura 7.2: Modelado de datos empíricos con GR y RGT

Distribución	parámetros estimados	95 % I.C.	-LL
GTR	$\hat{\alpha} = 1.173$	[0.8971, 1.451]	112.443
	$\hat{\beta} = 1.579$	[1.239, 1.394]	
	$\hat{\lambda} = -0.681$	[-0.917, -0.445]	
GR	$\hat{\alpha} = 1.579$	[1.346, 1.811]	119.457
	$\hat{\beta} = 1.250$	[1.326, 1.175]	
TR	$\hat{\sigma} = 0.555$	[0.528, 0.582]	121.224
	$\hat{\lambda} = -0.7718095$	[-0.914, -0.629]	
R	$\hat{\sigma} = 0.6475387$	[0.618, 0.687]	142.3572

Tabla 2: Parámetros estimados de la distribución de Rayleigh, RT,GR y GTR para los datos de medición de nicotina

Distribución	Kolmogorov-Smirnov	AIC	AICC	BIC
GR	0.281	242.914	242.949	250.606
GTR	0.122	230.886	230.956	242.425

Tabla 3: Criterios de comparación usados para GR y TGR.

7.0.3. Uso en análisis de software

En administración de software con fines tanto comerciales como divulgativos es de suma importancia para un administrador estimar de la manera más precisa posible el tiempo, el esfuerzo y los potenciales fallos de un sistema próximo a ser lanzado; Por ejemplo en el proceso de desarrollo de un software, teniendo aproximadamente unos 9 errores por cada mil líneas de código, surge la pregunta ¿es acertado lanzar el software?; Este tipo de cuestiones son abordadas por Laird y Brennan en [15], en dicho texto se realizan este tipo de estimaciones basándose en el modelo Rayleigh (ver pág 124 de [15]), para ejemplificar el mejor ajuste del modelo RGT contra el modelo Rayleigh se realiza un breve análisis de los fallos de el programa R con el cual se han realizado las simulaciones y pruebas para el presente trabajo, en la pagina web (bugs.r-project.org/bugzilla) se encuentra un listado de los bugs del software

R reportados por los usuarios con diversas clasificaciones, por ejemplo, por fecha, por versión, por componente, por sistema operativo, por plataforma, entre otros. Resultado de un breve análisis de bugs por fecha de R versión 2.11.0 se tiene el siguiente histograma y el correspondiente modelado por los métodos antes expuestos.

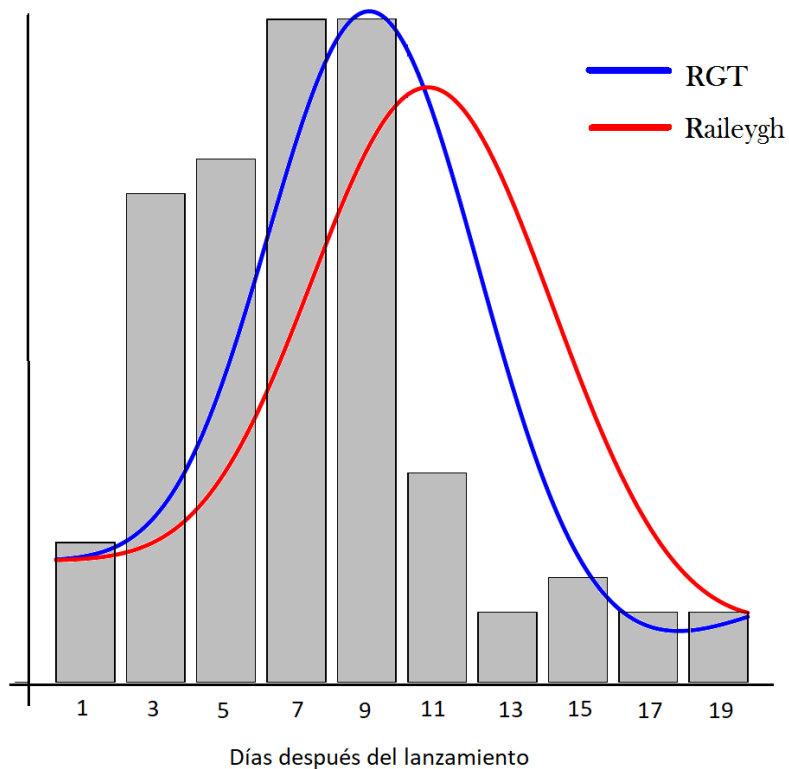


Figura 7.3: Comparación gráfica de ajuste

Capítulo 8

Conclusiones

- En el presente trabajo se evidencia la relación de la distribución RGT con otras importantes distribuciones como lo son la distribución chi cuadrado, la de Maxwell, distribución Half-normal o las diferentes variantes de la distribución de Rayleigh.
- Con el cálculo de los coeficientes de asimetría y de curtosis respectivamente se ha probado el hecho de que la distribución RGT es capaz de ajustar datos con estructura tipo Rayleigh generalizada pero con mejor ajuste en los datos.
- Dado que la distribución RGT posee un mejor ajuste que las distribuciones Rayleigh, RG y RT, constituye una mejor herramienta para análisis en variadas poblaciones de datos.
- Desde el punto de vista pragmático, dado que en el análisis de software se ha trabajado tradicionalmente con la distribución de Rayleigh (como se ve en [15]), la distribución RGT es un instrumento más apropiado para tal fin ya que todo aplicativo de software presenta defectos y el modelado de errores en fase de pruebas e implementación permite adelantar el total de defectos presentes optimizando de mejor manera los recursos.

- Desde una perspectiva más purista resulta fascinante dar cuenta de cómo hechos aparentemente inconexos al igual que un grupo de datos y una función con ciertas características guarden una relación tan estrecha lo que es una gran motivación para ahondar en el estudio de esta relación y encontrar otras.

Capítulo 9

Referencias

- 1 Shaw W. and Buckley I., "The alchemy of probability distributions: beyond gramcharlier expansions, and a skew-kurtotic-normal distribution from a rank transmutation map", IMA Primera Conferencia sobre Finanzas Computacionales, Cornell University,(2007).
- 2 Vodă VG., "Inferential procedures on a generalized Rayleigh variate I", *Apl. Mat.* 21 (1976), no. 6, 395-412.
- 3 Robert B. Ash, Catherine A. Doléans-Dade. *Probability and Measure Theory*, 2nd. Ed., Academic Press, (2000)
- 4 Iriarte, Yuri A, Astorga, Juan M. (2015). A version of transmuted generalized Rayleigh distribution. *Revista Integración*, 33(1), 83-95.
- 5 Aryal G.R. and Tsokos C.P., "On the transmuted extreme value distribution with applica- tion", *Nonlinear Anal.* 71 (2009), no. 12, 1401-1407.
- 6 Aryal G.R. and Tsokos C.P., "Transmuted Weibull distribution: a generalization of the Weibull probability distribution", *Eur. J. Pure Appl. Math.* 4 (2011), no. 2, 89-102.
- 7 Merovci F., "Transmuted Rayleigh distribution", *Aust. J. Stat.* 42 (2013), no. 1, 21-31.

- 8 Merovci F., “Transmuted Generalized Rayleigh Distribution”, *J. Stat. Appl. Prob.* 3 (2014), no. 1, 9-20.
- 9 C. Roberts and S. Gesser, A necessary and sufficient condition for the square of a random variable to be gamma, *Biometrika* 53 (1966), 275–278.
- 10 Warren Gilchrist, *Statistical modelling with quantile functions*, CRC Press Inc, 2000.
- 11 Johnson N.L., Kotz S. and Balakrishnan N., *Continuous univariate distributions*, John Wiley Sons, Inc., New York, 1994.
- 12 Hogg R.V. and Tanis E.A., *Probability and statistical inference*, Macmillan Publishing, New York, 1993.
- 13 Johnson N.L., Kotz S. and Balakrishnan N., *Continuous univariate distributions*, John Wiley Sons, Inc., New York, 1994.
- 14 Schwarz, Gideon E. , Estimating the dimension of a model, *Annals of Statistics*, 6 (2): 461–464, (1978)
- 15 L.M. Laird and M.C. Brennan. *Software Measurement and Estimation: A Practical Approach*. IEEE Computer Society and John Wiley Sons, Inc., Hoboken, New Jersey, 2006.