



Universidad Nacional Autónoma de México

---

---

Facultad de Ciencias

Sobre el método del cociente

T E S I S

QUE PARA OBTENER EL TÍTULO DE:  
MATEMÁTICO

PRESENTA:  
Emmanuel Delgadillo Garcia

DIRECTOR DE TESIS:  
Dr. Luis Antonio Rincón Solís



Cd.Mx. 2024



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

*Eso creo muy bien –dijo el cura–, que ya yo sé de experiencia que los montes crean letrados y las cabañas de los pastores encierran filósofos.*

Miguel de Cervantes, Don Quijote de la mancha.

# Agradecimientos

A la persona que más debo agradecer es a aquel hombre a quien la fortuna me ha permitido poder llamar padre; hombre gallardo, bonachón y generoso. Miembro de aquel género de hombres que nacen muy de vez en cuando y que van dejando tras de sí un rastro luminoso.

También debo agradecer infinitamente a mi madre, la cual hubiera podido llegar a ser reina, si la corona debiesen ceñirla solo las cabezas más lindas e inteligentes.

Doy un especial agradecimiento a mis amigos Francisco Javier Alvarado Cabrera, Alberto Isaac Estrella Madrigal, Lorenzo Antonio Alvarado Cabrera, Juan Carlos Luna García, Pedro Astudillo Ramírez, Eduardo Hernández Aguilar y Hugo Victor García Martínez. Porque sin aquellas buenas charlas en los pasillos, aquellas largas sobremesas y nuestras reuniones ocasionales, no hubiera sido tan placentero mi paso por la Facultad de Ciencias.

Corazón en mano, doy un sincero agradecimiento al doctor Luis Antonio Rincón Solís, sin cuya guía, paciencia y apoyo, este proyecto no lo hubiese podido llevar a cabo, porque como dice Don Cervates: "nadie nace enseñado, y de los hombres se hacen los obispos, que no de las piedras".

Finalmente, quiero agradecer a tres personas a quien la providencia me ha mandado, cuyo apoyo y calidez agradeceré toda la vida; mi tío Francisco Piliado Romo, mi hermana Ana Laura Delgadillo García, y la Dra. Elsa Puente Vázquez.

# Índice general

|  |           |
|--|-----------|
| <b>Introducción</b>  | <b>1</b>  |
| <b>1. Método del cociente</b>  | <b>4</b>  |
| 1.1. Método de aceptación y rechazo . . . . .                        | 4         |
| 1.2. Método del cociente de uniformes . . . . .                      | 11        |
| 1.3. Descripción en coordenadas polares . . . . .                    | 15        |
| 1.4. Recíproco del método del cociente . . . . .                     | 19        |
| 1.5. Transformación suma de una constante . . . . .                  | 24        |
| 1.6. Transformación producto por una constante . . . . .             | 27        |
| 1.7. Distribución exponencial . . . . .                              | 30        |
| 1.8. Distribución Cauchy . . . . .                                   | 34        |
| 1.9. Distribución normal . . . . .                                   | 39        |
| <b>2. Método general del cociente</b>                                | <b>45</b> |
| 2.1. Extensión del cociente de uniformes . . . . .                   | 46        |
| 2.2. Distribución exponencial bivariada . . . . .                    | 48        |
| 2.3. Distribución de densidad parabólica truncada . . . . .          | 50        |
| 2.4. Método del cociente de factor $r$ . . . . .                     | 53        |
| 2.5. La transformación $T_r$ . . . . .                               | 56        |
| 2.6. Optimización del valor de $r$ . . . . .                         | 65        |
| 2.7. Método general del cociente . . . . .                           | 71        |
| 2.8. Descripción en coordenadas polares . . . . .                    | 79        |
| 2.9. La transformación $T_g$ . . . . .                               | 82        |
| 2.10. Una región rectangular $S_g$ para la distribución normal . . . | 89        |
| 2.11. La región $S_g$ en ciertos casos particulares . . . . .        | 93        |

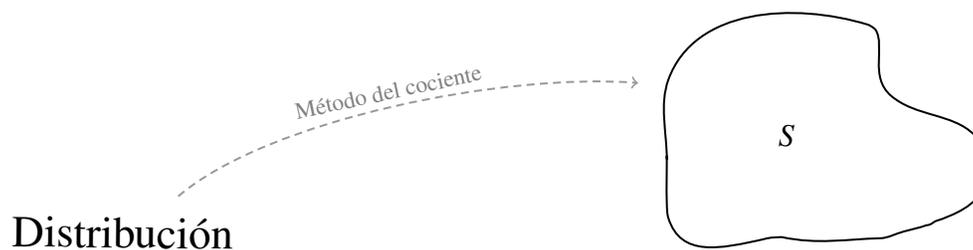
|                     |            |
|---------------------|------------|
| <b>Conclusiones</b> | <b>103</b> |
| <b>Apéndice 1</b>   | <b>110</b> |
| <b>Apéndice 2</b>   | <b>114</b> |
| <b>Bibliografía</b> | <b>116</b> |

# Introducción

Indudablemente, el método del cociente de uniformes es uno de los métodos de simulación de variables aleatorias por excelencia. No solo por ser sencillo de implementar computacionalmente, sino porque responde a una idea intuitiva que, sin exagerar, todos hemos tenido en algún momento: "subir una dimensión simplifica las cosas".

El método del cociente, en su forma mas elemental, establece una relación entre una distribución unidimensional y una distribución uniforme bidimensional. Dicha relación permite recuperar la distribución unidimensional a partir de la uniforme bidimensional; y el medio de pasar de una distribución a otra será un cociente, de donde proviene el nombre de este método.

Una distribución uniforme queda totalmente determinada por la región o conjunto sobre la cual toma valores; por lo que, implícitamente, el método del cociente asocia a cada distribución unidimensional una región en el plano, de tal manera que la distribución uniforme en dicha región nos permite recuperar la distribución unidimensional a partir de un cociente.



Este trabajo pretende mostrar los resultados básicos acerca del método del cociente, pero también se discutirán ciertos resultados que pueden ayudar a entender la relación entre distribuciones y regiones que establece el método del cociente de uniformes.

El capítulo 1 está destinado a estudiar la versión más elemental del método del cociente, la propuesta por Kinderman y Mohanan en [7]. Este capítulo está dividido en cuatro bloques. El primer bloque, constituido por las Secciones 1.1 y 1.2, estudia y exhibe los mecanismos que hacen funcionar al método del cociente. El segundo bloque, constituido únicamente de la Sección 1.3, describe a las regiones del método del cociente en coordenadas polares, lo cual, además de ser natural, permite determinar varias propiedades de las distribuciones a partir de sus regiones del método del cociente. El tercer bloque, constituido por las Secciones 1.5 y 1.6, estudian a las regiones provenientes de distribuciones que resultan de sumar o multiplicar una constante, y se demuestra que estas regiones son iguales salvo un isomorfismo lineal, es decir, que existe un isomorfismo lineal que lleva una región en la otra. Finalmente, el último bloque está destinado únicamente a ejemplos y está constituido por las Secciones 1.7, 1.8 y 1.9.

El capítulo 2, a diferencia del capítulo 1, es de una constitución más robusta en términos de contenido, y es el capítulo principal y final de este trabajo.

El capítulo 2 se encuentra dividido en tres bloques. El primer bloque extiende el método del cociente a varias dimensiones, y se desarrolla a lo largo de las Secciones 2.1, 2.2 y 2.3. El segundo bloque, constituido por las Secciones 2.4, 2.5 y 2.6, extiende aún más el método del cociente al introducir un factor  $r$ , el cual juega un activo papel determinando a las regiones resultantes. El tercer bloque, y el más importante, se desarrolla en las Secciones 2.7, 2.8, 2.9, 2.10 y 2.11, en donde se estudia el método del cociente en su versión más general, la cual introduce una función auxiliar  $g(x)$  que viene a generalizar al factor  $r$ .

En el segundo bloque del capítulo 2 se estudiará escrupulosamente al factor  $r$  que es introducido, y ello nos permitirá observar que las regiones que se obtienen de aplicar el método del cociente para los distintos valores de  $r$  son

iguales, salvo un difeomorfismo que lleva una región en la otra. Además, este análisis del factor  $r$  nos proporcionará los medios por los cuales se podrá llevar a cabo una optimización de este factor.

Con respecto al tercer bloque del capítulo 2, el trabajo se centrará en el análisis de la función auxiliar  $g(x)$ , la cual, como veremos, determina activamente a la región del método del cociente. Dicho análisis estará orientado a hallar funciones auxiliares para las cuales las regiones resultantes sean sencillas de simular, tales como rectángulos o secciones circulares. Esta tarea algunas veces se podrá lograr al plantear ecuaciones diferenciales ordinarias, aunque esto dependerá intrínsecamente de la distribución que busquemos simular.

# Capítulo 1

## Método del cociente

En diversas áreas de la ciencia, problemas complejos son abordados y entendidos a partir de algunos más sencillos. En este sentido se presenta el método del cociente, método por el cual, bajo determinadas condiciones, el problema de simular variables aleatorias es reducido a uno más simple, el de generar valores de distribución uniforme. Sin embargo, para que el método del cociente sea factible, se requiere del método de aceptación y rechazo. En consecuencia, para lograr un mayor entendimiento del método del cociente, es conveniente estar familiarizado con el método de aceptación y rechazo.

El presente capítulo se consagra a un escrupuloso estudio del método del cociente de uniformes propuesto por Kinderman y Monahan en [7]. Dicho estudio comienza por habituarnos al método de aceptación y rechazo.

### 1.1. Método de aceptación y rechazo

El método de aceptación y rechazo es estudiado en una gran variedad de trabajos, pues resulta particularmente útil cuando los métodos directos no pueden ser aplicados o son computacionalmente ineficientes. Entre esta gran variedad de trabajos podemos encontrar las principales fuentes de esta sección: el trabajo de Liang, Liu y Carrol en [8] y el trabajo de Martino, Luengo y Míguez en [10].

Existen distribuciones más complejas que otras, en el sentido de que son más difíciles de simular. Un nítido ejemplo de esto es proporcionado por la distribución Bernoulli de parámetro  $1/2$  y la distribución Weibull. Para generar valores de la distribución Bernoulli de parámetro  $1/2$  basta con tener una moneda justa, lanzarla, y tomar el valor 0 si cae águila y el valor 1 si cae sol. Por otro lado, el generar valores de la distribución Weibull, sin importar el valor de sus parámetros, puede llegar a ser desafiante experimentalmente, ya que no son evidentes los experimentos aleatorios que se llevan a cabo con distribución Weibull.

El método de aceptación y rechazo permite generar valores de una distribución auxiliándose de otra distribución más sencilla de simular. Evidentemente, ciertas condiciones son impuestas para que el método sea aplicable a cierto par de distribuciones, pero éstas son admisibles.

**Proposición 1.1 (Método de aceptación y rechazo)**

Sean  $h(x)$  y  $g(x)$  funciones de densidad, y suponga que existe una constante  $M \geq 1$  tal que  $h(x) \leq Mg(x)$ . Si  $X \sim g$  y  $U \sim Unif(0,1)$  son independientes, entonces

$$X \left| \left( U \leq \frac{h(X)}{Mg(X)} \right) \right. \sim h.$$

Es decir, si  $u$  y  $x$  son valores de la distribución uniforme sobre el intervalo  $(0,1)$  y de la distribución  $g$ , respectivamente, y además satisfacen la desigualdad

$$u \leq \frac{h(x)}{Mg(x)},$$

entonces  $x$  es un valor de la distribución  $h$ .

**Demostración.**

Notemos que para  $x \in \mathbb{R}$ ,

$$\mathbb{P} \left( X \leq x \left| U \leq \frac{h(X)}{Mg(X)} \right. \right) = \frac{\mathbb{P} \left( X \leq x, U \leq \frac{h(X)}{Mg(X)} \right)}{\mathbb{P} \left( U \leq \frac{h(X)}{Mg(X)} \right)}.$$

Por otro lado,

$$\begin{aligned} \mathbb{P}\left(X \leq x, U \leq \frac{h(X)}{Mg(X)}\right) &= \int_{-\infty}^x g(t) \cdot \mathbb{P}\left(U \leq \frac{h(X)}{Mg(X)} \mid X = t\right) dt \\ &= \int_{-\infty}^x g(t) \cdot \mathbb{P}\left(U \leq \frac{h(t)}{Mg(t)}\right) dt \end{aligned} \quad (1.1)$$

$$\begin{aligned} &= \int_{-\infty}^x g(t) \cdot \frac{h(t)}{Mg(t)} dt \\ &= \frac{1}{M} \int_{-\infty}^x h(t) dt. \end{aligned} \quad (1.2)$$

Así,

$$\begin{aligned} \mathbb{P}\left(U \leq \frac{h(X)}{Mg(X)}\right) &= \lim_{x \rightarrow \infty} \mathbb{P}\left(X \leq x, U \leq \frac{h(X)}{Mg(X)}\right) \\ &= \lim_{x \rightarrow \infty} \frac{1}{M} \int_{-\infty}^x h(t) dt \\ &= \frac{1}{M}. \end{aligned} \quad (1.3)$$

Consecuentemente, por (1.2) y (1.3), concluimos que

$$\mathbb{P}\left(X \leq x \mid U \leq \frac{h(X)}{Mg(X)}\right) = \int_{-\infty}^x h(t) dt,$$

es decir,  $X \mid \left(U \leq \frac{h(X)}{Mg(X)}\right) \sim h$ . □

### Observación 1.1

*En la demostración de la Proposición 1.1 no se menciona explícitamente dónde es usada la hipótesis de independencia de  $U$  y  $X$ , pero es fácil notar que dicha hipótesis es la que establece la igualdad en (1.1).*

En la siguiente proposición se generaliza el método de aceptación y rechazo a funciones de densidad conjunta  $h(x_1, \dots, x_n)$  y  $g(x_1, \dots, x_n)$  que cumplen la condición  $h(x_1, \dots, x_n) \leq Mg(x_1, \dots, x_n)$ , para una constante  $M \geq 1$ .

**Proposición 1.2 (Método general de aceptación y rechazo)**

Sean  $h(x_1, \dots, x_n)$  y  $g(x_1, \dots, x_n)$  funciones de densidad conjunta, y suponga que existe una constante  $M \geq 1$  tal que  $h(x_1, \dots, x_n) \leq Mg(x_1, \dots, x_n)$ . Si  $(X_1, \dots, X_n) \sim g$  y  $U \sim \text{Unif}(0, 1)$  son independientes, entonces

$$(X_1, \dots, X_n) \left| \left( U \leq \frac{h(x_1, \dots, x_n)}{Mg(x_1, \dots, x_n)} \right) \right. \sim h.$$

**Demostración.**

La demostración para el caso  $n = 2$  es exactamente igual, *mutatis mutandi*, a la del caso general. Así que, por brevedad en la escritura, se presentará únicamente la demostración para el caso  $n = 2$ .

Notemos que para  $(x_1, x_2) \in \mathbb{R}^2$ ,

$$\mathbb{P} \left( X_1 \leq x_1, X_2 \leq x_2 \mid U \leq \frac{h(X_1, X_2)}{Mg(X_1, X_2)} \right) = \frac{\mathbb{P} \left( X_1 \leq x_1, X_2 \leq x_2, U \leq \frac{h(X_1, X_2)}{Mg(X_1, X_2)} \right)}{\mathbb{P} \left( U \leq \frac{h(X_1, X_2)}{Mg(X_1, X_2)} \right)}.$$

Por otro lado,

$$\begin{aligned} \mathbb{P} \left( X_1 \leq x_1, X_2 \leq x_2, U \leq \frac{h(X_1, X_2)}{Mg(X_1, X_2)} \right) &= \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} g(t, s) \mathbb{P} \left( U \leq \frac{h(X_1, X_2)}{Mg(X_1, X_2)} \mid (X_1, X_2) = (t, s) \right) ds dt \\ &= \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} g(t, s) \mathbb{P} \left( U \leq \frac{h(t, s)}{Mg(t, s)} \right) ds dt \\ &= \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} g(t, s) \cdot \frac{h(t, s)}{Mg(t, s)} ds dt \\ &= \frac{1}{M} \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} h(t, s) ds dt. \end{aligned}$$

De esta manera,

$$\begin{aligned} \mathbb{P}\left(U \leq \frac{h(X_1, X_2)}{Mg(X_1, X_2)}\right) &= \lim_{x_2 \rightarrow \infty} \lim_{x_1 \rightarrow \infty} \mathbb{P}(X_1 \leq x_1, X_2 \leq x_2, U \leq \frac{h(X_1, X_2)}{Mg(X_1, X_2)}) \\ &= \frac{1}{M} \end{aligned}$$

Con ello concluimos que

$$\mathbb{P}\left(X_1 \leq x_1, X_2 \leq x_2 \mid U \leq \frac{h(X_1, X_2)}{Mg(X_1, X_2)}\right) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} h(t, s) ds dt,$$

es decir,  $(X_1, X_2) \mid (U \leq \frac{h(X_1, X_2)}{Mg(X_1, X_2)}) \sim h$ . □

La Proposición 1.2 justifica el siguiente algoritmo para generar valores de distribución  $h$  cuando la hipótesis  $h(x_1, \dots, x_n) \leq Mg(x_1, \dots, x_n)$  es satisfecha por alguna función de densidad  $g(x_1, \dots, x_n)$  y una constante  $M \geq 1$ .

• **Algoritmo de aceptación y rechazo**

1. Generar un valor  $(x_1, \dots, x_n)$  de distribución  $g$  y un valor  $u$  de distribución uniforme en el intervalo  $(0, 1)$ .
2. Tomar a  $(x_1, \dots, x_n)$  como valor de la distribución  $h$  si  $u \leq \frac{h(x_1, \dots, x_n)}{Mg(x_1, \dots, x_n)}$ .
3. Repetir los pasos 1 y 2 tantas veces como sea necesario.

**Ejemplo 1.1 (Distribución beta)**

Sea  $g$  la densidad de la distribución uniforme sobre el intervalo  $(0, 1)$  y denote por  $h$  la función de densidad de la distribución beta( $a, b$ ), esto es,

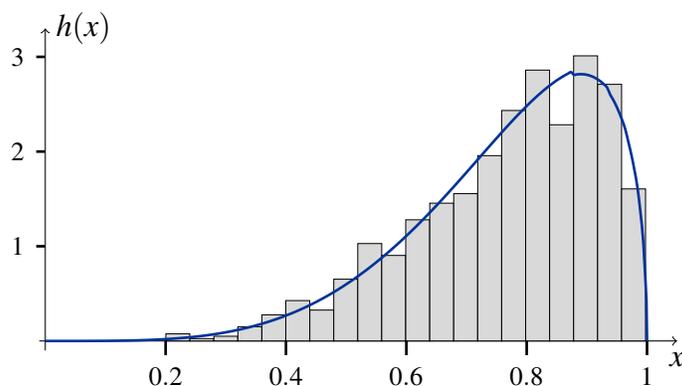
$$h(x) = \frac{1}{B(a, b)} x^{a-1} (1-x)^{1-b}, \quad 0 < x < 1,$$

con  $a = 5$  y  $b = 3/2$ . Se puede comprobar fácilmente que  $h$  es una función acotada, a decir verdad, se encuentra acotada por 3. De esta manera, en virtud de que el soporte de  $h$  es el intervalo  $(0, 1)$ , se tiene que  $h(x) \leq 3g(x)$ .

Por lo tanto, para generar valores de distribución  $h$ , es suficiente generar un valor  $x$  de la distribución  $g$  y un valor  $u$  de la distribución uniforme sobre intervalo  $(0, 1)$ , y aceptar  $x$  sólo si satisface

$$u \leq \frac{h(x)}{3g(x)} = \frac{1}{3}h(x). \quad (1.4)$$

En la Figura 1.1 se compara la gráfica de  $h$  y el histograma de 1000 valores de distribución  $h$  generados a partir del método de aceptación y rechazo, con  $g(x)$  como función de densidad auxiliar. Es decir, se generaron valores  $x$  y  $u$  de distribución uniforme sobre  $(0, 1)$  hasta obtener 1000 valores que satisficieran la condición (1.4).



**Figura 1.1 .** Histograma vs densidad  $beta(5, 3/2)$ .

○

### **Observación 1.2**

*Note que en el Ejemplo 1.1 se puede contemplar plenamente la fortaleza y eficacia del método de aceptación y rechazo. En dicho ejemplo hemos sido capaces de generar valores de una distribución cuya simulación directa puede ser particularmente difícil, y esto lo logramos al llevar a cabo una tarea mucho más simple, el generar valores de distribución uniforme en el intervalo  $(0, 1)$ .*

### Ejemplo 1.2 (Distribución uniforme)

Sean  $A$  y  $B$  un par regiones de  $\mathbb{R}^n$  con volumen finito distinto de cero tal que  $A \subseteq B$  (ver Figura 1.2).

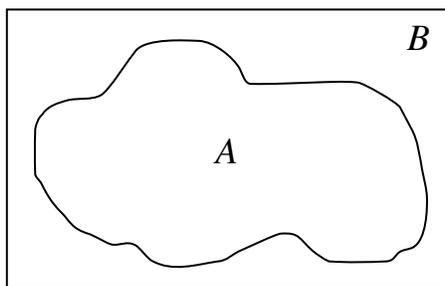


Figura 1.2 . Ejemplo con  $n = 2$  y  $B$  un rectángulo.

Las funciones de densidad

$$h(x) := \frac{1}{\text{vol}(A)} \mathbb{1}_A(x) \quad \text{y} \quad g(x) := \frac{1}{\text{vol}(B)} \mathbb{1}_B(x), \quad x \in \mathbb{R}^n,$$

de las distribuciones uniformes sobre  $A$  y  $B$ , respectivamente, satisfacen  $h(x) \leq \frac{\text{vol}(B)}{\text{vol}(A)} g(x)$ . Por lo tanto, para generar un valor de distribución uniforme sobre  $A$  basta generar valores  $x$  de distribución uniforme sobre  $B$  y  $u$  de distribución uniforme sobre el intervalo  $(0, 1)$  hasta que se satisfaga

$$u \leq \frac{h(x)}{\frac{\text{vol}(B)}{\text{vol}(A)} g(x)} = \mathbb{1}_A(x)$$

En consecuencia, para generar valores de distribución uniforme sobre la región  $A$ , es suficiente el generar valores de distribución uniforme sobre la región  $B$  y aceptar sólo aquellos que estén en  $A$ . Esto es particularmente útil cuando la región  $B$  es un rectángulo (o hiperectángulo en caso de que  $n > 2$ ), pues el generar valores de distribución uniforme en un rectángulo es fácil, lo que implica que podemos generar valores de distribución uniforme en  $A$ , sin importar cuán irregular o compleja sea la región  $A$ .  $\circ$

## 1.2. Método del cociente de uniformes

Ya hemos dicho que el método del cociente, bajo determinadas condiciones, permite simular variables aleatorias a partir de valores de distribución uniforme. Sin embargo, es pertinente mencionar que la región sobre la que se generan los valores de la distribución uniforme es totalmente determinada por la función de densidad de la distribución que buscamos simular. Este hecho es notable porque, en la práctica, puede que se desconozca la distribución, pero se conozca la distribución empírica, por lo que podemos aplicar el método del cociente a la distribución empírica, si es que cumple con los requerimientos del método.

### Proposición 1.3 (Método del cociente de uniformes)

Sea  $h(x)$  una función de densidad acotada y suponga que  $\sup_{x \in \mathbb{R}} |x| \sqrt{h(x)} < \infty$ .

Entonces la región

$$S = \left\{ (u, v) \in \mathbb{R}^2 : 0 < u < \sqrt{h(v/u)} \right\} \quad (1.5)$$

es de área finita y si  $(U, V)$  es un vector con distribución uniforme sobre  $S$ , entonces  $V/U \sim h$ .

### Demostración.

El valor absoluto del cociente de dos números reales es igual al cociente del valor absoluto de dichos números. Por lo que, en particular, se tiene que si  $u > 0$ , entonces

$$|v/u| = |v|/u.$$

Por lo tanto, para todo  $(u, v)$  en  $S$ , ocurre que

$$|v| = (|v|/u) \cdot u < |v/u| \sqrt{h(v/u)}. \quad (1.6)$$

Luego, debido a que  $\sup_{x \in \mathbb{R}} |x| \sqrt{h(x)} < \infty$ , se sigue de (1.6) que  $S$  está acotada y que, de hecho,

$$S \subseteq \left[ 0, \sup_{x \in \mathbb{R}} \sqrt{h(x)} \right] \times \left[ -\sup_{x \in \mathbb{R}} |x| \sqrt{h(x)}, \sup_{x \in \mathbb{R}} |x| \sqrt{h(x)} \right].$$

De esta manera, como el cálculo de áreas es una operación monótona, es decir, que preserva contenciones, concluimos que el área de  $S$  es finita.

Suponga que  $(U, V)$  es un vector aleatorio con distribución uniforme sobre  $S$ , entonces la función de densidad conjunta de  $(U, V)$  está dada por

$$f_{U,V}(u, v) = \frac{1}{|S|}, \quad \text{para } (u, v) \in S.$$

Por la fórmula general para la función de densidad del cociente de dos variables aleatorias continuas (véase el Apéndice 1), tenemos que, para  $-\infty < x < \infty$ , se cumplen la siguiente serie de igualdades

$$\begin{aligned} f_{V/U}(x) &= \int_{-\infty}^{\infty} f_{U,V}(u, xu) \cdot |u| du \\ &= \int_{(u, xu) \in S} \frac{1}{|S|} \cdot |u| du \\ &= \frac{1}{|S|} \cdot \int_0^{\sqrt{h(x)}} u du \\ &= \frac{1}{2|S|} \cdot h(x). \end{aligned}$$

Luego, como  $f_{V/U}(x)$  y  $h(x)$  son funciones de densidad, al integrar sobre  $(-\infty, \infty)$  deducimos que

$$|S| = \frac{1}{2}.$$

Por lo tanto,  $h(x) = f_{V/U}(x)$ . □

### Observación 1.3

- (I) En la prueba de la Proposición 1.3 es demostrado que la región  $S$  está contenida en un rectángulo. Esto es vital para la viabilidad del método; porque, en virtud del Ejemplo 1.2, con el método de aceptación y rechazo podemos generar fácilmente valores de distribución uniforme en  $S$ , sin importar cuan irregular o compleja sea la región  $S$ .
- (II) En la Proposición 1.3 no se imponen demasiadas condiciones a  $h(x)$ , a decir verdad, sólo se pide que las funciones  $|x|\sqrt{h(x)}$  y  $h(x)$  sean acotadas. Sin embargo, no hay un algoritmo bien determinado para verificar

que un conjunto o función es acotado, así que es conveniente hallar una condición suficiente que sea más computable.

En la Proposición 1.4 se establecen condiciones suficientes para que una función de densidad cumpla con las hipótesis de la Proposición 1.3. Estas condiciones son dadas a través del concepto de límite y continuidad.

**Proposición 1.4** *Sea  $h(x)$  una función de densidad continua. Si  $h(x)$  satisface que*

1.  $\lim_{x \rightarrow -\infty} x\sqrt{h(x)} < \infty$ ,
2.  $\lim_{x \rightarrow \infty} x\sqrt{h(x)} < \infty$ .

Entonces  $h(x)$  es acotada y  $\sup_{x \in \mathbb{R}} |x|\sqrt{h(x)} < \infty$ , es decir,  $h(x)$  satisface las hipótesis de la Proposición 1.3.

**Demostración.**

Sean  $a := \lim_{x \rightarrow -\infty} x\sqrt{h(x)}$  y  $b := \lim_{x \rightarrow \infty} x\sqrt{h(x)}$ . Haciendo uso de la desigualdad del triángulo inversa y la definición *épsilon-delta* de continuidad, tomemos  $r > 0$  de tal manera que para todo  $|x| > r$

$$|x|\sqrt{h(x)} \leq \max\{1 + |a|, 1 + |b|\}.$$

Sin pérdida de generalidad suponga que  $r \geq 1$ . Entonces para todo  $|x| > r$  se tiene que

$$h(x) = \left(\sqrt{h(x)}\right)^2 \leq \left(|x|\sqrt{h(x)}\right)^2 \leq (\max\{1 + |a|, 1 + |b|\})^2.$$

Esto demuestra que las funciones  $h(x)$  y  $|x|\sqrt{h(x)}$  son acotadas en  $(-\infty, -r) \cup (r, \infty)$ , por lo que es suficiente mostrar que estas dos funciones son también acotadas en el intervalo  $[-r, r]$  y para ello usaremos la hipótesis de continuidad de la función de densidad.

Veamos, debido a que la función  $h(x)$  es continua se tiene que las funciones  $|x|\sqrt{h(x)}$  y  $h(x)$  son continuas en el intervalo  $[-r, r]$ , de este modo ambas funciones son acotadas en dicho intervalo y, de hecho, por el teorema

de *Weierstrass* (ver [4]) alcanzan su máximo y mínimo en dicho intervalo. Consecuentemente, concluimos que las funciones  $h(x)$  y  $|x|\sqrt{h(x)}$  son acotadas.  $\square$

#### **Observación 1.4**

*En la demostración de la Proposición 1.4 podemos apreciar que las hipótesis de existencia y finitud de los límites sólo asegura que las funciones estén acotadas en  $(-\infty, -r) \cup (r, \infty)$  para algún  $r > 0$ . Mientras que la continuidad de  $h(x)$  nos asegura que están acotadas en el intervalo  $[-r, r]$ . Sin embargo, para que las funciones sean acotadas en  $[-r, r]$  es suficiente que  $h(x)$  sea discontinua excepto en un número finito de puntos y que en dichos puntos la función de densidad tenga límites finitos por la izquierda y por la derecha.*

La Proposición 1.3 justifica el siguiente algoritmo para generar valores de la distribución  $h$  cuando las hipótesis de dicha proposición son satisfechas.

#### **• Algoritmo del cociente de uniformes**

1. Generar valores  $(u, v)$  de distribución uniforme en la región  $S$  definida como en (1.5).
2. Tomar a  $v/u$  como valor de distribución  $h$ .

Tanto el método de aceptación y rechazo como el del cociente de uniformes requieren generar valores auxiliares. No obstante, el generar valores de distribución uniforme es más fácil que el generar valores de cualquier otra distribución. Por lo tanto, en este sentido, el método del cociente es más factible que el método de aceptación y rechazo.

Otra considerable ventaja teórica del método del cociente sobre el de aceptación y rechazo se halla en el hecho de que en el algoritmo de aceptación y rechazo algunos valores son descartados, mientras que en el algoritmo del cociente de uniformes cada valor auxiliar que se genera es aprovechado. Sin embargo, en la práctica dicha ventaja se verá perdida en gran medida, ya que para calcular valores de distribución uniforme en la región  $S$  (descrita en (1.5)) usualmente usaremos la Observación 1.3(I).

### 1.3. Descripción en coordenadas polares

La condición que determina a la región  $S$  del método del cociente es bastante concisa e inteligible. Sin embargo, describir a dicha región en coordenadas polares puede resultar ventajoso algunas veces, ya sea para lograr un mayor entendimiento de la región o para alcanzar una descripción mucho más compacta. Además, la descripción en coordenadas polares nos permite dilucidar el cómo la distribución  $h$  determina algunas características geométricas de su región  $S$ , posibilitando esbozar a la distribución  $h$  a partir de su región  $S$ .

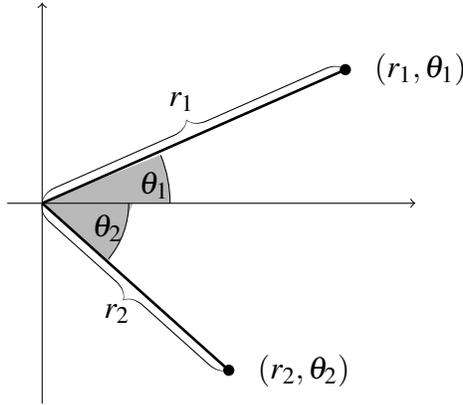
Para describir a la región  $S$  en coordenadas polares primero vamos a recordar brevemente este sistema de coordenadas y acordaremos la forma en que los ángulos serán medidos en este trabajo. Inmediatamente después, nos daremos a la tarea de interpretar en coordenadas polares la condición que determina a la región  $S$ .

#### Coordenadas polares

Las coordenadas polares de un punto en el plano (distinto del origen) es un par ordenado  $(r, \theta)$ , donde  $r$  es la longitud del segmento de recta que une al punto con el origen y  $\theta$  es el ángulo que forma dicho segmento con el eje  $x$ . En consecuencia, las coordenadas polares de un punto en el plano están sujetas a la forma en que los ángulos son medidos. A lo largo de todo este trabajo los ángulos serán medidos en radianes de  $-\pi$  a  $\pi$ , empleando la orientación usual, es decir, la dirección positiva será la contraria a las manecillas del reloj. Haciendo uso de esta convención, en la Figura 1.3 se ilustran dos puntos de coordenadas polares  $(r_1, \theta_1)$  y  $(r_2, \theta_2)$  con  $\theta_2 < 0 < \theta_1$ .

#### Notación

En este trabajo a la primer coordenada polar le llamaremos coordenada radial, mientras que a la segunda coordenada polar le llamaremos coordenada angular. Además, se reservará el uso de las letras  $r$  y  $\theta$  para denotar las coordenadas radiales y angulares. De esta manera un par ordenado  $(r, \theta)$  siempre denotará las coordenadas polares de un punto.



**Figura 1.3 .** Coordenadas polares de un par de puntos.

Todas las regiones que nos determina el método del cociente se encuentran confinadas a los cuadrantes 1 y 4 con el eje  $X$ . Este hecho es bastante conveniente, porque a los puntos de esta región fácilmente les podemos calcular su coordenada angular a partir de sus coordenadas cartesianas, a saber, si tenemos un punto con coordenadas cartesianas  $(x, y)$  (con  $x > 0$ ), entonces su coordenada angular, la cual conveniente denotaremos por  $\theta(x, y)$ , es exactamente  $\arctan(y/x)$ , es decir,  $\theta(x, y) = \arctan(y/x)$ . Así mismo, para mantener una notación consistente, denotaremos a la coordenada radial del punto  $(x, y)$  por  $\rho(x, y)$ .

### Interpretación en coordenadas polares

#### Lema 1.1 (Descripción en coordenadas polares)

Sea  $h(x)$  una función de densidad acotada y suponga que  $\sup_{x \in \mathbb{R}} |x| \sqrt{h(x)} < \infty$ .

Entonces

$$\begin{aligned}
 S &= \left\{ (u, v) \in \mathbb{R}^2 : 0 < u < \sqrt{h(v/u)} \right\} \\
 &= \left\{ (r, \theta) : \theta \in \left( -\frac{\pi}{2}, \frac{\pi}{2} \right), 0 < r < \frac{\sqrt{h(\tan(\theta))}}{\cos(\theta)} \right\}. \quad (1.7)
 \end{aligned}$$

**Demostración.**

Al cambiar de coordenadas cartesianas a polares obtenemos el siguiente par de equivalencias

$$\begin{aligned} 0 < u < \sqrt{h(v/u)} &\iff 0 < \rho(u, v) \cdot \cos(\theta(u, v)) < \sqrt{h(\tan(\theta(u, v)))} \\ &\iff 0 < \rho(u, v) < \frac{\sqrt{h(\tan(\theta(u, v)))}}{\cos(\theta(u, v))}. \end{aligned}$$

De donde se sigue inmediatamente el lema.  $\square$

En un principio podemos pensar que la descripción en coordenadas polares de  $S$ , la cual se encuentra en (1.7), parece ser mas complicada que la que teníamos en coordenadas cartesianas y puede que sea así en algunos casos, pero hay algunos otros donde la descripción en coordenadas polares se simplificará más y, por supuesto, también habrá casos donde ambas descripciones sean complicadas. Sin embargo, la descripción polar deriva en una serie de observaciones que la hacen un tanto más rica que la descripción cartesiana.

**Observación 1.5 (Regiones de haces)**

Si definimos en  $(-\pi/2, \pi/2)$  la función  $\rho(\theta) = \frac{\sqrt{h(\tan(\theta))}}{\cos(\theta)}$ , entonces

$$S = \left\{ (r, \theta) : \theta \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right), 0 < r < \rho(\theta) \right\}. \quad (1.8)$$

Además, el hecho de que  $h(x)$  y  $|x|\sqrt{h(x)}$  sean acotadas implica que la función  $\rho(\theta)$  que definimos sea acotada, pues de la definición de  $\rho(\theta)$  es fácil ver que

$$h(\tan(\theta)) = \rho^2(\theta) \cos^2(\theta) \quad \text{y} \quad |\tan(\theta)|\sqrt{h(\tan(\theta))} = |\sen(\theta)|\rho(\theta).$$

En consecuencia, para cada  $\theta \in (-\pi/2, \pi/2)$  el conjunto de todos los puntos de  $S$  con ángulo  $\theta$  es un segmento sin extremos que parte del origen y cuya longitud es  $\rho(\theta)$ . De esta manera, las regiones que son producidas a través del método del cociente son regiones de haces sin extremos que parten del origen y cuyos ángulos están entre  $-\pi/2$  y  $\pi/2$ .

La siguiente observación explica el efecto geométrico que tiene en  $S$  la continuidad de  $h(x)$  en un intervalo.

**Observación 1.6 (Regularidad)**

Suponga que  $-\infty \leq a < b \leq \infty$  son tales que  $h(x)$  es continua en el intervalo  $(a, b)$ . Recordando que la función tangente (restringida al intervalo  $(-\pi/2, \pi/2)$ ) es la inversa de la función arco tangente y que ambas son crecientes, se sigue fácilmente de la propia definición que  $\rho(\theta)$  es continua en el intervalo  $(\alpha, \beta)$  donde  $\alpha := \arctan(a)$  y  $\beta := \arctan(b)$ . Consecuentemente, la curva en coordenadas polares  $(\rho(\theta), \theta)$  que delimita a la región  $S$  en los rayos de pendiente  $a$  y  $b$  es continua, lo que hará que gráficamente la parte de  $S$  que se encuentra entre los rayos de pendiente  $a$  y  $b$  se vea como un sector circular deformado continuamente, tal como lo sugiere la Figura 1.4.

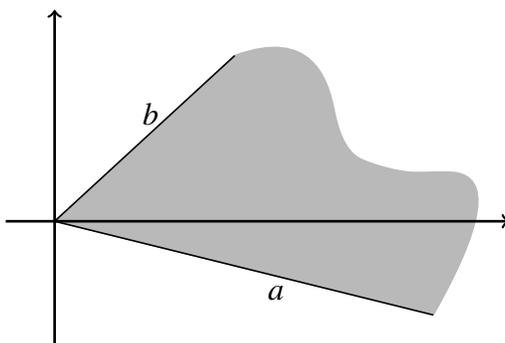


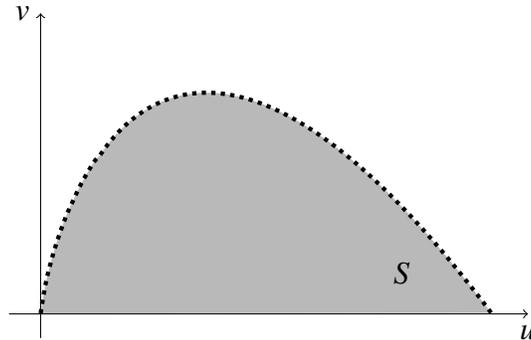
Figura 1.4 . Parte de  $S$  entre rectas de pendiente  $a$  y  $b$ .

**Observación 1.7 (Interpretación geométrica del soporte)**

Notemos que para toda  $x \in \mathbb{R}$  se tiene

$$\begin{aligned} \rho(\arctan(x)) = 0 &\iff \frac{\sqrt{h(\tan(\arctan(x)))}}{\cos(\arctan(x))} = 0 \\ &\iff h(x) = 0. \end{aligned}$$

En consecuencia,  $h(x) = 0$  si y solo si la intersección de  $S$  con la recta de pendiente  $x$  es igual al vacío. De esta manera, si la región  $S$  fuera como en la Figura 1.5, entonces el soporte de  $h(x)$  se encuentra en el intervalo  $(0, \infty)$  ya que la intersección de  $S$  con cualquier recta de pendiente negativa (o cero) es igual al vacío.



**Figura 1.5 .** Posible región  $S$ .

### Observación 1.8

Suponga que  $X$  es una variable aleatoria con función de densidad  $h$  y defina  $Y := \arctan(X)$ . Al aplicar el teorema de cambio de variable se deduce que la función de densidad  $f_Y(\theta)$  de  $Y$  está dada por

$$\begin{aligned} f_Y(\theta) &= (1 + \tan^2(\theta))h(\tan(\theta)) \cdot \mathbb{1}_{(-\pi/2, \pi/2)} \\ &= \frac{h(\tan(\theta))}{\cos^2(\theta)} \cdot \mathbb{1}_{(-\pi/2, \pi/2)}. \end{aligned}$$

Por ende,  $\rho^2(\theta)$  es igual a la función de densidad de  $\arctan(X)$ , o más correctamente dicho,  $\rho^2(\theta)$  iguala la función de densidad de  $\arctan(X)$  en el soporte de ésta.

Para dar por concluida esta sección, note que la descripción polar tiene una ventaja innata sobre la cartesiana, pues mientras que la descripción cartesiana requiere de dos variables, la descripción polar solo requiere de una.

## 1.4. Recíproco del método del cociente

Para cualquier región  $R \subseteq \mathbb{R}^2$  de área finita podemos tomarnos un vector aleatorio  $(U, V)$  de distribución uniforme sobre  $R$ , y a partir de este vector se determina una distribución particular, la correspondiente a la variable aleatoria  $X := V/U$ . Así que una cuestión interesante es la de saber cuándo dos regiones  $R_1$  y  $R_2$  determinan a una misma distribución a partir del cociente  $V/U$ .

**Observación 1.9**

Sea  $R \subseteq \mathbb{R}^2$  una región de área finita y  $(U, V) \sim \text{unif}(R)$ . Por la fórmula general para la función de densidad del cociente de variables aleatorias (ver Apéndice 1), la función de densidad de  $V/U$  está dada por

$$f_{V/U}(y) = \frac{1}{\text{Área}(R)} \int_{-\infty}^{\infty} |x| \mathbb{1}_R(x, xy) dx. \quad (1.9)$$

Si denotamos por  $l_y(R)$  a la intersección de la región  $R$  y la recta que pasa por el origen con pendiente  $y$ , entonces  $\mathbb{1}_{l_y(R)}(x) = \mathbb{1}_R(x, xy)$ , para toda  $y \in \mathbb{R}$ . Por lo tanto, la ecuación (1.9) se puede reescribir como:

$$f_{V/U}(y) = \frac{1}{\text{Área}(R)} \int_{-\infty}^{\infty} |x| \mathbb{1}_{l_y(R)}(x) dx. \quad (1.10)$$

De esta manera, la función de densidad de la variable  $V/U$  se encuentra totalmente determinada por los conjuntos  $l_y(R)$ .

**Definición 1.1** Para cada región  $R \subseteq \mathbb{R}^2$  de área finita y  $y \in \mathbb{R}$ , definimos el conjunto

$$l_y(R) := \left\{ (u, v) \in R : \frac{v}{u} = y \right\}.$$

Es decir,  $l_y(R)$  es la colección de puntos en  $R$  con pendiente igual a  $y$ .

**Observación 1.10**

Sabemos que los únicos subconjuntos conexos de una recta, dotada de su topología usual, son los segmentos de recta. De esta manera, para cada  $y \in \mathbb{R}$ , el conjunto  $l_y(R)$  se puede ver como la unión ajena de segmentos de la recta de pendiente igual a  $y$ . Además, cualquier subconjunto de la recta tiene a lo más una cantidad numerable de componentes conexas, por lo que sin pérdida de generalidad puede suponerse que esta unión ajena de intervalos es numerable.

**Nota:** Dado que cada segmento que conforma a  $l_y(R)$  es parametrizado a través de la asignación  $t \rightarrow (t, ty)$ , en la siguiente proposición entenderemos al conjunto  $l_y(R)$  como la unión de todas sus componentes conexas parametrizadas con la asignación  $t \rightarrow (t, ty)$ .

**Proposición 1.5** Sean  $R_1, R_2 \subseteq \mathbb{R}^2$  regiones de área finita. Para cualesquiera  $(U_1, V_1) \sim \text{unif}(R_1)$  y  $(U_2, V_2) \sim \text{unif}(R_2)$ , se tiene que las variables  $V_1/U_1$  y  $V_2/U_2$  son idénticamente distribuidas si y solo si, para cada  $y \in \mathbb{R}$ ,

$$\int_{l_y(R_2)} E dt = \frac{\text{Área}(R_2)}{\text{Área}(R_1)} \int_{l_y(R_1)} E dt, \quad (1.11)$$

donde  $E(x, y) = |x|$ .

**Demostración.**

Notemos que para cada  $y \in \mathbb{R}$ , si definimos la curva  $\gamma(t) = (t, ty)$  en un intervalo finito  $(a, b)$ , entonces

$$\begin{aligned} \int_{\gamma} E dt &= \int_a^b \|\gamma'(t)\| E(\gamma(t)) dt \\ &= \int_a^b \sqrt{1+y^2} E(t, ty) dt \\ &= \sqrt{1+y^2} \int_a^b |t| dt. \end{aligned}$$

De esta manera, para  $k \in \{1, 2\}$ , la función de densidad de  $V_k/U_k$  está dada por

$$f_{V_k/U_k}(y) = \frac{1}{\text{Área}(R) \sqrt{1+y^2}} \int_{l_y(R_k)} E dt.$$

Por lo tanto, las variables  $V_1/U_1$  y  $V_2/U_2$  son idénticamente distribuidas si y solo si, para toda  $y \in \mathbb{R}$ ,

$$\frac{1}{\text{Área}(R_1) \sqrt{1+y^2}} \int_{l_y(R_1)} E dt = \frac{1}{\text{Área}(R_2) \sqrt{1+y^2}} \int_{l_y(R_2)} E dt.$$

De donde se deriva inmediatamente la igualdad en (1.11).  $\square$

Evidentemente, la Proposición 1.5 da respuesta a la pregunta con la cual iniciamos esta sección, pues establece que dos regiones  $R_1$  y  $R_2$  van a determinar la misma distribución si, para cada  $y \in \mathbb{R}$ , la integral de línea de  $E(x, y)$  sobre  $l_y(R_1)$  y sobre  $l_y(R_2)$  son múltiplos escalares uno del otro y, de hecho, el escalar es la razón entre sus áreas.

**Observación 1.11**

Es sencillo constatar que la igualdad en (1.11) establece una relación de equivalencia en el conjunto de regiones de área finita en el plano, por lo cual la asociación entre distribuciones y regiones que establece el método del cociente puede ser trasladada al cociente de esta relación de equivalencia y, así, obtener una relación biunívoca. Sin embargo, no hay necesidad de trasladarse al cociente de la relación de equivalencia, porque podemos hallar un representante bastante amigable para cada una de estas clases.

**Proposición 1.6 (Recíproco del método del cociente)**

Sea  $\rho : (-\frac{\pi}{2}, \frac{\pi}{2}) \rightarrow [0, \infty)$  una función 2-integrable y  $\|\rho\|_2 = 1$ . Si las funciones  $\cos(\theta) \cdot \rho(\theta)$  y  $|\sen(\theta)| \cdot \rho(\theta)$  son acotadas, entonces la función  $h(x)$  definida en  $(-\infty, \infty)$  por

$$h(x) := \frac{\rho^2(\arctan(x))}{1+x^2} \quad (1.12)$$

cumple con las siguientes propiedades:

1. La función  $h(x)$  es de densidad.
2. Las funciones  $h(x)$  y  $|x|\sqrt{h(x)}$  son acotadas.
3.  $\left\{ (u, v) : 0 < u < \sqrt{h\left(\frac{v}{u}\right)} \right\} = \left\{ (r, \theta) : |\theta| < \frac{\pi}{2}, 0 < r < \rho(\theta) \right\}$ .

**Demostración.**

Por el teorema de cambio de variable tenemos que

$$\begin{aligned} \int_{-\infty}^{\infty} h(x) dx &= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} (1 + \tan^2(\theta)) \cdot h(\tan(\theta)) d\theta \\ &= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \rho^2(\theta) d\theta \\ &= 1. \end{aligned}$$

De esta manera, como es evidente que  $h(x)$  es no negativa, concluimos que  $h(x)$  es una función de densidad.

Para mostrar (2) basta observar que para toda  $\theta \in (-\pi/2, \pi/2)$ ,

$$h(\tan(\theta)) = \cos^2(\theta) \cdot \rho^2(\theta) \quad \text{y} \quad |\tan(\theta)|\sqrt{h(\tan(\theta))} = |\sin(\theta)|\rho(\theta).$$

Finalmente, para mostrar (3), observemos que para toda  $(u, v)$  en  $(0, \infty) \times \mathbb{R}$  se tiene que

$$\begin{aligned} 0 < u < \sqrt{h(v/u)} &\iff 0 < \sqrt{u^2 + v^2} \cdot \cos(\theta(u, v)) < \sqrt{h(\tan(\theta(u, v)))} \\ &\iff 0 < \sqrt{u^2 + v^2} < \rho(\theta(u, v)). \end{aligned}$$

Donde  $\theta(u, v)$  denota el ángulo del vector  $(u, v)$  en  $(-\pi/2, \pi/2)$ . □

### **Observación 1.12**

*En la Observación 1.5 se establece que toda región que surge al aplicar el método del cociente es de la forma descrita en (1.8). Mientras que en la Proposición 1.6 se demuestra que, de hecho, toda región de la forma descrita en (1.8) determina una función de densidad que cumple con los requerimientos del método del cociente y que, al aplicarle dicho método, la región  $S$  que surge coincide exactamente con la región que determinó a la función de densidad.*

### **Observación 1.13**

*La Proposición 1.6, en conjunto con la discusión llevada a cabo en la Sección 1.3, puede ser utilizada para producir distribuciones con ciertas características que deseemos. Además, aprovechando el carácter geométrico, y amparándonos del método de aceptación y rechazo, fácilmente puede diseñarse un programa computacional para llevar a cabo esta tarea, es decir, producir distribuciones con determinadas características a partir del método del cociente.*

*Antes de dar paso a la siguiente sección, en vista de la Proposición 1.6, nos permitimos decir que, en un carácter subjetivo, las regiones de haces, es decir, las regiones delimitadas por una función en coordenadas polares, son las regiones inherentes al método del cociente.*

## 1.5. Transformación suma de una constante

Esta sección es en esencia una observación, pero una notable, por ello la abordaremos con el carácter nominal de sección.

Suponga que  $h(x)$  es una función de densidad que cumple con los requerimientos del método del cociente y  $X \sim h$ . Sea  $a \in \mathbb{R}$ . Al definir  $Y := X + a$  obtenemos una nueva variable aleatoria con una distribución un tanto distinta de la de  $X$  y de la cual es interesante preguntarnos si cumple con los requerimientos del método del cociente y, si es así, ¿cuál es la región que nos arroja el método del cociente?

Observemos que  $Y = \phi(X)$  donde  $\phi(t) = t + a$  para toda  $t \in \mathbb{R}$ . De esta manera, como  $\phi(t)$  es un difeomorfismo, por el teorema de cambio de variable, se tiene que la función de densidad de  $Y$  está dada por

$$g(y) = h(\phi^{-1}(y)) \cdot |(\phi^{-1})'(y)|, \quad -\infty < y < \infty.$$

Por ende, en virtud de que  $\phi^{-1}(y) = y - a$ , para toda  $y \in \mathbb{R}$ , se deduce que

$$g(y) = h(y - a).$$

Es decir,  $g(y)$  es un simple desplazamiento de  $h(x)$  y, en consecuencia,  $g(y)$  cumple con las hipótesis de la Proposición 1.3.

Muy bien, ahora denotemos por  $S_{X+a}$  a la región que surge al aplicar el método del cociente a la densidad  $g(y)$ , y de la misma manera denotemos por  $S_X$  a la que surge al aplicarlo a  $h(x)$ .

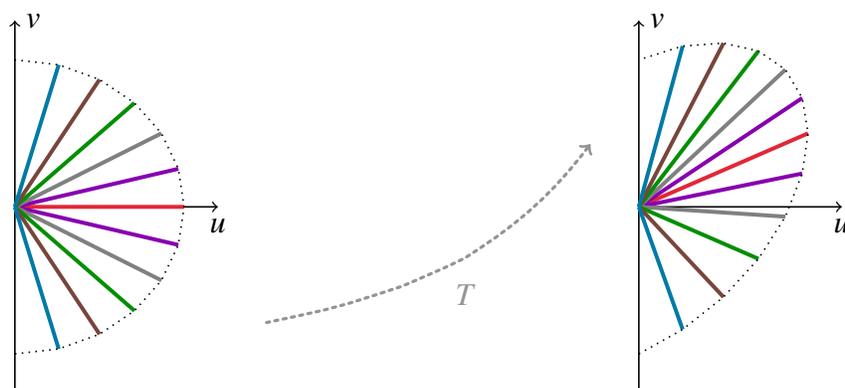
Al igual que  $g(y)$  se deriva de  $h(x)$ ,  $S_{X+a}$  se deriva de  $S_X$  a través de una transformación lineal y el describir esta transformación lineal nos revelará los efectos que tiene el sumar una simple constante.

Veamos, para  $(u, v) \in \mathbb{R}^2$  se tiene la siguiente cadena de equivalencias

$$\begin{aligned} (u, v) \in S_{X+a} &\iff 0 < u < \sqrt{g(v/u)} \\ &\iff 0 < u < \sqrt{h\left(\frac{v}{u} - a\right)} \\ &\iff (u, v - au) \in S_X. \end{aligned}$$

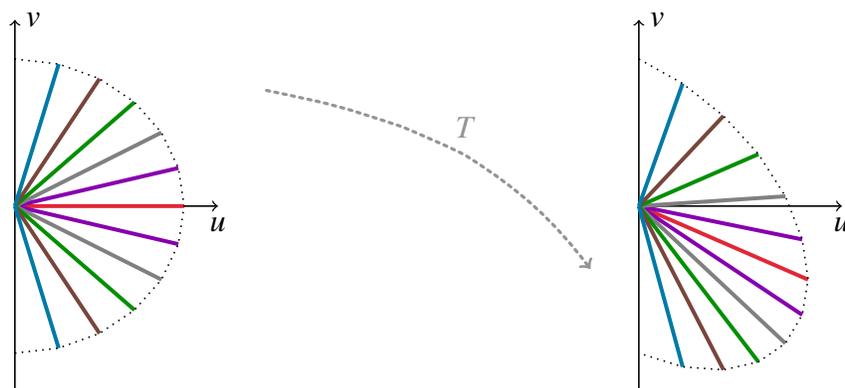


De forma particular, si  $a > 0$ , el desplazamiento de los cuadrantes 1 y 4 será hacia arriba. En la Figura 1.7 se muestra de manera gráfica el desplazamiento hacia arriba provocado por  $T(u, v)$  al tomar  $a = 1/2$  y también podemos observar que al tomar un valor pequeño de  $a$  el desplazamiento es también pequeño.



**Figura 1.7 .** Efecto de la transformación  $x + a$  con  $a > 0$ .

Asimismo, si  $a < 0$ , entonces el desplazamiento de los cuadrantes 1 y 4 será hacia abajo. En la Figura 1.8 se exhibe el desplazamiento provocado por  $T(u, v)$  al tomar  $a = -1/2$  y en dicha figura podemos observar que el desplazamiento es igual que cuando se toma  $a = 1/2$ , solo que ahora se lleva a cabo hacia abajo.



**Figura 1.8 .** Efecto de la transformación  $x + a$  con  $a < 0$ .

### Observación 1.14

Es evidente que hemos llevado a cabo el análisis de  $S_{X+a}$  en coordenadas cartesianas sin reparar siquiera en un análisis en coordenadas polares. La razón de esto es que en coordenadas cartesianas las expresiones son más sencillas de interpretar geoméricamente.

## 1.6. Transformación producto por una constante

Retomemos a  $h(x)$  una función de densidad que cumple los requerimientos del método del cociente y  $X \sim h$ . Si tomamos  $a \in \mathbb{R}$  distinto de cero y definimos  $Y := aX$ , entonces la variable  $Y$  tiene una función de densidad  $g(y)$  la cual también cumple los requerimientos del método del cociente y, además, su región puede verse como una deformación lineal de la correspondiente a  $h(x)$ .

Veamos, sea  $a \neq 0$ , entonces por el teorema de cambio de variable la función de densidad  $g(y)$  de  $Y$  está dada por

$$g(y) = \frac{1}{|a|} h\left(\frac{y}{a}\right), \quad -\infty < y < \infty. \quad (1.14)$$

De (1.14) se sigue inmediatamente que  $g(y)$  cumple con los requerimientos del método del cociente. En vista de esto, denote por  $S_{aX}$  a la región que surge al aplicar el método del cociente a  $g(y)$  y por  $S_X$  a la región que surge al aplicarlo a  $h(x)$ .

Dividamos el análisis en dos casos:  $a > 0$  y  $a < 0$ . Iniciemos abordando el caso en que  $a > 0$ .

Si  $a > 0$ , entonces para  $(u, v) \in \mathbb{R}^2$  se tiene la siguiente cadena de equivalencias

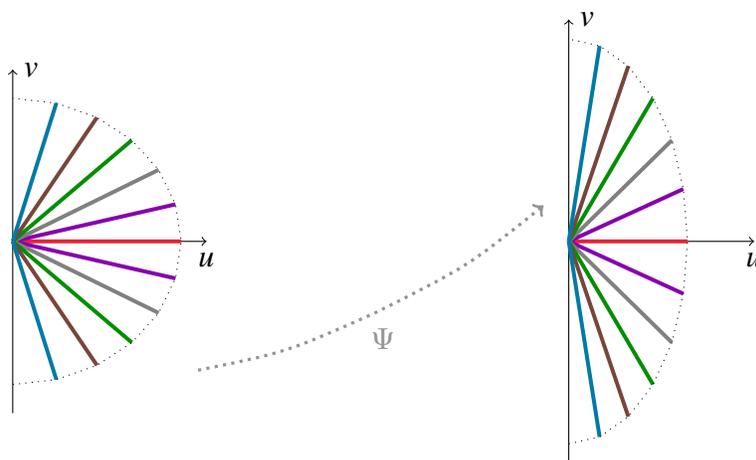
$$\begin{aligned} (u, v) \in S_{aX} &\iff 0 < u < \sqrt{g(v/u)} \\ &\iff 0 < u\sqrt{a} < \sqrt{h(v/au)} \\ &\iff (u\sqrt{a}, v/\sqrt{a}) \in S_X. \end{aligned}$$

De esta manera, concluimos que

$$S_{aX} = \{(u/\sqrt{a}, v\sqrt{a}) : (u, v) \in S_X\}. \quad (1.15)$$

En la ecuación (1.15) se describe a  $S_{aX}$  como la deformación de  $S_X$  bajo la transformación lineal  $\Psi(u, v) := (u/\sqrt{a}, v\sqrt{a})$ .

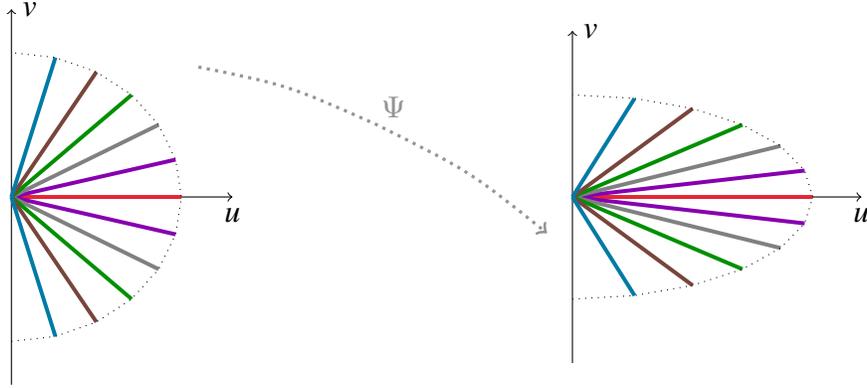
Es fácil observar que  $\Psi(1, 0) = (1/\sqrt{a}, 0)$  y  $\Psi(0, 1) = (0, \sqrt{a})$ , por lo que  $\Psi(u, v)$  es un isomorfismo lineal, el cual solo reescala (positivamente) a la base canónica. De esta manera, el efecto que tendrá sobre cualquier vector del plano será una reescalamiento de sus coordenadas, y el comportamiento de este reescalamiento dependerá del valor de  $a$ . Pero lo que sí podemos decir de este reescalamiento es que el efectuado en la coordenada  $x$  es inversamente proporcional al valor de  $a$  y, por el contrario, el efectuado en la coordenada  $y$  es directamente proporcional al valor de  $a$ . Es decir, entre más grande sea el valor de  $a$ , mayor será el reescalamiento de la coordenada  $y$  y menor el de la coordenada  $x$ .



**Figura 1.9 .** Efecto de la transformación  $ax$  con  $a > 0$ .

En la Figura 1.9 se muestra el comportamiento de  $\Psi(u, v)$  cuando  $a = 2$ . En dicho caso se tiene que  $a > 1$ , por lo que  $\Psi(u, v)$  contrae el eje  $x$ , mientras que expande el eje  $y$ , lo cual es apreciado en la Figura 1.9.

Por otro lado, si  $a < 1$  entonces  $\Psi(u, v)$  expandirá el eje  $x$  y contraerá el eje  $y$ . Este hecho puede ser apreciado en la Figura 1.10, en donde se muestra el comportamiento de  $\Psi(u, v)$  cuando  $a = 1/2$ .



**Figura 1.10 .** Efecto de la transformación  $ax$  con  $a < 0$ .

Muy bien, ahora supongamos que  $a < 0$ , entonces  $|a| = -a$ . En consecuencia, para todo  $(u, v) \in \mathbb{R}^2$  tenemos la siguiente cadena de equivalencias

$$\begin{aligned}
 (u, v) \in S_{aX} &\iff 0 < u < \sqrt{g(v/u)} \\
 &\iff 0 < u < \sqrt{\frac{1}{-a} \cdot h\left(\frac{v}{au}\right)} \\
 &\iff 0 < u\sqrt{-a} < \sqrt{h\left(\frac{-v/\sqrt{-a}}{u\sqrt{-a}}\right)} \\
 &\iff (u\sqrt{-a}, -v/\sqrt{-a}) \in S_X \\
 &\iff (u\sqrt{|a|}, -v/\sqrt{|a|}) \in S_X.
 \end{aligned}$$

Por lo tanto, concluimos que

$$S_{aX} = \left\{ (u/\sqrt{|a|}, -v/\sqrt{|a|}) : (u, v) \in S_X \right\}.$$

Así pues, la función que transforma  $S_X$  en  $S_{aX}$  es  $\xi(u, v) = (u/\sqrt{|a|}, -v/\sqrt{|a|})$ , la cual prácticamente es la transformación  $\psi(u, v)$  de parámetro  $|a|$ , salvo que el signo menos en la segunda coordenada invierte la orientación del eje  $Y$ . Por ende, concluimos que si  $a < 0$ , entonces  $S_{aX}$  es la reflexión con respecto al eje  $u$  de  $S_{|a|X}$

## 1.7. Distribución exponencial

Por la Observación 1.4, la distribución exponencial cumple con los requerimientos del método del cociente, así que en esta sección aplicaremos dicho método a la distribución exponencial de parámetro  $\lambda = 1$  y posteriormente lo haremos para algunas distribuciones que se derivan de ella, en particular la distribución exponencial de parámetro  $\lambda > 0$ .

### Distribución exponencial de parámetro $\lambda = 1$

Suponga entonces que  $h(x)$  es la función de densidad de la distribución exponencial de parámetro  $\lambda = 1$ , es decir,

$$h(x) = \begin{cases} e^{-x} & \text{para } x > 0, \\ 0 & \text{en otro caso.} \end{cases}$$

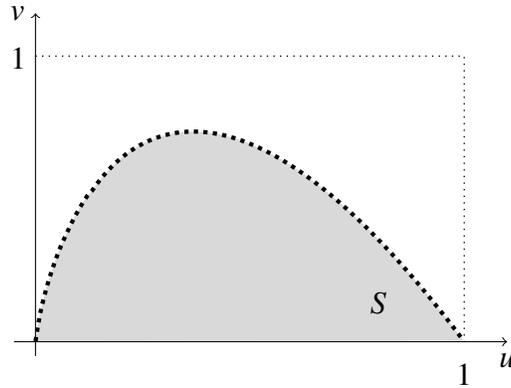
La región  $S$  que resulta de aplicar el método del cociente a  $h(x)$  está descrita en coordenadas cartesianas por

$$S = \{(u, v) : 0 < u < 1, 0 < v < -2u \ln(u)\}, \quad (1.16)$$

mientras que en coordenadas polares dicha región se encuentra descrita por

$$S = \left\{ (r, \theta) : \theta \in \left(0, \frac{\pi}{2}\right), 0 < r < \frac{1}{\cos(\theta)} \cdot e^{-\frac{\tan(\theta)}{2}} \right\}. \quad (1.17)$$

En la Figura 1.11 se encuentra ilustrada la región  $S$ , en donde también puede apreciarse que  $S$  se encuentra contenida en el rectángulo  $(0, 1) \times (0, 1)$ . Acorde al método del cociente, si generamos  $(u, v)$  de distribución uniforme sobre  $S$ , entonces  $v/u$  será de distribución  $h$ . Este hecho lo veremos puesto en práctica cuando al final esta sección simulemos a  $h(x)$  con el método del cociente.



**Figura 1.11 .** Región  $S$  para la distribución  $\exp(1)$ .

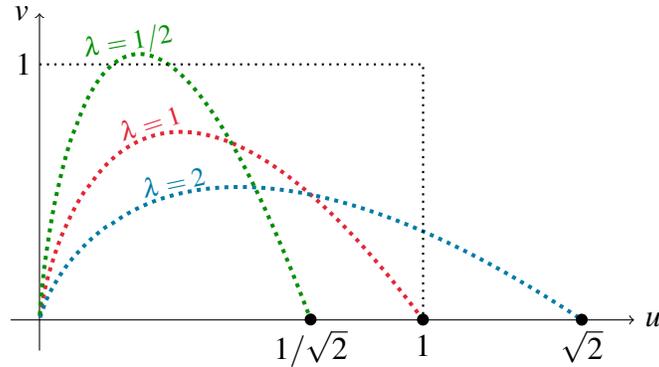
En esta ocasión  $S$  es descrita de forma mas sencilla en coordenadas cartesianas, pues en (1.16) se constata que  $S$  es la región bajo la gráfica de una función dada en términos de funciones elementales.

### Producto por una constante

La distribución exponencial de parámetro  $\lambda > 0$  es resultado de aplicar a la distribución  $\exp(1)$  la transformación que estudiamos en la Sección 1.6, pues si  $X \sim \exp(1)$ , entonces  $\frac{1}{\lambda}X \sim \exp(\lambda)$  para  $\lambda > 0$ . Por lo que, de lo discutido en la Sección 1.6, tenemos una idea clara de cómo será la región  $S_{\frac{1}{\lambda}}$  que resulta de aplicar el método del cociente a la distribución  $\exp(\lambda)$ . Realizando los cálculos correspondientes confirmamos esta intuición, pues la descripción en coordenadas cartesianas de esta región es

$$S_{\lambda} = \left\{ (u, v) \in \mathbb{R}^2 : 0 < u < \sqrt{\lambda}, 0 < v < -\frac{2u}{\lambda} \log \left( \frac{u}{\sqrt{\lambda}} \right) \right\}.$$

En la Figura 1.12 se ilustran las curvas que, con el eje  $x$ , delimitan a  $S_{\lambda}$  para ciertos valores de  $\lambda$ . En dicha figura notamos que si  $\lambda > 1$ , entonces  $1/\lambda < 1$ , con lo cual la región  $S_{\lambda}$  será una expansión de  $S$  en dirección del eje  $x$  y una contracción en dirección del eje  $y$ . Igualmente podemos notar que si  $\lambda < 1$ , entonces  $1/\lambda > 1$ , por lo que en dicho caso  $S_{\lambda}$  será una contracción de  $S$  en dirección del eje  $x$  y una expansión en dirección del eje  $y$ .



**Figura 1.12 .** Región  $S_\lambda$  para la distribución  $\exp(\lambda)$ .

En virtud de la Sección 1.6, la región correspondiente a la distribución de  $\frac{1}{\lambda}X$  con  $X \sim \exp(1)$  y  $\lambda < 0$  es la reflexión de  $S_{\frac{1}{|\lambda|}}$  con respecto al eje  $u$ . Dicho hecho puede ser comprobado fácilmente en este caso particular, pues si  $\lambda < 0$ , entonces tras una sencilla serie de cálculos se llega a que la región  $S_{\frac{1}{\lambda}}$  se encuentra dada por

$$\begin{aligned} S_{\frac{1}{\lambda}} &= \left\{ (u, v) : 0 < u < \sqrt{-\lambda}, -\frac{2u}{\lambda} \ln \left( \frac{u}{\sqrt{-\lambda}} \right) < v < 0 \right\} \\ &= \left\{ (u, v) : 0 < u < \sqrt{|\lambda|}, \frac{2u}{\sqrt{|\lambda|}} \ln \left( \frac{u}{\sqrt{|\lambda|}} \right) < v < 0 \right\}. \end{aligned}$$

### Suma de una constante

Suponga que  $X \sim \exp(1)$  y sea  $a \in \mathbb{R}$ . La función de densidad de  $Y := X + a$  esta dada por

$$g(y) = \begin{cases} e^{-y+a} & \text{para } y > a, \\ 0 & \text{en otro caso.} \end{cases}$$

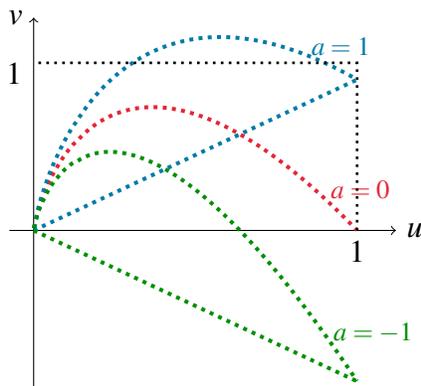
Denotemos por  $S_{X+a}$  la región que resulta de aplicar el método del cociente a  $g(y)$  y nuevamente denotemos por  $S$  a la que se obtiene al aplicar el método a la distribución  $\exp(1)$ . Por la discusión llevada a cabo en la Sección 1.5, sabemos que si  $a > 0$  entonces  $S_{X+a}$  será un desplazamiento hacia

arriba de  $S$  y la magnitud de este desplazamiento depende de la de  $a$ . Por otro lado, si  $a < 0$ , entonces  $S_a$  será un desplazamiento hacia abajo de  $S$  y la magnitud de dicho desplazamiento depende de la de  $|a|$ .

Pues bien, realizando el cálculo correspondiente se obtiene que

$$S_{X+a} = \{(u, v) : 0 < u < 1, au < v < -2u \ln(u) + au\}.$$

Ilustrando  $S_{X+a}$  para algunos valores de  $a$  (ver Figura 1.13) podemos observar que el desplazamiento lineal hacia arriba de  $S$  engañosamente se percibe como una rotación positiva y una modificación en la norma de sus puntos. Así mismo, el desplazamiento lineal hacia abajo sugiere engañosamente una simple rotación negativa y un modificación en la norma de sus puntos.

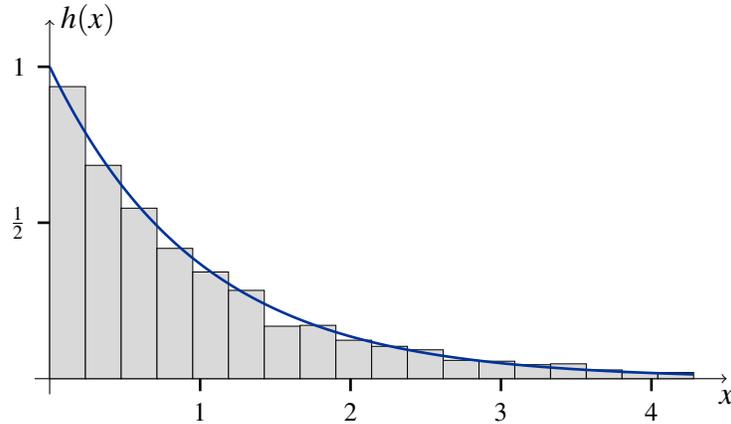


**Figura 1.13 .** Región  $S_{X+a}$  en el caso  $\lambda = 1$ .

## Simulación

En la Figura 1.14 se encuentra ilustrado el histograma de 1500 valores de distribución  $\exp(1)$  generados a partir del método del cociente. Sobre el histograma se sobrepone la gráfica de la función de densidad de esta misma distribución, con lo que podemos observar que el histograma se ajusta bastante bien a la gráfica como se esperaba.

Los valores de distribución uniforme sobre  $S$  que fueron requeridos se generaron a partir del método de aceptación y rechazo, auxiliandonos del paquete estadístico R para generar valores de distribución uniforme sobre el rectángulo  $(0, 1) \times (0, 1)$  que contiene a la región  $S$ .



**Figura 1.14 .** Histograma vs densidad de la distribución exponencial con  $\lambda = 1$

## 1.8. Distribución Cauchy

La distribución *Cauchy*( $\eta, \gamma$ ) es absolutamente continua y cumple con las hipótesis de la Proposición 1.4. Así que, tal como hicimos con la distribución exponencial, aplicaremos el método del cociente a la distribución de Cauchy con los parámetros más simples, la estándar, y de la región que de ello surge vamos a analizar las regiones que se obtienen al variar los parámetros.

### Distribución de Cauchy estándar

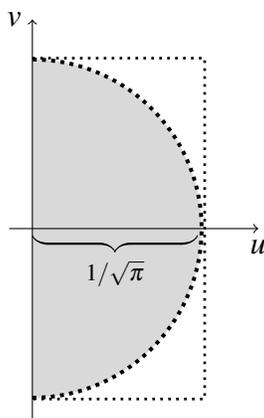
Considere  $h(x)$  la función de densidad de la distribución *Cauchy*(0, 1), esto es,  $h(x) = [\pi \cdot (1 + x^2)]^{-1}$ ,  $-\infty < x < \infty$ . La región  $S$  que surge al aplicar el método del cociente a  $h(x)$  esta descrita en coordenadas cartesianas por

$$S = \left\{ (u, v) \in \mathbb{R}^2 : 0 < u, u^2 + v^2 < \frac{1}{\pi} \right\}, \quad (1.18)$$

mientras que en coordenadas polares dicha región se encuentra dada por

$$S = \left\{ (\theta, r) : \theta \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right), 0 < r < \frac{1}{\sqrt{\pi}} \right\}.$$

Es evidente de (1.18) que  $S$  es el semicírculo derecho con centro en el origen y radio  $1/\sqrt{\pi}$  (ver Figura 1.15). Por ello, esta vez la región  $S$  es descrita de forma mucho más compacta y sencilla en coordenadas polares, pues en coordenadas polares se tiene que



**Figura 1.15** . Región  $S$  para la distribución  $\text{Cauchy}(0,1)$ .

### Observación 1.15

Acorde al método del cociente, si generamos  $(u, v)$  de distribución uniforme sobre  $S$ , entonces  $v/u$  será de distribución  $h$ . Esto resulta interesante para la distribución Cauchy, ya que si tuviera media finita, entonces el valor de dicha media sería la pendiente esperada de un punto  $(u, v)$  tomado con distribución uniforme en el semicírculo de la Figura 1.15. Por lo que, si un arquero disparara con fuerza y ángulo aleatorio (dentro de sus  $180^\circ$  de visión frontal), el peor ángulo para posicionarse sería el correspondiente a esta pendiente esperada. Sin embargo, debido a que la distribución de Cauchy no tiene media finita, tal "pendiente esperada" no existe.

Si bien no existe una "pendiente esperada", debe notarse que las pendientes cercanas a cero, las correspondientes a ángulos centrales, son las que tiene mayor densidad de probabilidad y dicha densidad disminuye conforme las pendientes son más pronunciadas. Esto se debe a que la distribución de Cauchy, al igual que la distribución normal (estándar), tiene forma de campana, siendo 0 el número con mayor densidad de probabilidad.

### Observación 1.16

La simetría de la distribución de  $Cauchy(0, 1)$  con respecto a 0 implica que su región  $S$  sea simétrica respecto al eje  $u$ , lo cual se sigue fácilmente de la expresión de la función  $\rho(\theta)$  que la delimita a su región  $S$  en coordenadas polares. De hecho, al examinar la expresión general de la función  $\rho(\theta)$  podemos observar fácilmente que la simetría respecto a 0 es otra característica de una distribución que puede ser determinada a partir de la región que se obtiene al aplicarle el método del cociente.

### Producto por una constante

Suponga que  $X \sim Cauchy(0, 1)$  y sea  $a > 0$ , entonces  $aX \sim Cauchy(0, a)$ . Si denotamos por  $S_{aX}$  a la región que resulta de aplicar el método del cociente a la distribución  $Cauchy(0, a)$ , entonces en coordenadas cartesianas tenemos que

$$S_{aX} = \left\{ (u, v) \in \mathbb{R}^2 : 0 < u, 0 < |v| < \sqrt{\frac{a}{\pi} - a^2 u^2} \right\}.$$

En la Figura 1.16 se ilustra  $S_{aX}$  para algunos valores de  $a$ . Lo observado en dicha figura es acorde al análisis llevado a cabo en la Sección 1.6, es decir,  $S_{aX}$  es resultado de una expansión-contracción de los ejes coordenados sobre  $S$ .

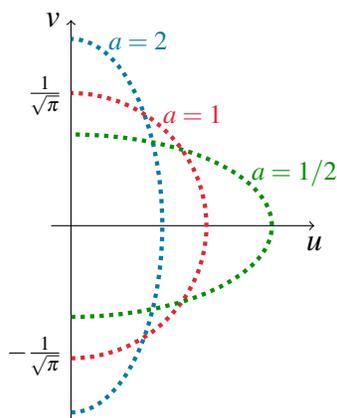


Figura 1.16 . Región  $S_{aX}$  con  $a > 0$ .

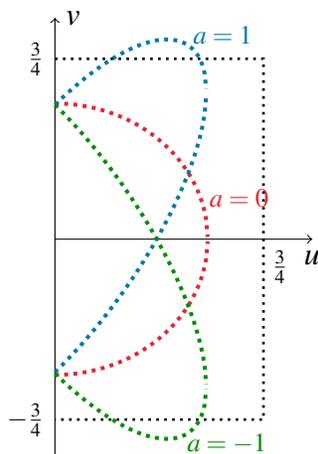
Por lo discutido al final de la Sección 1.6, sabemos que, para  $a < 0$ , la región  $S_{aX}$  es la reflexión con respecto al eje  $u$  de la región  $S_{|a|X}$ , pero debido a que en este caso particular  $S_{|a|X}$  es simétrica respecto al eje  $u$ , ambas regiones coinciden, es decir,  $S_{aX} = S_{|a|X}$ . Esto es totalmente natural, pues por la simetría de la distribución de Cauchy, se puede comprobar a través del teorema de cambio de variable que  $aX$  y  $|a|X$  son variables aleatorias idénticamente distribuidas.

### Suma de una constante

Suponga que  $X \sim \text{Cauchy}(0, 1)$  y sea  $a \in \mathbb{R}$ , entonces  $X + a \sim \text{Cauchy}(a, 1)$ . Denotemos por  $S_{X+a}$  a la región que surge de aplicarle el método del cociente a la distribución  $\text{Cauchy}(a, 1)$ . La descripción en coordenadas cartesianas de dicha región esta dada por

$$S_{X+a} := \left\{ (u, v) \in \mathbb{R}^2 : 0 < u, 0 < u^2 + (v - au)^2 < \frac{1}{\pi} \right\}. \quad (1.19)$$

La expresión que determina a  $S_{X+a}$  en (1.19) parece amigable, pero no es así, o al menos no como la que determina a  $S$  en (1.18). Para que el lector se percate de esto lo invitamos a que despeje  $u$  y  $v$  en la desigualdad presente en (1.19).



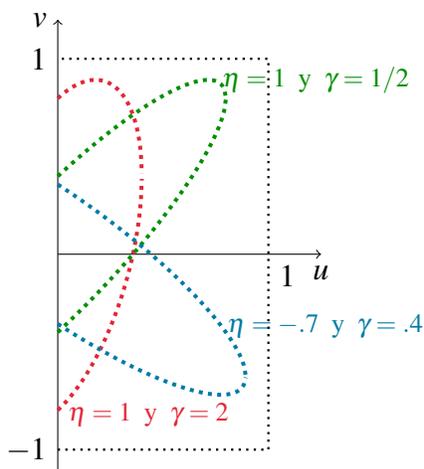
**Figura 1.17 . Región  $S_{X+a}$**

En la Figura 1.17 se ilustra la región  $S_{X+a}$  para algunos valores de  $a$ . En esta figura podemos apreciar que, al variar  $a$ , se obtienen regiones  $S_{X+a}$  que ya no pueden describirse como la región bajo la gráfica de una función. Además, al observar la expresión que determina a  $S_{X+a}$  en (1.19), podemos percatarnos que dicha ecuación oculta la dinámica, bastante simple, que se está llevando a cabo y que describimos en la Sección 1.5. Es decir, que  $S_{X+a}$  es un desplazamiento lineal hacia arriba o hacia abajo de la región  $S$ .

**Observación 1.17**

Si  $X \sim Cauchy(0, 1)$ , entonces  $\gamma X + \eta \sim Cauchy(\eta, \gamma)$ , para  $\gamma > 0$  y  $\eta \in \mathbb{R}$ . De esta manera, la región que resulta de aplicar el método del cociente a la distribución  $Cauchy(\eta, \gamma)$  puede verse como una doble deformación de la región  $S$  correspondiente a la distribución  $Cauchy(0, 1)$ : primero la contracción-expansión del parámetro  $\gamma$  que es multiplicado y después el desplazamiento hacia abajo o arriba del parámetro  $\eta$  que es sumado. Dichas deformaciones deben ser hechas en el orden en que fueron expuestas; ya que, como el lector recordará, la composición de funciones no es conmutativa.

En la Figura 1.18 se muestra la región  $S_{\gamma X + \eta}$  de la distribución de  $Cauchy(\eta, \gamma)$ , para algunos valores particulares de  $\eta$  y  $\gamma$ . En dicha figura podemos confirmar la veracidad de nuestro análisis llevado a cabo.



**Figura 1.18 .** Región  $S_{\gamma X + \eta}$

## Simulación

En la Figura 1.19 se ilustra el histograma de 3000 valores de distribución *Cauchy* estándar generados a partir del método del cociente, y la implementación se llevó a cabo con la ayuda del paquete estadístico R para generar los valores de distribución uniforme sobre la región  $S$ , correspondiente a dicha distribución. En la Figura 1.19 también se encuentra la gráfica de la densidad  $Cauchy(0, 1)$  que, al compararla con el histograma, su afinidad es considerable.

**Advertencia:** Debido a que la distribución de *Cauchy* es de colas pesadas, para que el histograma se ajuste apreciablemente a la gráfica de la función de densidad, es necesario tomar una considerable cantidad de intervalos y de valores de distribución *Cauchy*.

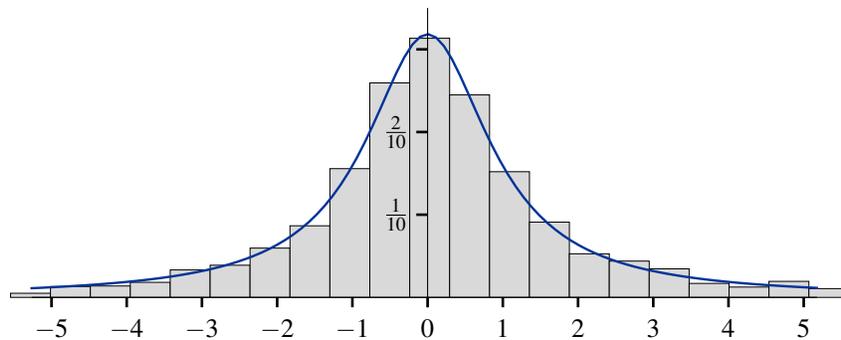


Figura 1.19 . Histograma vs densidad de la distribución  $Cauchy(0, 1)$ .

## 1.9. Distribución normal

Al igual que la distribución de *Cauchy*, la distribución normal  $N(\mu, \sigma^2)$  es absolutamente continua, por lo que cumple con los requerimientos del método del cociente (ver Observación 1.4). Así, como hemos hecho hasta ahora, vamos a aplicar el método del cociente a la distribución normal con los parámetros más simples, es decir, la distribución normal estándar y veremos la región que resulta de ello. Posteriormente analizaremos a las regiones correspondientes a la distribución  $N(\mu, \sigma^2)$  y las compararemos con la que se obtiene de la distribución  $N(0, 1)$ .

## Distribución normal estándar

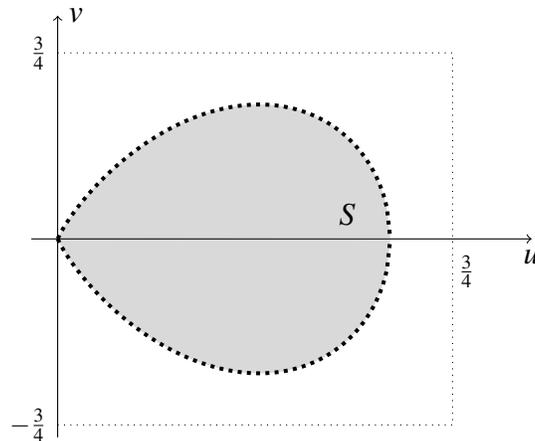
Considere  $h(x)$  la función de densidad de la distribución normal estándar, es decir,  $h(x) = e^{-\frac{x^2}{2}} / \sqrt{2\pi}$   $-\infty < x < \infty$ . La región  $S$  que surge al aplicar el método el cociente a  $h(x)$  esta dada en coordenadas cartesianas por

$$S = \left\{ (u, v) \in \mathbb{R}^2 : 0 < u < \frac{1}{\sqrt[4]{2\pi}}, 0 < v^2 < -2u^2 \ln(u^2 \sqrt{2\pi}) \right\}. \quad (1.20)$$

Mientras que en coordenadas polares dicha región esta dada por

$$S = \left\{ (\theta, r) : \theta \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right), 0 < r < \frac{e^{-\left(\frac{\tan(\theta)}{2}\right)^2}}{\sqrt[4]{2\pi} \cdot \cos(\theta)} \right\}. \quad (1.21)$$

En un primer vistazo a la Figura 1.20 donde se ilustra  $S$ , uno podría pensar que hay ángulos  $\theta \in (-\pi/2, \pi/2)$  para los cuales  $S$  no tiene puntos con dicho ángulo. Sin embargo, por (1.21), sabemos que esto no es verdad, pero la razón de esta ilusión óptica es que la exponencial decrece rápidamente a cero, por lo cual la función que delimita a  $S$  en coordenadas polares también tiende rápidamente a 0 conforme  $\theta$  tiende a  $-\pi/2$  y  $\pi/2$ .



**Figura 1.20 .** Región  $S$  para la distribución normal estándar.

### **Observación 1.18**

*Las distribución normal y de Cauchy son semejantes en el sentido de que ambas tienen un punto de mayor densidad y la densidad de los puntos restantes decrece continuamente a la izquierda y derecha de dicho punto. Es decir, la gráfica de las funciones de densidad de estas distribuciones tienen forma de campana. Pese a esto, la distribución de Cauchy carece de esperanza finita, mientras que la distribución normal sí posee esperanza finita.*

*La distribución normal tiene esperanza finita debido a que su función de densidad decrece rápidamente a la izquierda y derecha del punto de mayor densidad, por lo que, al integrar, es despreciable la contribución de los puntos demasiado alejados del de mayor densidad. En contraste, la densidad de la distribución de Cauchy decrece de forma un tanto más lenta, por lo que, al integrar, no es despreciable la contribución de los puntos alejados del de mayor densidad. Este hecho se traduce de forma gráfica en las regiones que les corresponden bajo el método del cociente, pues la función en coordenadas polares que delimita a la región de la distribución de Cauchy es constante, mientras que la función en coordenadas polares que delimita a la de la distribución normal tiende a 0 conforme  $\theta$  tiende a  $-\frac{\pi}{2}$  y  $\frac{\pi}{2}$ .*

### **Producto por una constante**

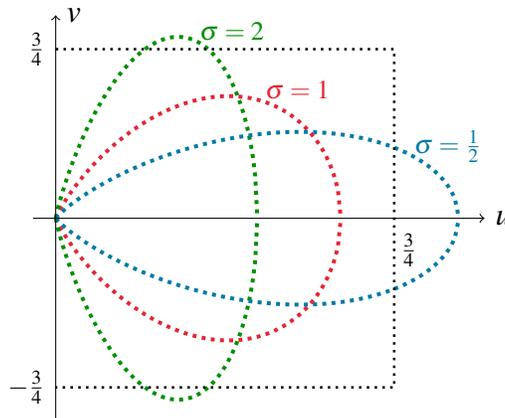
*Suponga que  $X \sim N(0, 1)$  y  $\sigma > 0$ , entonces  $\sigma X \sim N(0, \sigma^2)$ . Denote por  $S_\sigma$  a la región que resulta de aplicar el método del cociente a la distribución  $N(0, \sigma^2)$ . Entonces, de acuerdo al análisis llevado a cabo en la Sección 1.6,  $S_\sigma$  es resultado de aplicar una expansión-contracción a la región  $S$  correspondiente a la distribución  $N(0, 1)$ .*

*En la Sección 1.6 también comprobamos que la descripción cartesiana de  $S_\sigma$  es la que exhibe de forma más clara su relación con  $S$ . Sin embargo, dicha descripción en coordenadas cartesianas es*

$$S_\sigma = \left\{ (u, v) \in \mathbb{R}^2 : 0 < v^2 < -2u^2\sigma^2 \ln(u^2\sqrt{2\pi}) \right\},$$

*de donde difícilmente podríamos deducir que  $S_\sigma$  es una expansión-contracción de  $S$ . Esto exhibe la utilidad del panorama general que nos proporciona la discusión realizada en la Sección 1.6.*

En la Figura 1.21 se encuentra ilustrada  $S_\sigma$  para algunos valores particulares de  $\sigma$ . En esta figura podemos notar que  $S_\sigma$  es simétrica respecto al eje  $u$ . Así, pues, la región correspondiente a la distribución de la variable  $aX$  con  $a < 0$  es exactamente  $S_{-a}$ .



**Figura 1.21 .** Región  $S_\sigma$ .

### Suma de una constante

Nuevamente suponga que  $X \sim N(0, 1)$  y sea  $\mu \in \mathbb{R}$ , entonces  $X + \mu \sim N(\mu, 1)$ . En consecuencia, conforme a la Sección 1.5, si denotamos por  $S_{X+\mu}$  a la región que resulta de aplicar el método del cociente a la distribución  $N(0, 1)$ , entonces  $S_{X+\mu}$  es un desplazamiento lineal hacia arriba o hacia abajo de la región  $S$  correspondiente a la distribución  $N(0, 1)$ .

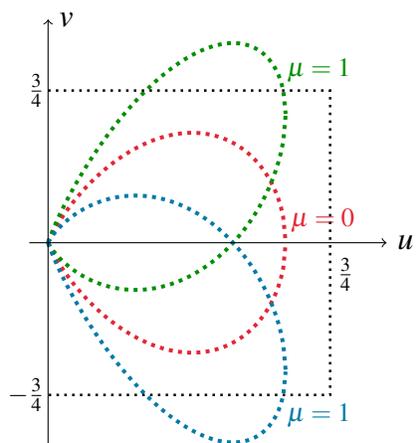
La descripción en coordenadas cartesianas de  $S_{X+\mu}$  es

$$S_{X+\mu} = \left\{ (u, v) \in \mathbb{R}^2 : 0 < \left( \frac{v}{u} - \mu \right)^2 < -2 \ln \left( u^2 \sqrt{2\pi} \right) \right\},$$

de la cual puede deducirse, tras ciertas simplificaciones que, en efecto,  $S_{X+\mu}$  es un desplazamiento lineal de  $S$ .

En la Figura 1.22 se ilustra la región  $S_{X+\mu}$  para algunos valores particulares de  $\mu$ . En esta figura podemos observar que para valores  $\mu \neq 0$ , la región

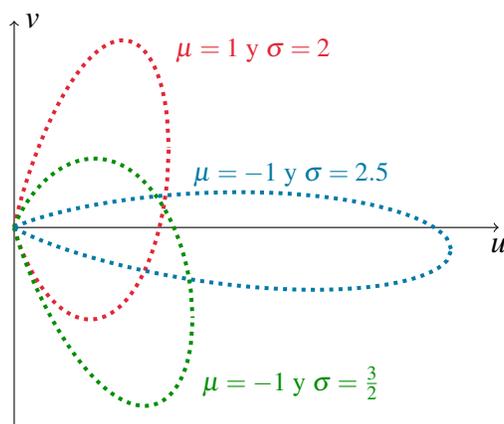
$S_{X+\mu}$  ya no es simétrica respecto al eje  $y$ , debido a que justamente la distribución  $N(\mu, 1)$  es simétrica respecto a cero si y solo si  $\mu = 0$ .



**Figura 1.22 .** Región  $S_{X+\mu}$ .

**Observación 1.19**

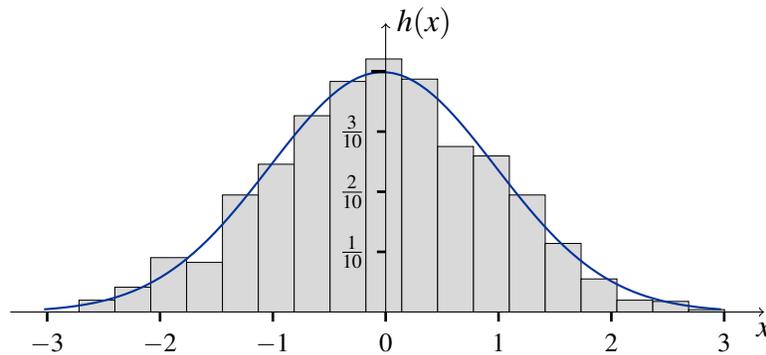
Si  $X \sim N(0, 1)$ , entonces  $\sigma X + \mu \sim N(\mu, \sigma^2)$ , para  $\mu \in \mathbb{R}$  y  $\sigma > 0$ . Por lo tanto, la región  $S_{\sigma X + \mu}$  que resulta de aplicar el método del cociente a  $N(\mu, \sigma^2)$  puede verse como una doble deformación de la región  $S$  correspondiente a la distribución  $N(0, 1)$ , a saber, primero la expansión-contracción correspondiente al parámetro  $\sigma$  que es multiplicado y después el desplazamiento lineal correspondiente al parámetro  $\mu$  que es sumado.



**Figura 1.23 .** Región  $S_{\sigma X + \mu}$ .

## Simulación

En la Figura 1.24 se ilustra el histograma de 1500 valores de distribución  $N(0, 1)$  generados a partir del método del cociente. Al igual que en nuestras secciones anteriores, para la implementación del método, los 1500 valores de distribución uniforme sobre la región  $S$  fueron generados a partir del paquete R.



**Figura 1.24 .** Histograma vs densidad.

## Notas y comentarios

La descripción en coordenadas polares, discutida en la Sección 1.3, esta basada en el libro de Martino, Luengo y Míguez [10] en donde se describe parametricamente a la frontera de la región  $S$  usando a la pendiente como parámetro.

Para generar valores de distribución uniforme sobre la región  $S$  hemos usado el método de aceptación y rechazo sobre un rectángulo que contiene a  $S$  porque generar valores de distribución uniforme sobre un rectángulo es relativamente sencillo, pero en realidad puede ser usada cualquier región que contenga a  $S$  y sobre la cual sea sencillo generar valores de distribución uniforme. Por ejemplo, en [10] se discute brevemente el uso de regiones poligonales en lugar de rectángulos.

## Capítulo 2

### Método general del cociente

*El generar valores de distribuciones multivariadas es un tanto complicado, pues el generar valores de una distribución multivariada requiere generar múltiples valores de distribuciones univariadas, posiblemente dependientes unas de las otras.*

*El método usual para generar valores de distribuciones multivariadas es el de aceptación y rechazo. Por supuesto, en esta situación las distribuciones auxiliares son también multivariadas y, por lo general, serán normales. Sin embargo, el hallar una adecuada distribución normal puede requerir una serie de cálculos que, al implementarse, entorpezcan nuestra simulación.*

*En este capítulo generalizamos el método del cociente a distribuciones multivariadas e introducimos ciertos parámetros que determinan a las regiones resultantes. Dicha generalización se llevará a cabo en tres etapas. La primera extiende el método a distribuciones multivariadas, la segunda incluye un factor  $r$  que determina a la región resultante, y la tercera introduce una función  $g(x)$  que generaliza al factor  $r$ .*

*El presente capítulo tiene como espina dorsal las ideas y el trabajo expuestos por Wakefield, Gelfand y Smith en [13]; Jones y Lunn en [6] y Martino, Luengo y Míguez en [10]. Sin embargo, las ideas presentadas por Marsaglia en [9]; Stefanescu y Vaduva en [12] y Liang, Liu y Carrol en [8], complementan y enriquecen a este capítulo.*

## 2.1. Extensión del cociente de uniformes

En el método del cociente se halla implícitamente una idea exquisita y grandiosa: las distribuciones uniformes en regiones de  $\mathbb{R}^2$  pueden describir elaboradas distribuciones univariadas. Esto responde a la idea general de que al subir una o más dimensiones los problemas se simplifican. Por ello, el extender el método del cociente a varias dimensiones debe generalizar esta idea, es decir, que las distribuciones uniformes en regiones de  $\mathbb{R}^{n+1}$  son capaces de describir elaboradas distribuciones de dimensión  $n$ . Además, como en el método del cociente el medio para lograr tal descripción es un cociente, lo idóneo es que en la extensión del método esta operación también sea la vía por la cual pasemos de  $n + 1$  dimensiones a  $n$  dimensiones.

En lo subsecuente es indispensable el uso de la norma euclidiana, la cual denotaremos por  $\|(x_1, \dots, x_n)\|$ , es decir,  $\|(x_1, \dots, x_n)\|^2 = x_1^2 + \dots + x_n^2$ .

### Proposición 2.1 (Método del cociente en varias variables)

Sea  $h(x_1, \dots, x_n)$  una función de densidad conjunta y suponga que

1.  $h(x_1, \dots, x_n)$  es acotada.
2.  $\sup_{(x_1, \dots, x_n)} \|(x_1, \dots, x_n)\|^{n+1} \sqrt{h(x_1, \dots, x_n)} < \infty$ .

Entonces la región

$$S = \left\{ (u, v_1, \dots, v_n) \in \mathbb{R}^{n+1} : 0 < u < \sqrt[n+1]{h\left(\frac{v_1}{u}, \dots, \frac{v_n}{u}\right)} \right\} \quad (2.1)$$

tiene hipervolumen finito y si  $(U, V_1, \dots, V_n)$  es un vector con distribución uniforme sobre  $S$ , entonces  $(V_1/U, \dots, V_n/U) \sim h$ .

### Demostración.

De la homogeneidad de la norma euclidiana se tiene que, si  $u > 0$ ,

$$\left\| \left( \frac{v_1}{u}, \dots, \frac{v_n}{u} \right) \right\| = \frac{1}{u} \|(v_1, \dots, v_n)\|.$$

Por lo tanto, para todo  $(u, v_1, \dots, v_n) \in S$ , tenemos que

$$0 < \|(v_1, \dots, v_n)\| < \left\| \left( \frac{v_1}{u}, \dots, \frac{v_n}{u} \right) \right\| \sqrt[n+1]{h\left(\frac{v_1}{u}, \dots, \frac{v_n}{u}\right)}.$$

Así pues, concluimos que  $S$  tiene volumen finito. Denotemos por  $|S|$  al volumen de  $S$ . Entonces la función de densidad de  $(U, V_1, \dots, V_n)$  esta dada por

$$f_{U, V_1, \dots, V_n}(u, x_1, \dots, x_n) = \frac{1}{|S|} \quad \text{sobre } S.$$

Por la fórmula para obtener la densidad de cociente de uniformes (ver Apéndice 1), tenemos que, para  $(x_1, \dots, x_n) \in \mathbb{R}^n$ ,

$$\begin{aligned} f_{V_1/U, \dots, V_n/U}(x_1, \dots, x_n) &= \int_{-\infty}^{\infty} |u|^n f_{U, V_1, \dots, V_n}(u, ux_1, \dots, ux_n) du \\ &= \int_0^{n+1\sqrt{h(x_1, \dots, x_n)}} \frac{1}{|S|} |u|^n du \\ &= \frac{1}{(n+1)|S|} \left| n+1\sqrt{h(x_1, \dots, x_n)} \right|^{n+1} \\ &= \frac{h(x_1, \dots, x_n)}{(n+1)|S|}. \end{aligned}$$

Luego, como  $f_{V_1/U, \dots, V_n/U}(y_1, \dots, y_n)$  y  $h(y_1, \dots, y_n)$  son funciones de densidad, al integrar sobre  $\mathbb{R}^n$  deducimos que

$$|S| = \frac{1}{n+1}.$$

Por lo tanto,  $h(x_1, \dots, x_n) = f_{V_1/U, \dots, V_n/U}(x_1, \dots, x_n)$ . □

### Observación 2.1

- (I) El hecho de que  $|S| = \frac{1}{n+1}$  provoca que el volumen de  $S$  tienda a 0 conforme  $n$  tiende a infinito, por lo que es tentadora la idea de que la Proposición 2.1 funciona mejor con valores de  $n$  pequeños. Sin embargo, dicha intuición esta errada, ya que siempre podemos reescalar  $S$  y la región que resulta de ello, al tomar el cociente de las ultimas coordenadas con las primeras, también tiene distribución  $h$ .
- (II) Es evidente que las demostración de las Proposiciones 2.1 y 1.3 son extremadamente idénticas. De hecho, en ambas el teorema de cambio

de variable es la piedra angular de su demostración, y la única diferencia estriba en el difeomorfismo sobre el cual es aplicado el teorema de cambio de variable.

(III) En la prueba de la Proposición 2.1, implícitamente es demostrado que la región  $S$  es acotada, por lo que en particular está contenida en un hiperectángulo. Esto es vital para la viabilidad del método; porque, al hacer uso del Ejemplo 1.2, somos capaces de generar valores de distribución uniforme en la región  $S$ , sin importar cuán irregular o complejo sea  $S$ .

La Proposición 2.1 justifica el siguiente algoritmo para generar valores de distribución (multivariada)  $h(x_1, \dots, x_n)$  cuando las hipótesis de dicha proposición son satisfechas.

- **Algoritmo del cociente de uniformes en varias variables**

1. Generar valores  $(u, v_1, \dots, v_n)$  de distribución uniforme en la región  $S$  definida como en (2.1).
2. Tomar a  $(v_1/u, \dots, v_n/u)$  como valor de distribución  $h$ .

## 2.2. Distribución exponencial bivariada

Si  $X \sim \exp(\lambda_1)$  y  $Y \sim \exp(\lambda_2)$  son variables aleatorias independientes, entonces escribiremos  $(X, Y) \sim \exp(\lambda_1, \lambda_2)$  y diremos que se distribuye exponencial bivariada de parámetros  $\lambda_1$  y  $\lambda_2$ .

La distribución exponencial bivariada puede ser simulada de forma sencilla sin recurrir al método del cociente. Sin embargo, en esta sección aplicaremos el método del cociente a esta distribución, pues, al ser una distribución sencilla, es un ejemplo excelente para familiarizarnos con la extensión a varias dimensiones.

Es evidente que la distribución exponencial bivariada satisface las hipótesis de la Proposición 2.1 para cualquier valor de sus parámetros, pues la densidad conjunta de variables independientes es el producto de sus marginales

y la distribución exponencial satisface los requerimientos del método del cociente.

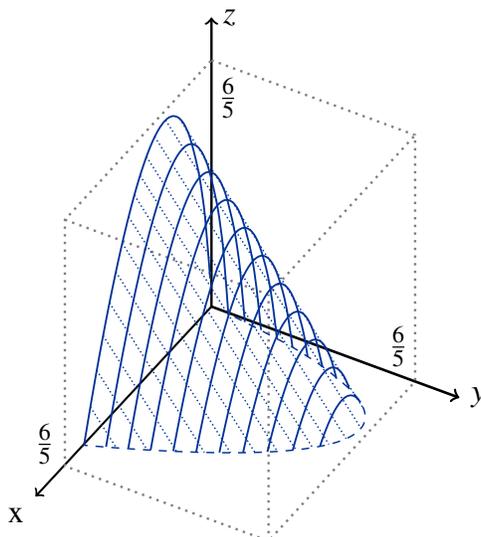
Sea  $h(x,y)$  la función de densidad conjunta de la distribución exponencial bivariada de parámetros  $\lambda_1 = 1 = \lambda_2$ , es decir,

$$h(x,y) = \begin{cases} e^{-(x+y)} & \text{para } x,y > 0, \\ 0 & \text{en otro caso.} \end{cases}$$

Tras hacer un desarrollo similar al realizado para la distribución exponencial en la Sección 1.7, hallamos sin mucho esfuerzo que la región  $S$  que resulta de aplicar el método del cociente a la distribución  $\text{exp}(1, 1)$  esta dada por

$$S = \{(u, v_1, v_2) : 0 < u < 1, 0 < v_1, 0 < v_2, v_1 + v_2 < -3u \ln(u)\}. \quad (2.2)$$

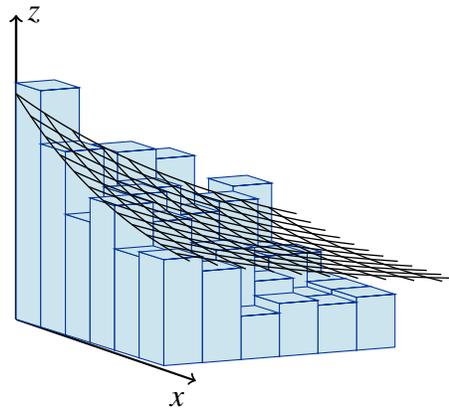
Lo primero que podemos observar en (2.2) es que  $S$  se encuentra en el primer octante. Luego, si en la expresión que describe a  $S$  en (2.2) fijamos el valor de  $v_1$ , entonces se obtiene una expresión para la variable  $v_2$  similar a la presente en (1.16) para la variable  $v$ . De esta manera, la intersección de  $S$  con el plano  $Y = v_1$  es el conjunto vacío o una región en el plano  $Y = v_1$  similar a la región que se encuentra en la Figura 1.16, siempre y cuando se mire perpendicularmente a dicho plano, y lo mismo puede ser dicho de la intersección de  $S$  con el plano  $Z = v_2$  (ver Figura 2.1).



**Figura 2.1** . Región  $S$  para la distribución exponencial bivariada.

## Simulación

En la Figura 2.2 se sobrepone la gráfica de  $h(x,y)$  en el histograma de 1500 valores de distribución  $\text{exp}(1,1)$  generados con el método del cociente. Los valores de distribución uniforme sobre  $S$  se obtuvieron usando el paquete estadístico R para generar valores de distribución uniforme sobre el cubo  $(0,6/5) \times (0,6/5) \times (0,6/5)$  que contiene a la región  $S$ .



**Figura 2.2 .** Histograma vs densidad de la exponencial bivariada.

Es un tanto difícil percibir en la Figura 2.2 cómo el histograma se ajusta a la función de densidad, pues las ilustraciones de  $\mathbb{R}^3$  son proyecciones del espacio en el plano. Por esta razón, invitamos al lector a llevar a cabo por su cuenta la simulación y comprobar que en efecto el histograma se ajusta considerablemente bien a la gráfica de  $h(x,y)$ , aunque se debe tener en cuenta el número de subdivisiones y valores generados.

### 2.3. Distribución de densidad parabólica truncada

Considere la función  $h(x,y)$  dada por

$$h(x,y) = \begin{cases} \frac{3}{2}(x^2 + y^2) & \text{para } x,y \in (0,1), \\ 0 & \text{en otro caso.} \end{cases}$$

Al integrar podemos comprobar fácilmente que  $h(x,y)$  es una función de densidad conjunta. Además,  $h(x,y)$  tiene soporte compacto y es continua en él, por lo que cumple con las hipótesis de la Proposición 2.1.

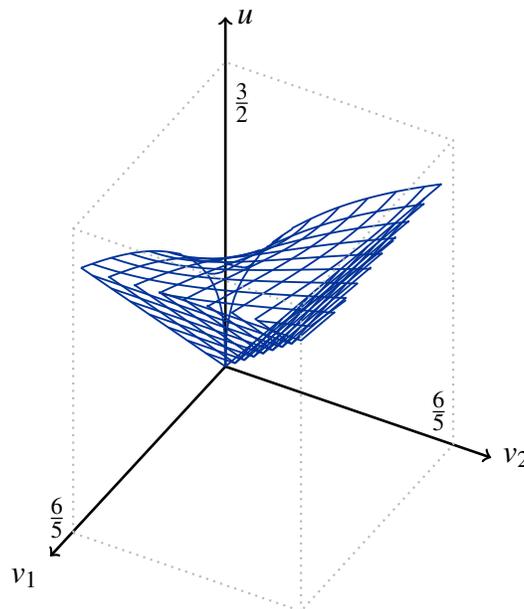
Realizando una sencilla serie de cálculos se puede comprobar que la región  $S$  que resulta de aplicar el método del cociente a la distribución  $h$  esta dada por

$$S = \left\{ (u, v_1, v_2) : \max(v_1, v_2) < u < \sqrt[5]{\frac{3}{2}(v_1^2 + v_2^2)} \right\}. \quad (2.3)$$

Inmediatamente podemos apreciar en (2.3) que  $S$  es la región que se encuentra entre las gráficas de dos funciones sobre las variables  $v_1$  y  $v_2$ . Además, dado que

$$\max(v_1, v_2) = \frac{v_1 + v_2 + |v_1 - v_2|}{2},$$

podemos intuir que la región  $S$  es considerablemente regular. Sin embargo, para dar detalles precisos de  $S$ , es necesario ahondar en las funciones que la definen, pero ello no aporta a nuestro entendimiento del método del cociente, por lo que nos conformamos con ilustrar a  $S$  en la Figura 2.3.



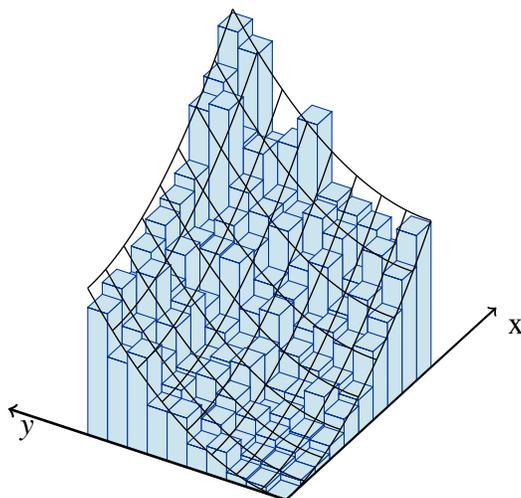
**Figura 2.3** . Región  $S$  para la distribución de densidad parabólica truncada.

### Observación 2.2

Es evidente que en la Figura 2.3 el sistema de coordenadas presenta una orientación distinta de la usual. La razón de esto es que, como  $S$  es la región entre las gráficas de funciones sobre las variables  $v_1$  y  $v_2$ , el usar la orientación de la Figura 2.3 es favorable visual y estéticamente.

## Simulación

Apoyándonos del paquete estadístico R para generar valores de distribución uniforme sobre el rectángulo  $(0, 3/2) \times (0, 3/2) \times (0, 6/5)$  fácilmente podemos simular a la distribución parabólica truncada a partir del método del cociente. En la Figura 2.4 se ilustra el histograma de 2500 valores de distribución parabólica generados con el método del cociente y se compara con la grafica de la función de densidad  $h(x, y)$ .



**Figura 2.4 .** Histograma vs densidad de la distribución parabólica truncada.

### Observación 2.3

La orientación de los ejes en la Figura 2.4 no es la usual, pero como se mencionó antes en otras figuras, el cambio de orientación es para lograr una mayor estética y una mejor interpretación visual por parte del lector.

## 2.4. Método del cociente de factor $r$

Para pasar de  $n + 1$  dimensiones a  $n$  dimensiones podemos usar el cociente  $V_k/U^r$  con  $r \geq 0$ , en lugar de  $V_k/U$ . Esto produce una versión mas general del método del cociente y en la cual  $r$  juega un activo papel determinando a la región resultante.

En esta sección estudiaremos el método del cociente de factor  $r$  (ver [10] y [13]) y describiremos los efectos de este factor  $r$  sobre las regiones que determina la aplicación de dicho método.

### Proposición 2.2 (Método del cociente de factor $r$ )

Sea  $h(x_1, \dots, x_n)$  una función de densidad conjunta y suponga que para  $r \geq 0$  se cumple que

1.  $h(x_1, \dots, x_n)$  es acotada.
2.  $\sup_{(x_1, \dots, x_n)} \|(x_1, \dots, x_n)\| [h(x_1, \dots, x_n)]^{\frac{r}{nr+1}} < \infty$ .

Entonces la región

$$S_r = \left\{ (u, v_1, \dots, v_n) \in \mathbb{R}^{n+1} : 0 < u < \left[ h\left(\frac{v_1}{u^r}, \dots, \frac{v_n}{u^r}\right) \right]^{\frac{1}{nr+1}} \right\} \quad (2.4)$$

tiene hipervolumen finito y si  $(U, V_1, \dots, V_n)$  es un vector de distribución uniforme sobre  $S$ , entonces  $(V_1/U^r, \dots, V_n/U^r) \sim h$ .

### Demostración.

De la homogeneidad de la norma euclidiana se tiene que, si  $u > 0$ ,

$$\left\| \left( \frac{v_1}{u}, \dots, \frac{v_n}{u} \right) \right\| = \frac{1}{u} \|(v_1, \dots, v_n)\|.$$

Por lo tanto, para todo  $(u, v_1, \dots, v_n) \in S_r$ , tenemos que

$$0 < \|(v_1, \dots, v_n)\|^{\frac{1}{r}} < \left\| \left( \frac{v_1}{u^r}, \dots, \frac{v_n}{u^r} \right) \right\|^{\frac{1}{r}} \left[ h\left(\frac{v_1}{u^r}, \dots, \frac{v_n}{u^r}\right) \right]^{\frac{1}{nr+1}}.$$

Así pues, concluimos que  $S_r$  es acotado y, por ende, que es de hipervolumen finito. Denotemos por  $|S_r|$  al volumen de  $S_r$ . Entonces la función de densidad del vector  $(U, V_1, \dots, V_n) \sim \text{unif}(S_r)$  está dada por

$$f_{U, V_1, \dots, V_n}(u, x_1, \dots, x_n) = \frac{1}{|S_r|} \quad \text{sobre } S_r.$$

Por la formula para obtener la densidad de cociente de uniformes (ver Apéndice 1), tenemos que, para  $(x_1, \dots, x_n) \in \mathbb{R}^n$ ,

$$\begin{aligned} f_{V_1/U^r, \dots, V_n/U^r}(x_1, \dots, x_n) &= \int_{-\infty}^{\infty} |u|^{nr} f_{U, V_1, \dots, V_n}(u, u^r x_1, \dots, u^r x_n) du \\ &= \int_0^{[h(x_1, \dots, x_n)]^{\frac{1}{nr+1}}} \frac{1}{|S_r|} |u|^{nr} du \\ &= \frac{1}{(nr+1)|S_r|} \left| h(x_1, \dots, x_n)^{\frac{1}{nr+1}} \right|^{nr+1} \\ &= \frac{h(x_1, \dots, x_n)}{(nr+1)|S_r|}. \end{aligned}$$

Luego, como  $f_{V_1/U^r, \dots, V_n/U^r}(x_1, \dots, x_n)$  y  $h(x_1, \dots, x_n)$  son funciones de densidad, al integrar sobre  $\mathbb{R}^n$  deducimos que

$$|S_r| = \frac{1}{nr+1}. \quad (2.5)$$

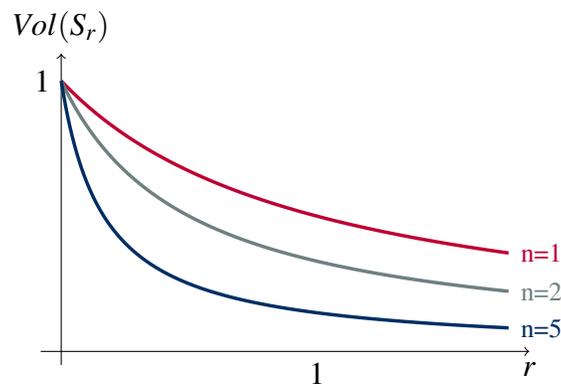
Por lo tanto,  $h(x_1, \dots, x_n) = f_{V_1/U^r, \dots, V_n/U^r}(x_1, \dots, x_n)$ .  $\square$

#### Observación 2.4

En la ecuación (2.5) podemos ver con nitidez que el volumen de  $S_r$  decrece conforme  $r$  tiende a infinito, es decir, que entre más grande es el valor de  $r$ , más pequeño es el volumen de  $S_r$ . Además, debido a que  $r$  esta multiplicando a  $n$ , esta relación entre  $r$  y el volumen de  $S_r$  se acentúa si  $n$  es grande.

Asimismo, en (2.5), podemos apreciar que el volumen de  $S_r$  nunca excederá 1, de hecho, el volumen de  $S_r$  será exactamente igual a 1 únicamente cuando  $r = 0$ . En realidad, en este caso particular, la región  $S_r$  resulta ser la región bajo la gráfica de la función de densidad  $h(x_1, \dots, x_n)$ .

En la Figura 2.5 podemos ver la gráfica del volumen de  $S_r$  en función de  $r$ , para ciertos valores de  $n$ . Dicha figura nos proporciona una buena intuición del ritmo con que decrece el volumen de  $S_r$  con respecto de  $r$ .



**Figura 2.5 .** Volumen de  $S$  en función de  $r$ .

La Proposición 2.2 justifica el siguiente algoritmo para generar valores de distribución (multivariada)  $h(x_1, \dots, x_n)$  cuando las hipótesis de dicha proposición son satisfechas.

• **Algoritmo del cociente de uniformes de factor  $r$**

1. Generar valores  $(u, v_1, \dots, v_n)$  de distribución uniforme en la región  $S_r$  definida como en (2.4).
2. Tomar a  $(v_1/u^r, \dots, v_n/u^r)$  como valor de distribución  $h$ .

**Observación 2.5**

Note que de la Proposición 2.1 se sigue fácilmente de la Proposición 2.2 al tomar  $r = 1$ . Por lo tanto, el método planteado en la Proposición 2.1 es un caso particular del método del cociente de factor  $r$ .

## 2.5. La transformación $T_r$

En la Observación 2.4 se exhibe la evidente relación que tiene  $r$  sobre el volumen de la región  $S$ . Sin embargo, este no es el único efecto que tiene  $r$  sobre  $S_r$  y al plantear a la región descrita en (2.4) como una deformación de la descrita en (2.1), podremos dilucidar el amplio efecto que tiene  $r$  a partir de la transformación que lleva una de las regiones en la otra.

**Definición 2.1** Para cada  $n \geq 1$ , definimos el difeomorfismo  $T_r(u, v_1, \dots, v_n)$  de  $R := \{(u, v_1, \dots, v_n) \in \mathbb{R}^n : u > 0\}$  en sí mismo por

$$T_r(u, v_1, \dots, v_n) = \left( u^{\frac{n+1}{nr+1}}, v_1 \cdot u^{\frac{r-1}{nr+1}}, \dots, v_n \cdot u^{\frac{r-1}{nr+1}} \right).$$

**Proposición 2.3** Suponga que  $h(x_1, \dots, x_n)$  es una función de densidad conjunta que cumple con las hipótesis de la Proposición 2.2 para  $r > 0$  y, en particular, para  $r = 1$ . Si  $S$  y  $S_r$  denotan las regiones que surgen al aplicar el método del cociente de factores 1 y  $r$ , respectivamente, entonces

$$S_r = \{T_r(u, v_1, \dots, v_n) \in \mathbb{R}^{n+1} : (u, v_1, \dots, v_n) \in S_1\} \quad (2.6)$$

es decir,  $S_r = T_r[S]$ .

### **Demostración.**

Basta notar que, para cada  $(u, v_1, \dots, v_n) \in \mathbb{R}^{n+1}$ , se tiene la siguiente cadena de equivalencias

$$\begin{aligned} (u, v_1, \dots, v_n) \in S &\iff 0 < u < \left[ h\left(\frac{v_1}{u}, \dots, \frac{v_n}{u}\right) \right]^{\frac{1}{n+1}} \\ &\iff 0 < u < \left[ h\left(\frac{v_1}{u}, \dots, \frac{v_n}{u}\right) \right]^{\frac{1}{nr+1} \cdot \frac{nr+1}{n+1}} \\ &\iff 0 < u^{\frac{n+1}{nr+1}} < \left[ h\left(\frac{v_1}{u}, \dots, \frac{v_n}{u}\right) \right]^{\frac{1}{nr+1}} \\ &\iff 0 < u^{\frac{n+1}{nr+1}} < \left[ h\left(\frac{v_1 \cdot u^{\frac{r-1}{nr+1}}}{\left(u^{\frac{n+1}{nr+1}}\right)^r}, \dots, \frac{v_n \cdot u^{\frac{r-1}{nr+1}}}{\left(u^{\frac{n+1}{nr+1}}\right)^r}\right) \right]^{\frac{1}{nr+1}} \\ &\iff \left( u^{\frac{n+1}{nr+1}}, v_1 \cdot u^{\frac{r-1}{nr+1}}, \dots, v_n \cdot u^{\frac{r-1}{nr+1}} \right) \in S_r. \end{aligned}$$

□

En virtud del Lema 2.3, nuestra tarea ahora es describir el comportamiento de la función  $T_r(u, v_1, \dots, v_n)$ . Sin embargo, esta descripción no puede llevarse a cabo de la misma manera que para las transformaciones  $\Phi(u, v)$  y  $\Psi(u, v)$  de las Secciones 1.5 y 1.6, pues la función  $T_r(u, v_1, \dots, v_n)$  no es lineal.

## Análisis local de la transformación $T_r$

Una función que es diferenciable puede estudiarse localmente a partir de su derivada y dicha descripción es relativamente sencilla. Por esta razón, estudiaremos localmente  $T_r(u, v_1, \dots, v_n)$  y posteriormente daremos una descripción general con base en la realizada localmente.

Al realizar una sencilla sustitución, podemos comprobar fácilmente que si  $r = 1$ , la función  $T_r(u, v_1, \dots, v_n)$  resulta ser la función identidad de  $R$ . Así que, sin pérdida de generalidad, en lo subsecuente supondremos que  $r \neq 1$ .

### Notación

La matriz jacobiana de la función  $T_r(u, v_1, \dots, v_n)$  es la matriz asociada respecto a las bases canónicas de la derivada global de  $T_r(u, v_1, \dots, v_n)$ , por lo que, en esencia, son objetos distintos. Sin embargo, denotaremos por  $DT_r(u, v_1, \dots, v_n)$  tanto a la derivada global como a la matriz jacobiana de  $T_r(u, v_1, \dots, v_n)$  en el punto  $(u, v_1, \dots, v_n)$ .

Al derivar parcialmente podemos comprobar fácilmente que la matriz jacobiana de la función  $T_r(u, v_1, \dots, v_n)$  se encuentra dada por

$$\begin{bmatrix} \frac{n+1}{nr+1} \cdot u^{\frac{n-nr}{nr+1}} & 0 & \dots & 0 & 0 \\ \frac{r-1}{nr+1} v_1 \cdot u^{\frac{r(1-n)-2}{nr+1}} & u^{\frac{r-1}{nr+1}} & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{r-1}{nr+1} v_{n-1} \cdot u^{\frac{r(1-n)-2}{nr+1}} & 0 & \dots & u^{\frac{r-1}{nr+1}} & 0 \\ \frac{r-1}{nr+1} v_n \cdot u^{\frac{r(1-n)-2}{nr+1}} & 0 & \dots & 0 & u^{\frac{r-1}{nr+1}} \end{bmatrix}. \quad (2.7)$$

Sea  $(u, v_1, \dots, v_n) \in \mathbb{R}^{n+1}$  con  $u > 0$  y denote por  $e_k$  al  $k$ -ésimo elemento de la base canónica de  $\mathbb{R}^{n+1}$ , para toda  $k \in \{1, \dots, n+1\}$ . De la igualdad en (2.7) podemos apreciar que  $DT_r(u, v_1, \dots, v_n)$  (vista como la derivada global) únicamente reescala al vector  $e_k$ , para toda  $k \in \{2, \dots, n+1\}$ , y los reescala a todos ellos por el mismo factor  $u^{\frac{r-1}{nr+1}}$ . Asimismo, de la igualdad en (2.7), podemos apreciar que la imagen de  $e_1$  bajo  $DT_r(u, v_1, \dots, v_n)$  es el vector

$$\left( \frac{n+1}{nr+1} u^{\frac{n-nr}{nr+1}}, \frac{r-1}{nr+1} v_1 \cdot u^{\frac{r(1-n)-2}{nr+1}}, \dots, \frac{r-1}{nr+1} v_n \cdot u^{\frac{r(1-n)-2}{nr+1}} \right). \quad (2.8)$$

Notemos que las últimas entradas de este vector están siendo multiplicadas por un mismo factor. Por ello, es conveniente expresarlo como la combinación lineal

$$\frac{n+1}{nr+1} u^{\frac{n-nr}{nr+1}} (1, 0, \dots, 0) + \frac{r-1}{nr+1} u^{\frac{r(1-n)-2}{nr+1}} (0, v_1, \dots, v_n). \quad (2.9)$$

Esto significa que  $DT_r(u, v_1, \dots, v_n)$  tiene el siguiente efecto sobre el vector  $e_1$ : primero lo reescala por el factor  $\frac{n+1}{nr+1} u^{\frac{n-nr}{nr+1}}$  y luego le suma el vector  $(0, v_1, \dots, v_n)$  reescalado por el factor  $\frac{r-1}{nr+1} u^{\frac{r(1-n)-2}{nr+1}}$ .

A partir de este análisis realizado sobre  $DT_r(u, v_1, \dots, v_n)$ , podemos detallar su comportamiento al variar  $(u, v_1, \dots, v_n)$ . Para ello es suficiente describir la evolución en el comportamiento de  $DT_r(u, v_1, \dots, v_n)$  sobre los vectores  $e_1, \dots, e_{n+1}$ .

El efecto que  $DT_r(u, v_1, \dots, v_n)$  tiene sobre los vectores  $e_2, \dots, e_{n+1}$  aparentemente depende solo de  $u$ . Sin embargo, resulta que  $r$  juega un interesante papel aquí y, de hecho, esto provoca que dividamos nuestro análisis en función del valor de  $r$ .

Supongamos primero que  $r < 1$ , entonces  $u^{\frac{r-1}{nr+1}}$  es decreciente como función de  $u$ . Por lo tanto, en este caso particular, se tiene que:

- $DT_r(u, v_1, \dots, v_n)$  expande los vectores  $e_2, \dots, e_{n+1}$ , si  $u < 1$ .
- $DT_r(u, v_1, \dots, v_n)$  contrae los vectores  $e_2, \dots, e_{n+1}$ , si  $u > 1$ .
- $DT_r(u, v_1, \dots, v_n)$  deja intactos a los vectores  $e_2, \dots, e_{n+1}$ , si  $u = 1$ .

Por otro lado, si  $r > 1$ , entonces  $u^{\frac{r-1}{nr+1}}$  es creciente como función de  $u$ . Así que en este caso se tiene que

- $DT_r(u, v_1, \dots, v_n)$  expande los vectores  $e_2, \dots, e_{n+1}$ , si  $u > 1$ .
- $DT_r(u, v_1, \dots, v_n)$  contrae los vectores  $e_2, \dots, e_{n+1}$ , si  $u < 1$ .
- $DT_r(u, v_1, \dots, v_n)$  deja intactos a los vectores  $e_2, \dots, e_{n+1}$ , si  $u = 1$ .

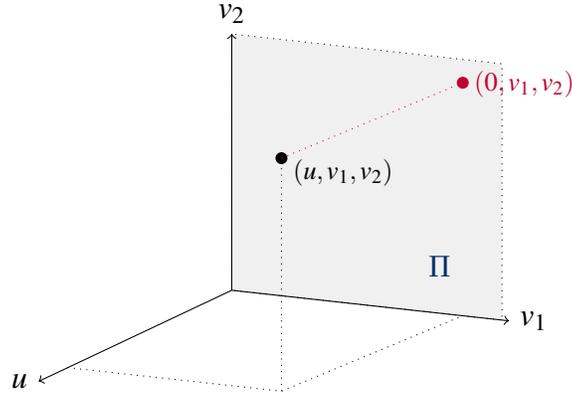
Es oportuno mencionar que en nuestra discusión anterior se manifiesta de forma palpable la influencia que  $r$  ejerce sobre  $T_r(u, v_1, \dots, v_n)$ , pues en ella podemos apreciar que el valor de  $r$  establece el marco general de acción de  $DT_r(u, v_1, \dots, v_n)$ , y la variación de  $(u, v_1, \dots, v_n)$ , particularmente de  $u$ , solo lo atenúa o acentúa.

Muy bien, ahora analicemos el efecto que  $DT_r(u, v_1, \dots, v_n)$  tiene sobre  $e_1$ . Veamos, en (2.8) podemos apreciar que el efecto de  $DT_r(u, v_1, \dots, v_n)$  sobre  $e_1$  depende de  $u, v_1, \dots, v_n$  y  $r$ , haciéndolo considerablemente mas complejo. Sin embargo, al apoyarnos de (2.9), podemos desentrañar la forma en que cada una de estas variables altera el efecto que tiene  $DT_r(u, v_1, \dots, v_n)$  sobre el vector  $e_1$ .

Comencemos por analizar el vector  $(0, v_1, \dots, v_n)$  presente en la combinación lineal de (2.9). Lo primero que puede ser dicho acerca de este vector es que se encuentra confinado al hiperplano

$$\Pi := \{(0, y_1, \dots, y_n) \in \mathbb{R}^{n+1} : y_1, \dots, y_n \in \mathbb{R}\}. \quad (2.10)$$

De hecho, el vector  $(0, v_1, \dots, v_n)$  es la proyección del vector  $(u, v_1, \dots, v_n)$  sobre dicho hiperplano. En la Figura 2.6 se ilustra la proyección de un punto  $(u, v_1, \dots, v_n)$  sobre  $\Pi$  cuando  $n = 2$ , es decir, cuando  $(u, v_1, \dots, v_n)$  es un punto en el espacio euclidiano y  $\Pi$  es el plano  $V_1V_2$ .



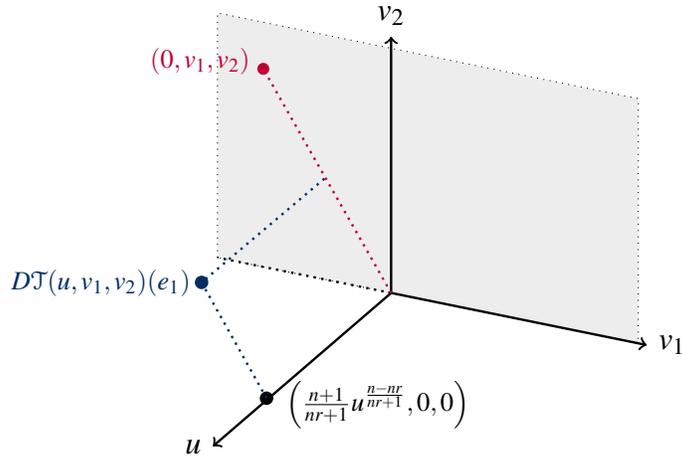
**Figura 2.6 .** Proyección de  $(u, v_1, \dots, v_n)$  sobre  $\Pi$  cuando  $n = 2$ .

Es evidente que el sentido del vector  $\frac{r-1}{nr+1}u^{\frac{r(1-n)-2}{nr+1}}(0, v_1, \dots, v_n)$  está totalmente determinado por el valor de  $r$ . Si  $r > 1$ , entonces el sentido del vector coincidirá con el de  $(0, v_1, \dots, v_n)$ , pero si  $r < 1$ , entonces el sentido del vector será contrario al de  $(0, v_1, \dots, v_n)$ .

Asimismo, independientemente del valor de  $n$  y  $r$ ,  $u^{\frac{r(1-n)-2}{nr+1}}$  es decreciente como función de  $u$ . Así que, entre más grande sea el valor de  $u$ , más pequeña será la norma del vector  $\frac{r-1}{nr+1}u^{\frac{r(1-n)-2}{nr+1}}(0, v_1, \dots, v_n)$  y, en consecuencia, su contribución será menor en la combinación lineal de (2.9). Por lo tanto, para valores pequeños de  $u$ , el efecto de  $DT_r(u, v_1, \dots, v_n)$  sobre  $e_1$  es similar a un reescalamiento.

En cuanto al vector  $\frac{n+1}{nr+1}u^{\frac{n-nr}{nr+1}}(1, 0, \dots, 0)$  podemos observar que es un reescalamiento positivo del vector  $e_1$  y la magnitud de dicho reescalamiento esta en función de  $u$ ; creciente, si  $r < 1$  y decreciente, si  $r > 1$ .

Con base en lo expuesto, concluimos que  $DT_r(u, v_1, \dots, v_n)$  primero reescala positivamente  $e_1$  en función de  $u$ ; creciente, si  $r < 1$  y decreciente, si  $r > 1$ . Luego, lo desvía con una suma vectorial en dirección de su componente en el hiperplano  $\Pi$ , conservando el mismo sentido si  $r > 1$  y cambiándolo si  $r < 1$ . Este hecho se encuentra ilustrado en la Figura 2.7 para el caso particular de  $n = 2$  y  $r > 1$ .



**Figura 2.7 .** La imagen de  $e_1$  bajo  $DT_r(u, v_1, \dots, v_n)$ .

A partir de la descripción de  $DT_r(u, v_1, \dots, v_n)$  al variar  $(u, v_1, \dots, v_n)$  se sigue de inmediato la descripción general de  $T_r(u, v_1, \dots, v_n)$ , pues infinitesimalmente la función  $T_r(u, v_1, \dots, v_n)$  alrededor del punto  $(u, v_1, \dots, v_n)$  se comporta como  $DT_r(u, v_1, \dots, v_n)$  alrededor del origen.

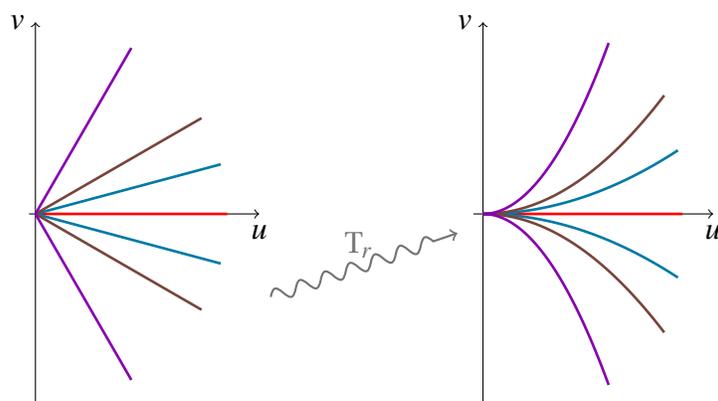
En el siguiente par de ejemplos se analiza a la función  $T_r(u, v_1, \dots, v_n)$  cuando ésta actúa sobre el plano y sobre el espacio para determinados valores de  $n$  y  $r$ .

**Ejemplo 2.1 (La función  $T_r$  sobre el plano)**

Bajo los valores particulares de  $n = 1$  y  $r = 2$  la transformación  $T_r$  es en dos variables y esta dada por

$$T_r(u, v) = (u^{\frac{2}{3}}, v \cdot u^{\frac{1}{3}}). \tag{2.11}$$

Así que,  $T_r(u, v)$  es en cierta medida una función simple y, por ende, podemos proceder de diversas formas para entenderla. Por ejemplo, en la Figura 2.8 se ilustra la imagen bajo  $T_r(u, v)$  de algunos segmentos de recta que parten del origen, lo cual describe enteramente su comportamiento. Esto supone una ventaja para nuestros fines, ya que al conocer tan bien a la función en (2.11) reconoceremos los alcances del análisis general que realizamos sobre la transformación  $T_r$ .



**Figura 2.8 .** La función  $T_r$  sobre el plano al tomar  $n = 1$  y  $r = 2$ .

Veamos, al tomar  $n = 1$  el hiperplano  $\Pi$  de (2.10) se degenera en el eje  $Y$ . Por lo tanto, como  $r > 1$ , se tiene  $DT_r(u, v)$  tras rescalar al vector  $(1, 0)$  lo desplaza linealmente hacia arriba o hacia abajo dependiendo de si  $v > 0$  o  $v < 0$ , respectivamente. Además, debido a que la magnitud del desplazamiento es inversamente proporcional a la magnitud de  $u$ , el desplazamiento se acentúa para valores pequeños de  $u$ . Por otro lado, dado que  $r > 1$ ,  $DT_r(u, v)$  reescala al vector  $(0, 1)$  por un factor mayor a uno si es que  $u > 1$  y por un factor menor a 1 si por el contrario  $u < 1$ .

Por otro lado, en virtud de que  $r > 1$ , resulta que  $DT_r(u, v)$  reescala al vector  $(0, 1)$  por un factor mayor a 1 cuando  $u > 1$  y por un factor menor a 1 cuando  $u < 1$ , es decir,  $DT_r(u, v)$  contraerá a  $(0, 1)$  cuando  $u < 1$  y lo expandirá cuando  $u > 1$ . De esta manera, podemos concluir el siguiente comportamiento de  $T_r(u, v)$

- En el cuadrante 1 el efecto de la función  $T_r(u, v)$  es principalmente el de desplazar hacia arriba, pero este desplazamiento se ve beneficiado o mermado de la contracción o expansión (local) que realiza en dirección del eje  $V$ .
- La parte positiva del eje  $U$ ,  $T_r(u, v)$  no presenta desplazamiento local ni hacia arriba ni hacia abajo del eje  $U$ , por lo que, alrededor de estos puntos,  $T_r(u, v)$  reescalará localmente a los ejes coordenados.

- En el cuadrante 4 el efecto de la función  $T_r(u, v)$  es en su mayoría un desplazamiento hacia abajo, pero este desplazamiento se ve acentuado o atenuado por la contracción o expansión (local) que realiza localmente.

De esta manera, se tiene un entendimiento claro de  $T_r(u, v)$  y podemos comprobar que los efectos de la función  $T_r(u, v)$  descritos en el análisis anterior concuerdan con los ilustrados en la Figura 2.8.  $\circ$

### Ejemplo 2.2 (La función $T_r$ sobre el espacio)

Al tomar los valores particulares de  $n = 2$  y  $r = 1/2$ , la función  $T_r$  esta dada por

$$T_r(u, v_1, v_2) = (u^{\frac{3}{2}}, v_1 \cdot u^{-\frac{1}{4}}, v_2 \cdot u^{-\frac{1}{4}}).$$

La función  $T_r(u, v_1, v_2)$  es en cierta medida simple, pues cada una de sus funciones coordenadas es un polinomio o un cociente de polinomios. Sin embargo, el hecho de que  $T_r(u, v_1, v_2)$  sea una función sobre  $\mathbb{R}^3$  conlleva cierta dificultad adicional, la cual yace fundamentalmente en que las gráficas sobre  $\mathbb{R}^3$  nos proporcionan una intuición sesgada al ser proyecciones del espacio sobre el plano. Por lo tanto, el detallado análisis local de  $T_r(u, v_1, v_2)$  acrecienta su viabilidad en esta situación.

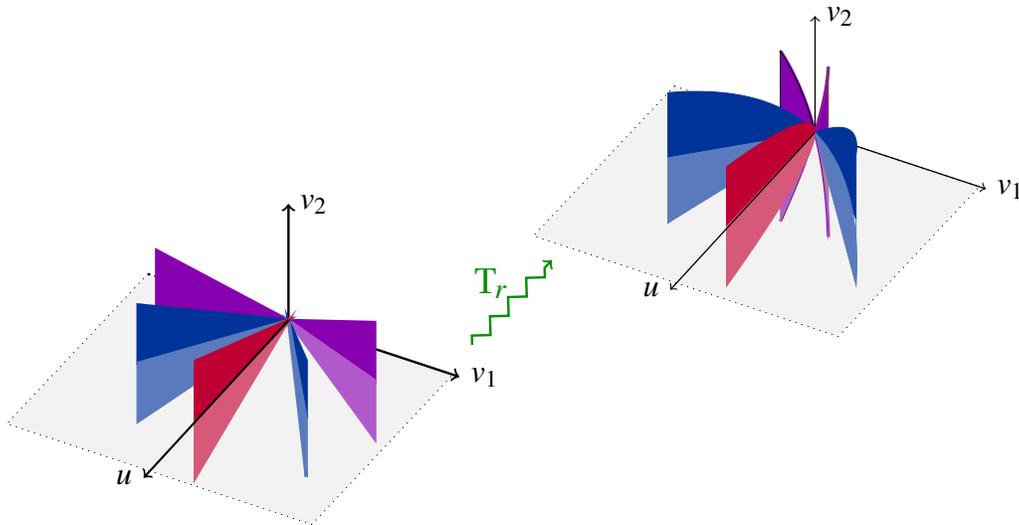
Veamos, la derivada  $T_r(u, v_1, v_2)$  reescala al vector  $e_1$  por un factor positivo y luego le suma un múltiplo escalar de  $(0, v_1, v_2)$  cuya magnitud es inversamente proporcional al valor de  $u$ . Sin embargo, debido a que  $r < 1$ , el múltiplo escalar de  $(0, v_1, v_2)$  que es sumado tiene orientación contraria a la de  $(0, v_1, v_2)$  y la magnitud del reescalamiento del vector  $e_1$  es creciente en función de  $u$ .

Ahora, como  $r < 1$ , la derivada  $DT_r(u, v_1, v_2)$  reescala a los vectores  $e_2$  y  $e_3$  por un factor que es decreciente en función de  $u$ . De esta manera, concluimos que el comportamiento de  $T_r(u, v_1, v_2)$  es el siguiente.

- En el primer y quinto octante el efecto de la función  $T_r(u, v_1, v_2)$  es principalmente un desplazamiento hacia el octavo y cuarto octante. Además, de forma local el desplazamiento es en dirección contraria  $(0, v_1, v_2)$ , por lo que el desplazamiento se acentúa al alejarse del plano  $UV_2$ , aunque también conforme  $u$  tiende a infinito.

- En el octavo y cuarto octante el efecto de la función  $T_r(u, v_1, v_2)$  es un desplazamiento hacia el primer y quinto octante. Además, sabemos que localmente el desplazamiento es en dirección contraria a  $(0, v_1, v_2)$ , por lo que el desplazamiento se acentúa al alejarse del plano  $UV_2$ , aunque también conforme  $u$  tiende a infinito.
- Sobre la parte del plano  $UV_2$  que se encuentra entre los octantes 1, 5, 4 y 8, la función  $T_r(u, v_1, v_1)$  reescala localmente a los ejes  $V_1$  y  $V_2$ . La magnitud de dicho desplazamiento es decreciente en función de  $u$ .

En la Figura 2.9 se muestran las imágenes bajo  $T_r(u, v_1, v_2)$  de algunas secciones triangulares perpendiculares al plano  $UV_1$ . En dicha figura podemos apreciar cómo las secciones triangulares en los octantes 4 y 8 son "doblad" hacia los octantes 1 y 5, mientras que las secciones triangulares en los cuadrantes 1 y 5 son dobladas hacia los octantes 4 y 8. También podemos observar cómo la sección que se encuentra contenida en el plano  $UV_2$  no es sacada de dicho plano y solo presenta una pequeña deformación en dirección del eje  $V_2$ .



**Figura 2.9** . La función  $T_r$  sobre el espacio al tomar  $n = 2$  y  $r = \frac{1}{2}$  .

○

## 2.6. Optimización del valor de $r$

En el Capítulo 1 vimos que, en general, al aplicar el método del cociente se obtienen regiones considerablemente regulares. A pesar de ello, puede ser desafiante el generar valores uniformemente distribuidos sobre estas región sin recurrir al método de aceptación y rechazo. Sin embargo, el método de aceptación y rechazo tiene una desventaja que puede tener un considerable efecto si deseamos generar demasiados valores. La cuestión es que, usando el método de aceptación y rechazo, algunos de los valores que son generados no son aceptados, lo cual entorpece el proceso cuando demasiados valores son rechazados. Pero ¿cual es la probabilidad de que un valor sea rechazado?

Sean  $A$  y  $B$  regiones en el plano con volumen distinto de cero y finito. Si  $A \subseteq B$  y se genera un valor  $x$  distribución uniforme sobre  $B$ , entonces  $x$  es aceptado como valor de distribución uniforme sobre  $A$  solo si  $x \in A$ . Por lo tanto, los valores que son rechazados son todos aquellos valores que son generados en  $B \setminus A$  y en consecuencia la probabilidad de que un valor sea rechazado es igual a

$$\frac{\text{vol}(B) - \text{vol}(A)}{\text{vol}(B)}.$$

De esta manera, para que no haya demasiados valores rechazados, es necesario que la diferencia entre el volumen de  $A$  y  $B$  sea pequeña, entre más pequeña mejor. Por lo tanto, diremos que el valor mas óptimo de  $r$  es con el que se tiene la menor diferencia entre los volúmenes de  $S_r$  y el rectángulo más pequeño que la contiene. Así mismo, informal e intuitivamente, diremos que un valor de  $r$  es óptimo si la diferencia entre los volúmenes de  $S_r$  y el rectángulo mínimo que la contiene es pequeña.

### Optimización de $r$ a partir de su efecto en la transformación $T_r$

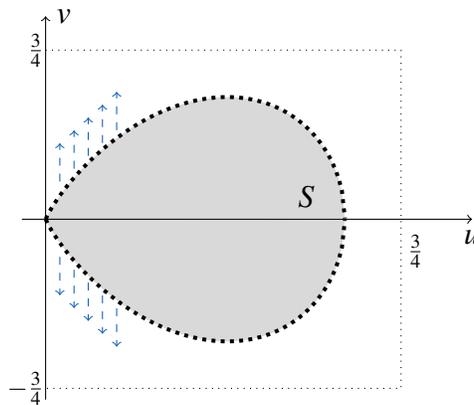
A partir del análisis realizado en la Sección 2.5 podemos proponer valores considerablemente óptimos de  $r$ , pues en dicha sección se estudian los efectos de  $r$  sobre la transformación  $T_r(u, v_1, \dots, v_n)$ , que deforma a  $S_1$  en  $S_r$ .

En el Ejemplo 2.3 se abordará nuevamente a la distribución normal estándar, pero esta vez para ilustrar la manera en que el análisis de la transformación  $T_r$  debe ser empleado para la elección de valores óptimos de  $r$ .

**Ejemplo 2.3 (Distribución normal)**

En la Sección 1.20 aplicamos el método del cociente a la distribución normal estándar. Al hacerlo, se obtuvo una región parecida a un pétalo y la cual se ilustra una vez mas en la Figura 2.10. Si a los puntos de  $S$  los expandiéramos verticalmente hacia arriba y abajo como lo sugieren las flechas de la Figura 2.10, entonces la región resultante se acoplaría mejor al mínimo rectángulo que la contiene. Por lo tanto, los valores de  $r$  a tomar en cuenta son aquellos con los cuales la transformación  $T_r(u, v)$  realice esta expansión.

Veamos, al tomar los valores  $n = 1$  y  $r < 1$ , la transformación  $T_r$  principalmente desplaza verticalmente a los puntos, hacia arriba a los que se encuentren en el cuadrante 1 y hacia abajo a los que se encuentren en el cuadrante 4. Pero también desplaza a todos los puntos horizontalmente, a la izquierda si están a la izquierda de la recta  $x = 1$  y a la derecha si se encuentran a la derecha de la recta  $x = 1$ . El desplazamiento vertical se acentúa en los puntos cercanos al eje  $V$ , mientras que el desplazamiento horizontal se atenúa al alejarse de la recta  $x = 1$ .



**Figura 2.10 .** Región  $S$  para la distribución normal estándar.

Por lo tanto, encontramos buenos prospectos para el factor  $r$  en el intervalo

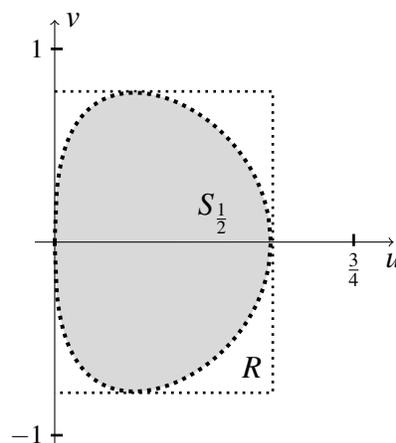
$(0, 1)$  y como el comportamiento de  $T_r(u, v)$  descrito arriba es más preciso para valores no tan pequeños de  $r$ , un excelente propuesta es aplicar el método del cociente con  $r = 1/2$  a la distribución normal.

Al aplicar el método del cociente de factor  $r = \frac{1}{2}$  a la distribución normal estándar se obtiene la región  $S_{\frac{1}{2}}$  dada por

$$S_{\frac{1}{2}} = \left\{ (u, v) : 0 < u < \frac{1}{\sqrt[3]{2\pi}}, 0 < v^2 < -2u \ln(u^{\frac{3}{2}} \sqrt{2\pi}) \right\}.$$

En la Figura 2.11 se encuentra ilustrada  $S_{\frac{1}{2}}$ . En esta figura también se trazan las líneas del rectángulo  $R$  más pequeño que contiene  $S_{\frac{1}{2}}$  y con lados paralelos a los ejes. Con una precisión de 3 decimales, la diferencia en el área de  $R$  y  $S_{\frac{1}{2}}$  es 0.1722 la cual, al ser dividida por el área de  $R$  nos da un aproximado de 0.20465, es decir, al generar un valor uniformemente distribuido sobre  $R$ , se tiene un 79.53% de probabilidad de que el valor generado se encuentre en  $S_{\frac{1}{2}}$ .

En contraste, al realizar los cálculos pertinentes, es fácil comprobar que al generar un valor de distribución uniforme en el rectángulo más pequeño que contiene a  $S$  (también con lados paralelos a los ejes), la probabilidad de que el punto esté en  $S$  es únicamente del 73.05%, lo cual supone un descenso en la probabilidad de aceptación de alrededor de 6%, respecto a  $r = \frac{1}{2}$ .



**Figura 2.11** . Región  $S_r$  para la distribución normal estándar con  $r = \frac{1}{2}$ .

### Observación 2.6

Como podemos observar en el Ejemplo 2.11, el análisis de la transformación  $T_r$  nos permite hacer una considerablemente buena elección, aunque intuitiva, de valores de  $r$  para las cuales la región  $S_r$  se acopla bastante bien a su rectángulo mínimo. Sin embargo, para hacer esta elección intuitiva de  $r$  es necesario conocer de antemano a la región  $S$ , lo cual entorpece esta técnica empírica para hallar valores óptimos de  $r$ .

En el ejemplo 2.3 ha sido fructífero el análisis que se ha realizado acerca del factor  $r$ . Sin embargo, debido a los múltiples efectos que tiene  $r$  sobre la transformación  $\mathcal{T}(u, v)$ , si no se consideran los suficientes efectos, entonces la elección de  $r$  en base a este análisis puede fallar y llevarnos a tomar valores de  $r$  poco óptimos.

### Ejemplo 2.4 (Distribución Cauchy)

En la Sección 1.8 vimos que la región  $S$  correspondiente a la distribución Cauchy (estándar) es el semicírculo derecho de radio  $\frac{1}{\sqrt{\pi}}$ . A partir de esta región y del análisis del factor  $r$ , propondremos un valor óptimo de  $r$  tal como se hizo en el Ejemplo 2.3, pero esta vez seremos laxos considerando los efectos de  $r$  en la transformación  $T_r$ .

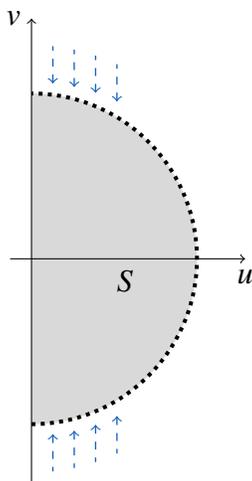


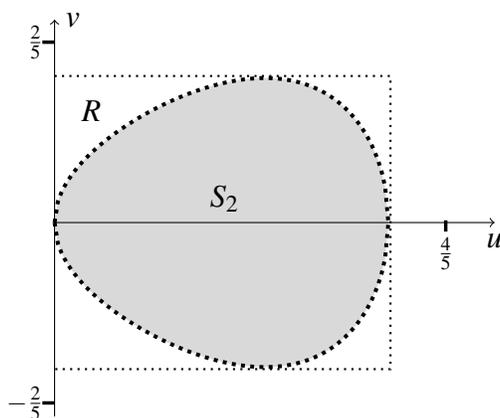
Figura 2.12 . Región  $S$  para la distribución Cauchy(0,1).

Veamos, en nuestro análisis del factor  $r$  descubrimos que para valores  $r > 1$  la transformación  $T_r(u, v)$  contrae localmente el eje  $Y$  si  $u < 1$  y este efecto se va reduciendo en intensidad conforme  $u$  se acerca a 1 que es el caso límite en el cual localmente  $T_r(u, v)$  no expande ni contrae el eje  $Y$ . Por lo tanto, de lo anterior, se deduce que la función  $T_r(u, v)$  con  $r = 2$  tiene el efecto sugerido en la Figura 2.12 sobre  $S$ , es decir, que va a contraer verticalmente a los puntos cercanos al eje  $Y$  y que esta contracción va ser menos intensa en los puntos lejanos del eje  $Y$ . De esta manera, parece ser que una adecuada elección de  $r$  es 2, pero lamentablemente en el fortuito análisis realizado no se considera la magnitud de la contracción la cual es considerablemente grande, lo que provoca que la región  $S_2$  no presente las características deseadas, es decir, que no se acople mejor que  $S$  al rectángulo mínimo que lo contiene.

La probabilidad de aceptación de  $S$  y  $S_2$  en los respectivos rectángulos mínimos que los contienen es de 78.54% y 76.18%, por lo que la diferencia no es grande, pero establece a  $r = 2$  como poco óptimo.

En la Figura 2.13 se ilustra a la región  $S_2$ , la cual se encuentra dada por

$$S_2 = \left\{ (u, v) : 0 < u < \frac{1}{\sqrt[3]{\pi}}, 0 < |v| < \sqrt{\frac{u}{\pi} - u^4} \right\}. \quad (2.12)$$



**Figura 2.13 .** Región  $S_2$  para la distribución  $Cauchy(0, 1)$ .

Tanto en (2.12) como en la Figura 2.13 se puede apreciar la formidable magnitud de la contracción local cerca del eje  $Y$ , responsable de que nuestra propuesta de  $r = 2$  no resulte ser óptima como suponíamos.  $\circ$

## Optimización analítica de $r$

El Ejemplo 2.4 muestra que la elección de  $r$  a partir de sus efectos en  $T_r$  depende del grado de minuciosidad del análisis y que para un análisis limitado, dicha elección puede resultar bastante mala. Por ello, en pro de evitar la incertidumbre, podemos intentar hallar valores óptimos de  $r$  de forma analítica y la dificultad de esta tarea dependerá de la distribución con la que se esté trabajando, pero el plan de acción es el mismo: calcular la probabilidad de aceptación de la región  $S_r$  en el mínimo rectángulo (o hiperectángulo) que lo contiene y minimizarla como función de  $r$ .

### Ejemplo 2.5

Denote por  $S_r$  a la región que resulta de aplicar el método general del cociente con  $r > 0$  a la distribución normal estándar. Independientemente del valor de  $r$ , el volumen de  $S_r$  es exactamente igual a  $\frac{1}{r+1}$  (ver Observación 2.4), por lo que concentrémonos en calcular el volumen del rectángulo mínimo que contiene a  $S_r$ .

Tras una serie de sencillos cálculos, puede comprobarse que la región  $S_r$  se encuentra dada por

$$S_r = \left\{ (u, v) : 0 < u < (2\pi)^{-\frac{1}{2(r+1)}}, 0 < |v| < u^r \sqrt{-2\ln(u^{r+1}\sqrt{2\pi})} \right\}.$$

De esta manera, el rectángulo  $\mathcal{R}_r = (0, a_r) \times (-b_r, b_r)$  es el más pequeño que contiene a  $S_r$ , donde

$$a_r := (2\pi)^{-\frac{1}{2(r+1)}} \quad \text{y} \quad b_r := \sup_u u^r \sqrt{-2\ln(u^{r+1}\sqrt{2\pi})}.$$

Ahora, notemos que  $b_r$  puede calcularse fácilmente, pues es el supremo de una función diferenciable en un intervalo. Por lo tanto, aplicando el criterio de la primer y segunda derivada para calcular su supremo, obtenemos que

$$b_r := e^{-\frac{1}{2}} \cdot (2\pi)^{-\frac{r}{2(r+1)}} \cdot \sqrt{\frac{r+1}{r}}.$$

En consecuencia, el volumen de  $\mathcal{R}_r$  se encuentra dado por

$$\text{Vol}(\mathcal{R}_r) = 2a_r b_r = 2\sqrt{\frac{r+1}{2r\pi e}}. \quad (2.13)$$

Si denotamos por  $p_A(r)$  a la probabilidad de aceptación de  $S_r$  en  $\mathcal{R}_r$ , entonces

$$p_A(r) = \frac{\frac{1}{r+1}}{\frac{2\sqrt{(r+1)}}{\sqrt{2r\pi e}}} = \frac{\sqrt{2r\pi e}}{2(r+1)^{\frac{3}{2}}}.$$

De esta manera, como  $p_A(r)$  es una función diferenciable, al aplicar el criterio de la primera y segunda derivada para maximizar a  $p_A(r)$ , se tiene que  $r = \frac{1}{2}$  maximiza a la función  $p_A(r)$ .  $\circ$

Optimizar analíticamente el valor de  $r$  es lo ideal, ya que al hacerlo obtenemos la certeza de que el valor  $r$  que elegimos es el óptimo. Sin embargo, en algunos casos, obtener analíticamente el valor de  $r$  puede llegar a ser complicado. Por esta razón, es recomendable que primero se intente hallar analíticamente el valor óptimo de  $r$ , pero si esta tarea es demasiado laboriosa, puede darse propuestas bastante buenas de  $r$  usando el análisis de la transformación  $T_r$ .

## 2.7. Método general del cociente

En la demostración de la Proposición 2.2 resulta fundamental el teorema de cambio de variable. Sin embargo, existe un segundo ingrediente que permite dicha demostración, pero que se presenta de manera tan sutil que requiere de cierto esfuerzo para ser consciente de él. Este sutil ingrediente nos proporciona el método general del cociente.

El método general del cociente presentado en la Proposición 2.4 es sugerida en su forma unidimensional en el trabajo de Wakefield en [13]. Dicha sugerencia es seguida en el extraordinario trabajo de Martino, Luengo y Miguez en [10], estableciendo la Proposición 2.4 en su versión unidimensional. Por lo tanto, el trabajo realizado en la Proposición 2.4 se reduce a una simple extensión a varias dimensiones del método general del cociente.

**Proposición 2.4 (Método general del cociente)**

Sean  $h(x_1, \dots, x_n)$  una función de densidad y  $g(x)$  una función dos veces diferenciable y estrictamente creciente definida sobre un intervalo  $[0, a)$  (o  $[0, a]$ ), para alguna  $a > 0$ , tal que  $g(0) = 0$ . Además, suponga que  $c > 0$  es tal que, para cada  $(x_1, \dots, x_n) \in \mathbb{R}^n$ ,

$$c \cdot h(x_1, \dots, x_n) < \sup_x g(x). \quad (2.14)$$

Si la región

$$S_g := \left\{ (u, v_1, \dots, v_n) : 0 < u < g^{-1} \left( c \cdot h \left( \frac{v_1}{\sqrt[n]{g'(u)}}, \dots, \frac{v_n}{\sqrt[n]{g'(u)}} \right) \right) \right\} \quad (2.15)$$

tiene hipervolumen finito y  $(U, V_1, \dots, V_n)$  es un vector con distribución uniforme sobre  $S_g$ , entonces  $(V_1/\sqrt[n]{g'(U)}, \dots, V_n/\sqrt[n]{g'(U)}) \sim h$ .

**Demostración.**

Si  $(U, V_1, \dots, V_n)$  es un vector con distribución uniforme sobre  $S_g$ , entonces su función de densidad  $\rho(u, v_1, \dots, v_n)$  está dada por

$$\rho(u, v_1, \dots, v_n) = \frac{1}{\text{vol}(S_g)} \mathbb{1}_{S_g}(u, v_1, \dots, v_n).$$

Notemos que la función

$$H(y, x_1, \dots, x_n) = \left( y, \frac{x_1}{\sqrt[n]{g'(y)}}, \dots, \frac{x_n}{\sqrt[n]{g'(y)}} \right),$$

definida sobre  $[0, a) \times \mathbb{R}^n$ , es un difeomorfismo y que su función inversa esta dada por

$$H^{-1}(y, x_1, \dots, x_n) = (y, \sqrt[n]{g'(y)}x_1, \dots, \sqrt[n]{g'(y)}x_n).$$

Ahora, tras derivar parcialmente, podemos comprobar con facilidad que la matriz jacobiana de  $H^{-1}(y, x_1, \dots, x_n)$  se encuentra dada por

$$\begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ \frac{1}{n}x_1g''(y)(g'(y))^{\frac{1-n}{n}} & \sqrt[n]{g'(y)} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{1}{n}x_{n-1}g''(y)(g'(y))^{\frac{1-n}{n}} & 0 & \cdots & \sqrt[n]{g'(y)} & 0 \\ \frac{1}{n}x_n g''(y)(g'(y))^{\frac{1-n}{n}} & 0 & \cdots & 0 & \sqrt[n]{g'(y)} \end{bmatrix}.$$

De esta manera, se tiene que  $|\det(DH^{-1}(y, x_1, \dots, x_n))| = g'(y)$  y, por el teorema de cambio de variable, concluimos que la función de densidad  $q(y, x_1, \dots, x_n)$  del vector aleatorio  $(U, V_1/\sqrt[n]{g'(U)}, \dots, V_n/\sqrt[n]{g'(U)})$  está dada por

$$q(y, x_1, \dots, x_n) = \frac{g'(y)}{\text{vol}(S_g)} \mathbb{1}_{S_g}(y, x_1 \sqrt[n]{g'(y)}, \dots, x_n \sqrt[n]{g'(y)}).$$

Por lo tanto, la función de densidad del vector  $(V_1/\sqrt[n]{g'(U)}, \dots, V_n/\sqrt[n]{g'(U)})$  se encuentra dada por

$$\begin{aligned} f(x_1, \dots, x_n) &= \int_{-\infty}^{\infty} \frac{g'(y)}{\text{vol}(S_g)} \mathbb{1}_{S_g}(y, x_1 \sqrt[n]{g'(y)}, \dots, x_n \sqrt[n]{g'(y)}) dy \\ &= \frac{1}{\text{vol}(S_g)} \int_0^{g^{-1}(c \cdot h(x_1, \dots, x_n))} g'(y) dy \\ &= \frac{1}{\text{vol}(S_g)} g(y) \Big|_0^{g^{-1}(c \cdot h(x_1, \dots, x_n))} \\ &= \frac{c}{\text{vol}(S_g)} h(x_1, \dots, x_n). \end{aligned}$$

Finalmente, debido a que  $f(x_1, \dots, x_n)$  y  $h(x_1, \dots, x_n)$  son funciones de densidad, al integrar sobre  $\mathbb{R}^n$  obtenemos que  $\text{vol}(S_g) = c$ . Así, pues, concluimos que  $f(y_1, \dots, y_n) = h(y_1, \dots, y_n)$ .  $\square$

### Observación 2.7

La hipótesis expresada en (2.14) no se usa explícitamente en la demostración de la Proposición 2.4. Sin embargo, dicha hipótesis es necesaria para

que  $c \cdot h(x_1, \dots, x_n)$  este en el dominio de  $g^{-1}(y)$  para cualquier  $(x_1, \dots, x_n) \in \mathbb{R}^n$  y así la expresión que define a  $S_g$  no sea ambigua.

El método general del cociente en la Proposición 2.4 presenta una diferencia considerable con las versiones del método del cociente que hemos estudiado previamente. Esta diferencia radica en que en las versiones previas del método del cociente no teníamos que preocuparnos de la finitud del volumen de las regiones involucradas, mientras que en el método general del cociente es necesario verificar esta condición para poder hacer uso del método. Esta diferencia se debe a que en las versiones previas del método del cociente hemos establecido condiciones que aseguran las regiones sean acotadas y por ende de volumen finito.

En la siguiente proposición se imponen ciertas condiciones suplementarias a las hipótesis del método general del cociente. Estas condiciones son suficientes para que la región  $S_g$  sea acotada y, por ende, de hipervolumen finito.

**Proposición 2.5 (Versión débil del método general del cociente)**

Sean  $h(x_1, \dots, x_n)$  una función de densidad acotada y  $g(x)$  una función diferenciable definida sobre un intervalo  $[0, a)$ , para alguna  $a > 0$ . Si para una constante  $c > 0$  se satisface que

1.  $g(x)$  estrictamente creciente y convexa.
2.  $g(0) = 0$ .
3.  $c \cdot h(x_1, \dots, x_n) < \sup_x g(x)$  para cada  $(x_1, \dots, x_n) \in \mathbb{R}^n$ .
4.  $\sup_{(x_1, \dots, x_n)} \|(x_1, \dots, x_n)\| \sqrt[n]{g'(g^{-1}(c \cdot h(x_1, \dots, x_n)))} < \infty$ .

Entonces la región

$$S_g := \left\{ (u, v_1, \dots, v_n) : 0 < u < g^{-1} \left( c \cdot h \left( \frac{v_1}{\sqrt[n]{g'(u)}}, \dots, \frac{v_n}{\sqrt[n]{g'(u)}} \right) \right) \right\}$$

es acotada y si  $(U, V_1, \dots, V_n)$  es un vector con distribución uniforme sobre  $S_g$ , entonces  $(V_1/\sqrt[n]{g'(U)}, \dots, V_n/\sqrt[n]{g'(U)}) \sim h$ .

**Demostración.**

Por la Proposición 2.4 solo resta probar que la región  $S_g$  es acotada y es precisamente lo que haremos.

Suponga que  $(u, v_1, \dots, v_n) \in S_g$  y defina

$$x_k := \frac{v_k}{\sqrt[n]{g'(u)}} \quad (1 \leq k \leq n). \quad (2.16)$$

De esta manera, se tiene que

$$\|(v_1, \dots, v_n)\| = \sqrt[n]{g'(u)} \|(x_1, \dots, x_n)\|. \quad (2.17)$$

Ahora, debido a que  $g(x)$  es convexa, se tiene que  $g'(x)$  es creciente y, por lo tanto,

$$g'(0) \leq g'(u) \leq g' \left( g^{-1} \left( c \cdot h \left( \frac{v_1}{\sqrt[n]{g'(u)}}, \dots, \frac{v_n}{\sqrt[n]{g'(u)}} \right) \right) \right).$$

Así, por (2.16) y (2.17), se concluye que

$$\|(v_1, \dots, v_n)\| \leq \|(x_1, \dots, x_n)\| \sqrt[n]{g'(g^{-1}(c \cdot h(x_1, \dots, x_n)))}. \quad (2.18)$$

Asimismo, por (2.16), se tiene que

$$0 < u < g^{-1}(c \cdot h(x_1, \dots, x_n)). \quad (2.19)$$

De las condiciones 1 y 2 de la Proposición 2.4 así, como de la arbitrariedad con que elegimos  $(u, v_1, \dots, v_n)$  en  $S_g$ , se sigue inmediatamente de (2.18) y (2.19) que  $S_g$  es acotada.  $\square$

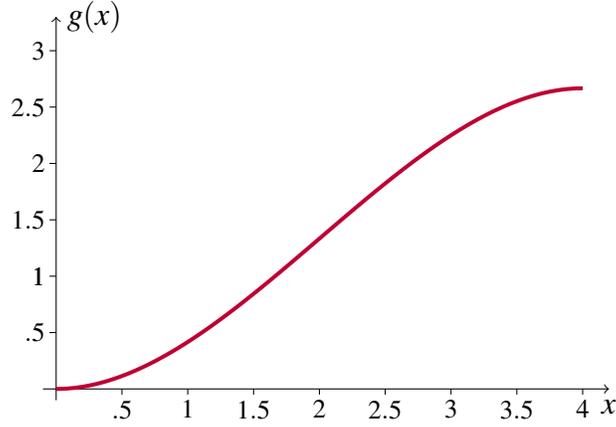
*Las condiciones suplementarias de la Proposición 2.5, aunque suficientes, no son necesarias para que la región  $S_g$  sea acotada, y esto es expuesto en el Ejemplo 2.6.*

**Ejemplo 2.6 (Condiciones suficientes, pero no necesarias)**

Defina a la función  $g(x)$  en el intervalo  $[0, 4)$  por

$$g(x) := x - \frac{(x-2)^3}{12} - \frac{2}{3}. \quad (2.20)$$

Al derivar se puede verificar sin esfuerzo que la función  $g(x)$  es convexa en  $[0, 2)$  y cóncava en  $(2, 4)$ , siendo 2 el punto de inflexión de la función. Este comportamiento también puede ser apreciado en la gráfica de la función, la cual se encuentra ilustrada en la Figura 2.14.



**Figura 2.14 .** Gráfica de la función  $g(x)$ .

Sea  $h(x)$  la función de densidad de la distribución exponencial de parámetro  $\lambda = 1$ . Es fácil comprobar que la función  $g(x)$  definida en (2.20) satisface las hipótesis de la Proposición 2.4 para la constante  $c = 1$  y la densidad  $h(x)$ .

Aprovechando la monotonía de  $g(x)$  y que el soporte de  $h(x)$  es  $(0, \infty)$ , se tiene la siguiente cadena de equivalencias.

$$\begin{aligned}
 0 < u < g^{-1} \left( h \left( \frac{v}{g'(u)} \right) \right) &\iff 0 < g(u) < h \left( \frac{v}{g'(u)} \right) \\
 &\iff 0 < v, 0 < g(u) < e^{-\frac{v}{g'(u)}} \\
 &\iff 0 < v, 0 < u < g^{-1}(1), \ln(g(u)) < -\frac{v}{g'(u)} \\
 &\iff 0 < u < g^{-1}(1), 0 < v < -g'(u) \ln(g(u)).
 \end{aligned}$$

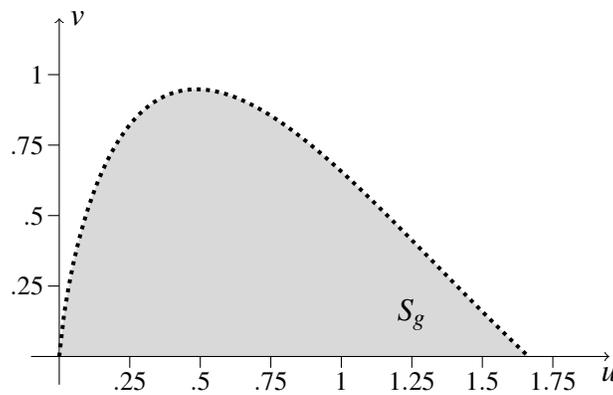
Por lo tanto, la región  $S_g$  para  $h(x)$  esta dada por

$$S_g = \{(u, v) : 0 < u < g^{-1}(1), 0 < v < -g'(u) \ln(g(u))\}. \quad (2.21)$$

Por ende,  $S_g$  es la región bajo la gráfica de la función  $-g'(x)\ln(g(x))$  en el intervalo  $(0, g^{-1}(1))$ , y dicha función es acotada, por lo que  $S_g$  es una región acotada.

De esta manera, a pesar de que  $g(x)$  no cumple las hipótesis de la Proposición 2.5, pues no es convexa, se tiene que la región  $S_g$  de la distribución exponencial es acotada. Como resultado, las condiciones suplementarias de la Proposición 2.5 son suficientes, mas no necesarias, para que la región resultante  $S_g$  sea acotada.

En la Figura 2.15 se ilustra la región  $S_g$  para la distribución exponencial, en donde se puede apreciar una extrema semejanza con la región  $S$  de la Figura 1.11.



**Figura 2.15 .** Región  $S_g$  de la distribución exponencial.

Para finalizar con nuestro ejemplo, note que la expresión en (2.21) para la región  $S_g$  de la distribución exponencial es válida para cualquier  $g(x)$ , pues para llegar a dicha expresión solo se necesitó de la monotonía de la función  $g(x)$ . ○

*La Proposición 2.4 justifica el siguiente algoritmo para generar valores de distribución  $h$  cuando las hipótesis de dicha proposición son satisfechas por  $h(x_1, \dots, x_n)$  y  $g(x)$ .*

• **Algoritmo general del cociente de uniformes**

1. Generar valores  $(u, v_1, \dots, v_n)$  de distribución uniforme en la región  $S_g$  definida como en (2.4).
2. Tomar a  $\left(v_1/\sqrt[n]{g'(u)}, \dots, v_n/\sqrt[n]{g'(u)}\right)$  como valor de distribución  $h$ .

Las versiones del método del cociente estudiadas en las secciones anteriores son un caso particular de la versión débil del método general del cociente, y dicha particularidad radica en la elección de la función  $g(x)$ , lo cual se evidencia en la observación siguiente.

**Observación 2.8**

Consideremos la función  $g(x) = \frac{x^{nr+1}}{nr+1}$  definida sobre  $[0, \infty)$ . Tras una serie de sencillos cálculos podemos comprobar fácilmente que  $g'(x) = x^{nr}$  y  $g^{-1}(x) = [x(nr+1)]^{\frac{1}{nr+1}}$ . De esta manera, para una función de densidad  $h(x_1, \dots, x_n)$  se tiene la siguiente serie de igualdades para  $c = \frac{1}{nr+1}$

$$\begin{aligned} \sqrt[n]{g'(g^{-1}(c \cdot h(x_1, \dots, x_n)))} &= \sqrt[n]{(g^{-1}(c \cdot h(x_1, \dots, x_n)))^{nr}} \\ &= [(nr+1) \cdot c \cdot h(x_1, \dots, x_n)]^{\frac{r}{nr+1}} \\ &= [h(x_1, \dots, x_n)]^{\frac{r}{nr+1}}. \end{aligned}$$

Por lo tanto, en este caso particular, la condición 4 de la Proposición 2.4 se torna en

$$\sup_{(x_1, \dots, x_n)} \|(x_1, \dots, x_n)\| [h(x_1, \dots, x_n)]^{\frac{r}{nr+1}} < \infty. \quad (2.22)$$

De esta manera, al tomar  $g(x) = \frac{x^{nr+1}}{nr+1}$  y  $c = \frac{1}{nr+1}$  en la Proposición 2.5, las condiciones impuestas a  $h(x_1, \dots, x_n)$  son exactamente las que se le imponen en la Proposición 2.2. Pero esta no es la única similitud, pues al realizar la substituciones pertinentes podemos comprobar que

$$g^{-1} \left( c \cdot h \left( \frac{v_1}{\sqrt[n]{g'(u)}}, \dots, \frac{v_n}{\sqrt[n]{g'(u)}} \right) \right) = \left[ h \left( \frac{v_1}{u^r}, \dots, \frac{v_n}{u^r} \right) \right]^{\frac{1}{nr+1}}.$$

Por lo tanto, en este caso particular,  $S_g$  y  $S_r$  coinciden, es decir,  $S_g = S_r$ . En consecuencia, el método del cociente de factor  $r$  al igual que sus variantes son un caso particular del método general del cociente.

### **Observación 2.9**

A partir de la Observación 2.8 podemos apreciar que nuestras versiones previas del método del cociente condicionan a la función de densidad para que las regiones resultantes sean acotadas. Esto se debe a que es en este contexto cuando el método del cociente resulta realmente útil, ya que es en esta situación en donde podemos simular fácilmente, a través del método de aceptación y rechazo, a la distribución uniforme sobre estas regiones.

Las versiones fuertes de nuestras versiones previas del método del cociente se siguen fácilmente de tomar la función  $g(x) = \frac{x^{nr+1}}{nr+1}$  y  $c = \frac{1}{nr+1}$  de la Observación 2.8 y con ella particularizar a la Proposición 2.4.

### **Observación 2.10**

A partir de la Proposición 2.4 es evidente que el segundo ingrediente fundamental del método del cociente es la función  $g(x)$ , la cual pasa un tanto desapercibida en versiones más débiles del método. Sin embargo, dicha función es de gran relevancia, ya que determina activamente a la región del método (ver (2.15)).

## **2.8. Descripción en coordenadas polares**

La idea clave en la descripción en coordenadas polares de la Sección 1.3 subyace en el hecho de que

$$\arctan\left(\frac{v}{u}\right) = \text{ángulo del vector } (u, v).$$

Por lo que, es natural que una descripción similar para  $S_g$  (en el caso univariable) se obtenga del hecho de que

$$\arctan\left(\frac{v}{g'(u)}\right) = \text{ángulo del vector } (g'(u), v).$$

Veamos, considere la función  $\mathcal{T}(u, v) = (g'(u), v)$  y volvamos a considerar las funciones  $\theta(u, v)$  y  $\rho(u, v)$  que a cada vector le asignan su ángulo y norma respectivamente. Entonces se tiene que

$$\frac{v}{g'(u)} = \tan(\theta(\mathcal{T}(u, v))) \quad y \quad g'(u) = \rho(\mathcal{T}(u, v)) \cos(\theta(\mathcal{T}(u, v))).$$

De esta manera, al suponer que  $g'(x)$  es estrictamente creciente, se obtiene que

$$0 < u < g^{-1}\left(ch\left(\frac{v}{g'(u)}\right)\right) \Leftrightarrow 0 < \rho(z) < \frac{g'(g^{-1}(ch(\tan(\theta(z)))))}{\cos(\theta(z))},$$

donde  $z := \mathcal{T}(u, v)$ . Por lo tanto, se concluye que

$$S_g = \left\{ \mathcal{T}^{-1}(r, \theta) : \theta \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right), 0 < r < \frac{g'(g^{-1}(c \cdot h(\tan(\theta))))}{\cos(\theta)} \right\}$$

En consecuencia, si definimos

$$\mathcal{B} := \left\{ (r, \theta) : \theta \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right), 0 < r < \frac{g'(g^{-1}(c \cdot h(\tan(\theta))))}{\cos(\theta)} \right\},$$

entonces la región  $S_g$  es la imagen directa de  $\mathcal{B}$  bajo  $\mathcal{T}^{-1}(u, v)$ , es decir,

$$S_g = \mathcal{T}^{-1}[\mathcal{B}]. \quad (2.23)$$

La descripción para  $S_g$  en (2.23) no está dada en coordenadas polares, pero debido a que es el resultado de la extensión natural del análisis de la Sección 1.3, parece natural referirnos a la expresión en (2.23) como la descripción en coordenadas polares de la región  $S_g$  del método general del cociente.

### Observación 2.11

La expresión (2.23) hallada en la Sección 2.8, aunque compacta, no es tan simple como lo aparenta por dos razones:

1. La función  $\mathcal{T}^{-1}(u, v)$  esta dada a través de la función  $g'(x)$ .
2. El conjunto  $\mathcal{B}$  depende de  $g'(x)$  y  $g^{-1}(x)$ .

De esta manera, al hallar una función  $g(x)$  para la cual  $\mathcal{T}^{-1}(u, v)$  se simplifique, puede que la región  $\mathcal{B}$  se torne complicada, e inversamente, al hallar una función  $g(x)$  para la cual  $\mathcal{B}$  se simplifique, puede que la función  $\mathcal{T}^{-1}(u, v)$  se vuelva complicada. Por lo tanto, el simplificar a la región  $S_g$  a través de la expresión en (2.23) es una guerra en dos frentes, volviéndola una guerra difícil de ganar.

**Ejemplo 2.7 (Descripción polar para la función  $g(x) = \frac{x^2}{2}$  y  $c = \frac{1}{2}$ )**

Sea  $h(x)$  una función de densidad y, por conveniencia, suponga que satisface las hipótesis de la Proposición 1.3. Al aplicar el método general del cociente a  $h(x)$  con  $g(x) = x^2/2$  y  $c = 1/2$ , se obtiene la región del método del cociente de la Proposición 1.3, por lo que la descripción en (2.23) de esta región debe coincidir con la obtenida en la Sección 1.3. Comprobemos lo.

Veamos, debido a que  $g^{-1}(x) = \sqrt{2x}$  y  $g'(x) = x$ , se tiene que

$$\frac{g'(g^{-1}(c \cdot h(\tan(\theta))))}{\cos(\theta)} = \frac{\sqrt{h(\tan(\theta))}}{\cos(\theta)}. \quad (2.24)$$

Además, debido a que  $\mathcal{T}(u, v) = (g'(u), v)$ , la función  $\mathcal{T}^{-1}(u, v)$  es la función identidad y, por lo tanto, a partir de (2.24), se concluye que en este caso particular

$$\mathcal{T}^{-1}[\mathcal{B}] = \left\{ (r \cos(\theta), r \sin(\theta)) : \theta \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right), 0 < r < \frac{\sqrt{h(\tan(\theta))}}{\cos(\theta)} \right\}.$$

○

**Ejemplo 2.8 (Descripción polar para la función  $g(x) = e^x - 1$  y  $c = 1$ )**

Sea  $h(x)$  una función de densidad para la cual la función  $g(x) := e^x - 1$  y la constante  $c = 1$  satisfacen las hipótesis de la Proposición 2.4.

Al realizar una serie de sencillos cálculos, los cuales se dejan al lector, puede comprobarse que

$$g'(x) = e^x \quad \text{y} \quad g^{-1}(x) = \ln(1 + x).$$

Por lo tanto, se tiene

$$\frac{g'(g^{-1}(h(\tan(\theta))))}{\cos(\theta)} = \frac{1+h(\tan(\theta))}{\cos(\theta)} \quad \text{y} \quad \mathcal{T}^{-1}(u, v) = (\ln(u), v).$$

De esta manera, en este caso particular, se concluye que la descripción polar de la región  $S_g$  correspondiente a la distribución  $h$  esta dada por

$$S_g = \left\{ (\ln(r \cos(\theta)), r \sin(\theta)) : \theta \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right), 0 < r < \frac{1+h(\tan(\theta))}{\cos(\theta)} \right\}.$$

○

### Observación 2.12

*Es importante puntualizar que la igualdad en (2.23) requirió de la monotonía estricta de  $g'(x)$ . Así que la descripción en coordenadas polares de la región  $S_g$  del método general del cociente es valida solo cuando dicho método es aplicado con una función  $g(x)$  con derivada estrictamente creciente, o equivalentemente, con una función  $g(x)$  estrictamente convexa.*

*Ahora, como uno de las hipótesis de la Proposición 2.5 es que la función  $g(x)$  sea convexa, la versión débil del método general del cociente es una excelente alternativa si se desean obtener regiones que puedan ser expresadas como en (2.15).*

## 2.9. La transformación $T_g$

*En la Sección 2.5 fue interesante estudiar cómo la región  $S$  era deformada en la región  $S_r$ , por lo que es natural hacer lo mismo para la región  $S_g$ , es decir, estudiar la manera en que la región  $S$  es deformada en la región  $S_g$ .*

**Definición 2.2** *Sean  $h(x_1, \dots, x_n)$  una función de densidad y  $g(x)$  una función definida sobre  $[0, a)$ . Si las hipótesis de la Proposición 2.4 son satisfechas por  $h(x_1, \dots, x_n)$  y  $g(x)$  para la constante  $c > 0$ , entonces*

1. *Definimos sobre el intervalo  $\left[0, \sqrt[n+1]{a/c}\right)$  a la función*

$$\tau(u) = g^{-1}(cu^{n+1}).$$

2. Definimos sobre  $\left[0, \sqrt[n]{a/c}\right) \times \mathbb{R}^n$  a la función

$$T_g(u, v_1, \dots, v_n) = \left( \tau(u), \frac{v_1 \cdot \sqrt[n]{g'(\tau(u))}}{u}, \dots, \frac{v_n \cdot \sqrt[n+1]{g'(\tau(u))}}{u} \right).$$

**Proposición 2.6** *Suponga que  $h(x_1, \dots, x_n)$  es una función de densidad conjunta que satisface las hipótesis de la Proposición 2.4 para la función  $g(x)$  y la constante  $c > 0$ . Entonces*

$$S_g = \left\{ T_g(u, v_1, \dots, v_n) : (u, v_1, \dots, v_n) \in S \right\},$$

donde

$$S = \left\{ (u, v_1, \dots, v_n) : 0 < u < \sqrt[n+1]{h\left(\frac{v_1}{u}, \dots, \frac{v_n}{u}\right)} \right\}.$$

**Demostración.**

Notemos que, para cada  $(u, v_1, \dots, v_n) \in \mathbb{R}^{n+1}$ , se tiene la siguiente cadena de equivalencias

$$\begin{aligned} (u, v_1, \dots, v_n) \in S &\iff 0 < u < \left[ h\left(\frac{v_1}{u}, \dots, \frac{v_n}{u}\right) \right]^{\frac{1}{n+1}} \\ &\iff 0 < u^{n+1} < h\left(\frac{v_1}{u}, \dots, \frac{v_n}{u}\right) \\ &\iff 0 < cu^{n+1} < c \cdot h\left(\frac{v_1}{u}, \dots, \frac{v_n}{u}\right) \\ &\iff 0 < g^{-1}(cu^{n+1}) < g^{-1}\left(h\left(\frac{v_1}{u}, \dots, \frac{v_n}{u}\right)\right). \end{aligned}$$

Finalmente, debido a que para toda  $k \in \{1, \dots, n\}$

$$\frac{v_k}{u} = \frac{v_k \cdot \sqrt[n]{g'(g^{-1}(c \cdot u^{n+1}))}}{u \cdot \sqrt[n]{g'(g^{-1}(c \cdot u^{n+1}))}},$$

concluimos que  $(u, v_1, \dots, v_n) \in S$  si y solo si

$$\left( g^{-1}(cu^{n+1}), \frac{v_1 \cdot \sqrt[n]{g'(g^{-1}(cu^{n+1}))}}{u}, \dots, \frac{v_n \cdot \sqrt[n]{g'(g^{-1}(cu^{n+1}))}}{u} \right) \in S_g.$$

De donde se sigue inmediatamente la demostración.  $\square$

Por lo tanto, la función que deforma a  $S$  en  $S_g$  es  $T_g$ , la cual depende activamente de la función  $g$ , lo que la hace complicada de estudiar si no se especifica a la función  $g(x)$ .

**Observación 2.13**

Retomemos la función  $g(x) = \frac{x^{nr+1}}{nr+1}$  definida sobre  $[0, \infty)$  y  $c = \frac{1}{nr+1}$ . Entonces  $g'(x) = x^{nr}$  y  $g^{-1}(x) = [(nr+1)x]^{\frac{1}{nr+1}}$ . Así, en este caso particular, se tiene que

$$\begin{aligned} g^{-1}(cu^{n+1}) &= [(nr+1)cu^{n+1}]^{\frac{1}{nr+1}} \\ &= u^{\frac{n+1}{nr+1}}. \end{aligned}$$

Asimismo, tenemos que

$$\begin{aligned} \sqrt[n]{g'(g^{-1}(cu^{n+1}))} &= \sqrt[n]{(g^{-1}(cu^{n+1}))^{nr}} \\ &= [(nr+1)cu^{n+1}]^{\frac{r}{nr+1}} \\ &= u^{\frac{r(n+1)}{nr+1}} \end{aligned}$$

De esta manera,

$$\begin{aligned} \frac{v_k \cdot \sqrt[n]{g'(g^{-1}(cu^{n+1}))}}{u} &= v_k \cdot u^{\frac{n(r+1)}{nr+1}-1} \\ &= v_k \cdot u^{\frac{r-1}{nr+1}}. \end{aligned}$$

Por lo tanto, en este caso particular, tenemos que

$$T_g(u, v_1, \dots, v_n) = (u^{\frac{n+1}{nr+1}}, v_1 \cdot u^{\frac{r-1}{nr+1}}, \dots, v_n \cdot u^{\frac{r-1}{nr+1}}).$$

Por ende, al retomar la transformación  $T_r$  de la Sección 2.5, se tiene que en este caso particular  $T_g = T_r$  como esperaríamos.

La Observación 2.13 nos muestra que la transformación  $T_g$  que definimos es una generalización de la transformación  $T_r$  de la Sección 2.5. Sin embargo,

como ya se mencionó antes, la transformación  $T_g$  depende activamente de la función  $g$ , por lo que el desarrollar una técnica empírica para hacer que  $S_g$  se ajuste lo mejor posible a su rectángulo mínimo es una tarea titánica sino que imposible, ya que como  $T_g$  se expresa a través de la función  $g(x)$ , su derivada también se expresará a través de ella, así que en esta ocasión la derivada no puede ser utilizada como en la Sección 2.5.

¿Qué hay respecto al desarrollo de un técnica analítica? Bueno, lo primero que debe ser dicho es que, al trabajar sobre una función, una posible forma de encaminarse es intentar plantear una ecuación diferencial cuya solución sea una función  $g(x)$  con la que  $S_g$  sea una región sencilla de simular con distribución uniforme. Pero ¿por qué una ecuación diferencial supone una buena alternativa para atacar el problema? Bueno, la respuesta es sencilla, si observamos la expresión que determina a  $S_g$ , entonces podemos apreciar que dicha expresión depende de 4 objetos (matemáticos): la constante  $c$  y las funciones  $h(x)$ ,  $g(x)$  y  $g'(x)$ . Pero la función  $h(x)$  siempre es conocida, pues es la densidad de la distribución que deseamos simular, mientras que la constante  $c$  siempre la podemos fijar. Por lo que, si deseamos que dicha expresión nos proporcione determinada región, entonces naturalmente somos llevados a plantear ciertas condiciones sobre  $g(x)$  y  $g'(x)$ , las cuales pueden llegar a plantearse a través de una ecuación diferencial.

Si bien es cierto que es fácil ver que una ecuación diferencial es un buen prospecto para eficientar  $S_g$ , en la práctica puede carecer de viabilidad. Para exhibir esto, restrinjamonos al caso unidimensional.

Veamos, la condición para que un punto  $(u, v)$  esté en  $S_g$  es que se satisfaga la desigualdad

$$0 < u < g^{-1} \left( c \cdot h \left( \frac{v}{g'(u)} \right) \right). \quad (2.25)$$

Así que, si deseamos que  $S_g$  sea el rectángulo  $(0, 1) \times (0, 1)$  ¿que condiciones deben ser impuestas sobre  $g(x)$  y su derivada? La primer condición es que  $[0, 1)$  esté dentro del dominio de  $g(x)$  y que para cada  $u > 1$ , necesaria-

mente se cumpla que

$$g(u) \geq c \cdot h\left(\frac{v}{g'(u)}\right). \quad (2.26)$$

La segunda, y quizá fundamental, es que para cada  $v \in [0, 1)$  se cumpla que para todo  $u \in (0, 1)$

$$0 < g(u) < c \cdot h\left(\frac{v}{g'(u)}\right). \quad (2.27)$$

Sin embargo, el asentar las condiciones (2.26) y (2.27) en una ecuación diferencial sobre  $g(x)$  es bastante difícil, sino que imposible, desde un marco general. A pesar de ello, al restringirnos al contexto particular de algunas distribuciones, la condición en (2.25) puede simplificarse, tal como podemos apreciar en el siguiente ejemplo.

### Ejemplo 2.9 (Distribución logarítmica fraccionaria)

Considere la función  $h(x)$  dada por

$$h(x) = \begin{cases} \frac{1}{\ln(2)} \cdot \frac{1}{(1+x)(2+x)} & \text{para } x > 0, \\ 0 & \text{en otro caso.} \end{cases}$$

Debido a que para todo  $x > 0$  se tiene que

$$\frac{d}{dx} \ln\left(\frac{1+x}{2+x}\right) = \frac{1}{(1+x)(2+x)},$$

la función  $h(x)$  es de densidad. La gráfica de  $h(x)$  en el intervalo  $(0, \infty)$  se encuentra ilustrada en la Figura 2.16.

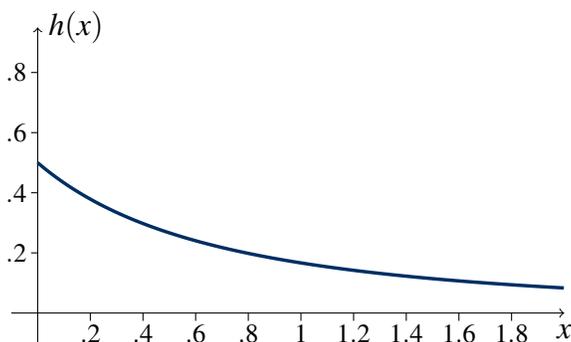


Figura 2.16 . Gráfica de la función de densidad  $h(x)$ .

Veamos, de la hipótesis de monotonía estricta de  $g(x)$ , trivialmente se tiene que

$$0 < u < g^{-1} \left( c \cdot h \left( \frac{v}{g'(u)} \right) \right) \iff 0 < \frac{g(u)}{c} < h \left( \frac{v}{g'(u)} \right). \quad (2.28)$$

Ahora, para desarrollar (2.28) necesitamos observar que la función  $h(x)$  es estrictamente decreciente en  $(0, \infty)$ , lo cual se comprueba fácilmente y puede apreciarse con nitidez en la Figura 2.16. De esta manera, la función  $h(x)$  es invertible en dicho intervalo. De hecho, su inversa, que denotaremos por  $h^{-1}(y)$ , esta dada por

$$h^{-1}(y) = \frac{-3 + \sqrt{9 - 4 \left( 2 - \frac{\ln(2)}{y} \right)}}{2}.$$

Así, al retomar (2.28), obtenemos el siguiente par de equivalencias

$$\begin{aligned} 0 < \frac{g(u)}{c} < h \left( \frac{v}{g'(u)} \right) &\iff 0 < \frac{v}{g'(u)} < h^{-1} \left( \frac{g(u)}{c} \right) \\ &\iff 0 < v < g'(u) h^{-1} \left( \frac{g(u)}{c} \right). \end{aligned}$$

Por lo tanto, para la distribución  $h$  tenemos que su región  $S_g$  del método general del cociente se encuentra dada por

$$S_g = \left\{ (u, v) : 0 < v < g'(u) h^{-1} \left( \frac{g(u)}{c} \right) \right\}.$$

Esto es bastante conveniente, porque significa que  $S_g$  es la región bajo la gráfica de una función y esta función se encuentra dada por  $g(x)$  y  $g'(x)$ . Así, pues, si dicha función es la constante 1, entonces la región  $S_g$  será el rectángulo  $(0, 1) \times (0, 1)$  y, por ende,  $S_g$  tendrá un 100% de eficiencia. Por lo tanto, para que esto ocurra, la función  $g(x)$  debe satisfacer la siguiente ecuación diferencial.

$$g'(u) h^{-1} \left( \frac{g(u)}{c} \right) = 1.$$

Por simplicidad fijaremos el valor de  $c$  a 1. Entonces la ecuación diferencial a resolver es

$$g'(u)h^{-1}(g(u)) - 1 = 0. \quad (2.29)$$

Esta ecuación diferencial es exacta y su solución general está dada por

$$\alpha(x, g(x)) = \text{constante}$$

en donde  $\alpha(x, y) = A(y) - x$  y  $A(y) = \int_0^y h^{-1}(t) dt$ . Por lo tanto, la solución implícita planteada en (2.29) es

$$A(g(x)) = x + \text{constante}.$$

Pero, debido a que la condición inicial es  $g(0) = 0$ , se tiene que

$$0 = A(0) = A(g(0)) = 0 + \text{constante} = \text{constante}.$$

Por lo tanto la función  $g(x)$  que estábamos buscando está dada implícitamente por

$$A(g(x)) = x \quad \text{para toda } x \in [0, 1].$$

Finalmente, debido a que  $A(y)$  es invertible, pues es estrictamente creciente, se obtiene que

$$g(x) = A^{-1}(x) \quad \text{para toda } x \in [0, 1]. \quad (2.30)$$

### **Observación 2.14**

*La función  $A^{-1}(x)$  en (2.30) puede ser dada a través de funciones elementales, pero esta expresión es demasiado extensa, por lo que no haremos su cálculo explícito en este trabajo.*

De esta manera, para la función  $g(x)$  dada en (2.30), se tiene que la región  $S_g$  de la distribución  $h(x)$  es exactamente el rectángulo  $(0, 1) \times (0, 1)$ .

Por lo tanto, al generar un valor  $(u, v)$  uniformemente distribuido en  $(0, 1) \times (0, 1)$  y tomar  $v/(A^{-1})'(u)$ , obtenemos un valor de distribución  $h(x)$ . Además, notemos que en virtud de que  $g(x) = A^{-1}(x)$  satisface la ecuación diferencial en (2.29),

$$\begin{aligned} \frac{v}{(A^{-1})'(u)} &= v \cdot h^{-1}(A^{-1}(u)) \\ &= v \cdot (A \circ h)^{-1}(u). \end{aligned}$$

○

En resumen, para hacer de  $S_g$  una región conocida, podemos intentar plantear una ecuación diferencial sobre  $g(x)$ . Sin embargo, la viabilidad de esto dependerá de la función de densidad  $h(x)$  y la condición que defina a su región  $S_g$ .

## 2.10. Una región rectangular $S_g$ para la distribución normal

Para la distribución normal estándar también puede plantearse una ecuación diferencial sobre  $g(x)$ , de tal manera que su correspondiente región  $S_g$  sea el rectángulo  $(0, 1) \times (-1/2, 1/2)$ . Sin embargo, al igual que en el Ejemplo 2.9, la expresión de  $g(x)$  no será amigable, de hecho, en este caso la función solución  $g(x)$  no podrá ser expresada en términos de funciones elementales.

Sea  $h(x)$  la función de densidad de la distribución normal estándar, es decir,  $h(x) = e^{-x^2/2} / \sqrt{2\pi}$ ,  $-\infty < x < \infty$ . Entonces la siguiente serie de igualdades se mantiene para cualquier función  $g(x)$  que satisface las hipótesis de la Proposición 2.4 para la distribución normal estándar con  $c = 1$ :

$$\begin{aligned} 0 < u < g^{-1} \left( h \left( \frac{v}{g'(u)} \right) \right) &\Leftrightarrow 0 < g(u) < \frac{1}{\sqrt{2\pi}} e^{-\frac{v^2}{2(g'(u))^2}} \\ &\Leftrightarrow 0 < u < d, \ln(g(u)\sqrt{2\pi}) < -\frac{v^2}{2(g'(u))^2}, \end{aligned}$$

donde  $d := g^{-1}(1/\sqrt{2\pi})$ . Por lo tanto, la región  $S_g$  de distribución normal estándar, correspondiente a la función  $g(x)$  y la constante  $c = 1$ , está dada por

$$S_g := \left\{ (u, v) : 0 < u < g^{-1} \left( \frac{1}{\sqrt{2\pi}} \right), 0 < v^2 < -2(g'(u))^2 \ln(g(u)\sqrt{2\pi}) \right\}$$

De esta manera, para que la región  $S_g$  sea el rectángulo  $(0, 1) \times (-1/2, 1/2)$  es necesario que la función satisfaga la siguiente ecuación diferencial en el intervalo  $(0, 1]$

$$g'(u) \sqrt{-\ln(2\pi \cdot g^2(u))} = \frac{1}{2}. \quad (2.31)$$

con condición inicial  $g(1) = 1/\sqrt{2\pi}$ . Pero no es difícil notar que la ecuación diferencial planteada en (2.31) es exacta. Veamos, la expresión general de una ecuación diferencial exacta sobre  $g(x)$  está dada por

$$M(x, g(x)) + N(x, g(x))g'(x) = 0,$$

donde  $M(x, y)$  y  $N(x, y)$  están definidas en una región simplemente conexa y en dicha región satisfacen la relación  $\frac{\partial M}{\partial x_2}(x, y) = \frac{\partial N}{\partial x_1}(x, y)$ .

En el caso de la ecuación planteada en (2.31), las funciones  $N(x, y)$  y  $M(x, y)$  están definidas en el rectángulo  $(0, 1) \times (0, \frac{1}{\sqrt{2\pi}})$  por

$$M(x, y) = -\frac{1}{2} \quad y \quad N(x, y) = \sqrt{-\ln(2\pi y^2)},$$

y es evidente que dichas funciones satisfacen la condición  $\frac{\partial M}{\partial x_2}(x, y) = \frac{\partial N}{\partial x_1}(x, y)$ , por lo que la ecuación diferencial planteada en (2.31) en efecto es exacta. Por lo tanto, la solución a la ecuación diferencial planteada en (2.31) está dada en forma general por

$$\alpha(u, g(u)) = \text{constante},$$

donde  $\alpha(x, y)$  satisface que

$$\frac{\partial \alpha}{\partial x_1}(x, y) = M(x, y) \quad y \quad \frac{\partial \alpha}{\partial x_2}(x, y) = N(x, y). \quad (2.32)$$

Pero ¿qué función satisface las dos igualdades en (2.32)?

En virtud del teorema fundamental del cálculo, la función  $B(y)$  definida en el intervalo  $(0, \frac{1}{\sqrt{2\pi}})$  por

$$B(y) := \int_0^y \sqrt{-\ln(2\pi t^2)} dt, \quad (2.33)$$

es diferenciable y  $B'(y) = \sqrt{-\ln(2\pi y^2)}$ .

### Observación 2.15

Notemos que la integral que define a  $B(y)$  en (2.33) es impropia, pues

$$\lim_{t \rightarrow 0^+} \sqrt{-\ln(2\pi t^2)} = \infty.$$

por lo que, en principio, el teorema fundamental del cálculo no puede ser aplicado directamente en dicha integral. Sin embargo, al notar que

$$\int_0^y \sqrt{-\ln(2\pi t^2)} dt = \int_0^{\frac{1}{\sqrt{2\pi}}} \sqrt{-\ln(2\pi t^2)} dt + \int_{\frac{1}{\sqrt{2\pi}}}^y \sqrt{-\ln(2\pi t^2)} dt$$

se sigue fácilmente que  $B(y)$  esta bien definida y  $B'(y) = \sqrt{-\ln(2\pi y^2)}$ .

Por lo tanto, la función  $\alpha(x, y) := B(y) - x/2$  satisface las igualdades (2.32) y en consecuencia, la solución implícita a la ecuación diferencial planteada en (2.31) está dada por

$$B(g(u)) - \frac{u}{2} = \alpha(u, g(u)) = \text{constante}.$$

Ahora, como la función  $B(y)$  es la función integral de una función positiva, la función  $B(y)$  es invertible. Así, pues, la función  $g(u)$  puede ser dada explícitamente en términos de  $B(y)$ , a saber,

$$g(u) = B^{-1}(u + \text{constante}).$$

Además, dado que  $B(1/\sqrt{2\pi}) = 1/2$ , la condición inicial  $g(1) = 1/\sqrt{2\pi}$  establece que la solución particular que buscamos es con  $\text{constante} = 0$  y, por ende, la función  $g(x)$  con la cual  $S_g = (0, 1) \times (-1/2, 1/2)$ , está dada por

$$g(u) = B^{-1}\left(\frac{u}{2}\right). \quad (2.34)$$

Por lo tanto, al aplicar el método del cociente a la distribución normal estándar con la función  $g(u)$  definida en (2.34) y con  $c = 1$ , la región  $S_g$  resultante es exactamente el rectángulo  $(0, 1) \times (-1/2, 1/2)$ . Así, en virtud del método del cociente, disponemos de un algoritmo para generar valores de distribución normal estándar a partir de valores de distribución uniforme sobre  $(0, 1) \times (-1/2, 1/2)$ .

**Lema 2.1** Si  $(u, v)$  es un valor de distribución uniforme sobre el rectángulo  $(0, 1) \times (-1/2, 1/2)$ , entonces  $2v \cdot \sqrt{-\ln(2\pi \cdot (B^{-1})^2(u/2))}$  es un valor de distribución normal estándar, donde  $B(y)$  es la función definida en (2.33).

**Observación 2.16**

*El volumen de la región  $S_g$  es igual a la constante  $c$ , por lo que al trabajar con  $c = 1$ , era indispensable que el rectángulo objetivo para  $S_g$  tuviera volumen 1. Además, en (2.10) puede observarse que la región  $S_g$  se encuentra delimitada por las gráficas de dos funciones; una negativa y otra positiva, por lo que era natural que el rectángulo objetivo para  $S_g$  tuviera su lado inferior en el cuadrante 4 y su lado superior en el cuadrante 1. De esta manera, el rectángulo  $(0, 1) \times (-1/2, 1/2)$  es la elección natural de rectángulo objetivo para  $S_g$  y esta es la razón por la que en esta sección hemos trabajado con él.*

**Observación 2.17**

*El que la función  $B(y)$  no esté dada explícitamente establece la principal dificultad para la viabilidad del algoritmo que acabamos de desarrollar (ver Lema 2.1), pues requerimos calcular la función  $B(u)$  para poder aplicar el algoritmo.*

**Observación 2.18**

*El algoritmo desarrollado en esta sección para simular a la distribución normal no es la única técnica de esta naturaleza. Por ejemplo, las transformaciones de Box-Muller (ver Apéndice 2) permite generar valores de distribución normal a partir de valores de distribución uniforme sobre el rectángulo  $(0, 1) \times (0, 1)$ . De hecho, las transformaciones de Box-Muller están dadas en términos de funciones elementales; funciones trigonométricas y logarítmicas, por lo que es más sencillo simular computacionalmente a la distribución normal a partir de ellas.*

*Antes de finalizar esta sección, es oportuno mencionar que a pesar de que podemos repetir el proceso de esta sección para un rectángulo distinto, digamos  $(0, 1) \times (0, 1)$ , difícilmente la función  $g(u)$  resultante estará dada en términos de funciones elementales, pues, como describimos en la Observación 2.16, cuando más se simplifican los cálculos es al trabajar con el rectángulo  $(0, 1) \times (-1/2, 1/2)$ . Por lo tanto, si al trabajar con el rectángulo  $(0, 1) \times (-1/2, 1/2)$  no se logró obtener a  $g(u)$  a través de funciones elementales, entonces difícilmente se logrará al trabajar con algún otro rectángulo.*

## 2.11. La región $S_g$ en ciertos casos particulares

En esta sección veremos un tipo especial de distribuciones, para las cuales podemos encontrar una función  $g(x)$  de tal manera que su región  $S_g$  sea un rectángulo, pero sin la necesidad de recurrir al planteamiento de una ecuación diferencial.

### Densidades decrecientes en $(0, \infty)$

Sea  $h(x)$  una función de densidad y  $g(u)$  una función real-valuada para las cuales se satisfacen las hipótesis de la Proposición 2.4 para la distribución  $h$  y la constante  $c > 0$ . Supongamos que el soporte de la función  $h(x)$  es el intervalo  $(0, \infty)$  en donde es decreciente y sobre el cual se tiene que  $h[(0, \infty)] = (0, a)$  para alguna  $a > 0$ . Entonces, bajo estas condiciones, se tiene que

$$\begin{aligned} 0 < u < g^{-1}\left(c \cdot h\left(\frac{v}{g'(u)}\right)\right) &\iff 0 < \frac{g(u)}{c} < h\left(\frac{v}{g'(u)}\right) \\ &\iff 0 < \frac{v}{g'(u)} < h^{-1}\left(\frac{g(u)}{c}\right) \\ &\iff 0 < v < g'(u)h^{-1}\left(\frac{g(u)}{c}\right). \end{aligned}$$

Donde  $h^{-1}(x)$  es la inversa de la función  $h(x)$  restringida al intervalo  $(0, \infty)$ . Por lo tanto, la región  $S_g$  de la distribución  $h$  está dada por

$$S_g := \left\{ (u, v) : 0 < v < g'(u)h^{-1}\left(\frac{g(u)}{c}\right) \right\}. \quad (2.35)$$

Muy bien, ahora consideremos una variable aleatoria  $X$  con función de densidad  $f(x)$  y sea  $M(x)$  una función estrictamente creciente. Por el teorema de cambio de variable, la función de densidad de la variable  $Y := M(X)$  está dada por

$$(M^{-1})'(y) \cdot f(M^{-1}(y)). \quad (2.36)$$

Si comparamos la función en (2.36) con la función cuya gráfica determina a  $S_g$  en (2.35), entonces podemos ver la considerable semejanza de ambas

funciones. De hecho, al tomar  $c = 1$ , la semejanza se torna en igualdad al tomar  $M^{-1}(x) = g(x)$  y  $h^{-1}(x) = f(x)$ . Por lo que, al tomar  $c = 1$ , se tiene que  $S_g$  es la región bajo la gráfica de la función de densidad de la variable  $g^{-1}(X)$  donde  $X \sim h^{-1}$ .

Lo anterior es particularmente provechoso, porque por el método de la función inversa, sabemos que si  $F(x)$  es una función de distribución invertible y  $X \sim F$ , entonces  $F(X) \sim \text{unif}(0, 1)$ .

**Proposición 2.7** Sea  $h(x)$  una función de densidad con soporte  $(0, \infty)$  y estrictamente decreciente en dicho intervalo tal que  $h[(0, \infty)] = (0, b)$  para algún  $b > 0$ . Defina

$$f(x) = \begin{cases} h^{-1}(x) & \text{para } x \in (0, b), \\ 0 & \text{en otro caso.} \end{cases}$$

En donde  $h^{-1}(x)$  es la inversa de la función  $h(x)$  restringida al intervalo  $(0, \infty)$ . Entonces la función  $f(x)$  es una función de densidad y su respectiva función de distribución  $F(x)$  es estrictamente creciente en  $[0, b]$ . Además, si denotamos por  $F^{-1}(x)$  a su inversa en dicho intervalo, entonces la función  $g(x) = F^{-1}(x)$  satisface las hipótesis de la Proposición (2.4) para la distribución  $h$  y la constante  $c = 1$ , y la región  $S_g$  es exactamente el rectángulo  $(0, 1) \times (0, 1)$ .

**Demostración.**

Debido a que la región bajo la gráfica de  $h(x)$  en  $(-\infty, 0)$  coincide con la región bajo la gráfica de  $h^{-1}(x)$ , se tiene que la función  $f(x)$  es de densidad. Además, como el soporte de  $f(x)$  es el intervalo  $(0, b)$ , se tiene que su función de distribución es estrictamente creciente en  $[0, b]$  y, de hecho,  $F[[0, b]] = [0, 1]$ . Denotemos por  $F^{-1}(x)$  a la inversa de  $F(x)$  en este intervalo. Entonces, debido a que  $F(x)$  es estrictamente creciente, la función  $g(x) = F^{-1}(x)$  también es estrictamente creciente, y  $g(x) = 0$  debido a que  $F(0) = 0$ .

Finalmente, dado que  $h[(0, \infty)] = (0, b)$  y  $(0, \infty)$  es el soporte de  $h(x)$ , para

cada  $t \in \mathbb{R}$ , se tiene que

$$h(t) < b = \sup_x g(x).$$

De esta manera, la función  $g(x) = F^{-1}(x)$  satisface las hipótesis de la Proposición (2.4) para la distribución  $h$  y la constante  $c = 1$ . Además, debido a que  $(F^{-1})'(u) = 1/f(F^{-1}(u))$ , se tiene que

$$\frac{v}{g'(u)} = v \cdot f(F^{-1}(u)).$$

Por lo tanto, para toda  $u \in (0, 1)$  se tiene la siguiente serie de equivalencias

$$0 < g(u) < h\left(\frac{v}{g'(u)}\right) \iff 0 < F^{-1}(u) < h(v \cdot f(F^{-1}(u))) \quad (2.37)$$

$$\iff 0 < v \cdot f(F^{-1}(u)) < f(F^{-1}(u)) \quad (2.38)$$

$$\iff 0 < v < 1.$$

Así, pues, concluimos que  $S_g = (0, 1) \times (0, 1)$ .  $\square$

### Observación 2.19

En la Proposición 2.7 se pide como hipótesis que  $h[(0, \infty)] = (0, b)$  para alguna  $b > 0$ , lo que puede parecer una hipótesis bastante restrictiva. Sin embargo, no lo es como lo aparenta, pues como  $h(x)$  es estrictamente decreciente en  $(0, \infty)$ , la imagen directa de este intervalo bajo  $h(x)$  necesariamente es un intervalo abierto. Además, como  $h(x)$  es no negativa, dicho intervalo abierto necesariamente debe estar contenido en  $(0, \infty)$ , es decir, de la monotonía estricta de  $h(x)$  se tiene que  $h[(0, \infty)] = (a, b)$  para algún  $0 \leq a$  y  $a < b$ . De esta manera, la hipótesis  $h[(0, \infty)] = (0, b)$  solo impone la condición de que  $a = 0$ .

### Observación 2.20

La demostración de la Proposición 2.7 no menciona explícitamente dónde es usada la hipótesis  $h[(0, \infty)] = (0, b)$ , por lo que puede cuestionarse su necesidad. Sin embargo, esta posible intuición está errada.

Veamos, sea  $h(x)$  una función de densidad con soporte  $(0, \infty)$  y estrictamente decreciente en dicho intervalo, pero esta vez suponga que  $h[(0, \infty)] =$

$(a, b)$  para algún  $a > 0$  y  $a < b$  (ver Observación 2.19). Entonces, tal como se hizo en la prueba de la Proposición 2.7, podemos definir la función

$$f(x) = \begin{cases} h^{-1}(x) & \text{para } x \in (a, b), \\ 0 & \text{en otro caso.} \end{cases}$$

En donde  $h^{-1}(x)$  es la inversa de la función  $h(x)$  restringida al intervalo  $(0, \infty)$ . La función  $f(x)$  nuevamente es de densidad, pero su soporte ahora será el intervalo  $(a, b)$ , en el cual es decreciente. Esto provoca que su función de distribución  $F(x)$  sea estrictamente creciente en el intervalo  $[a, b]$ , por lo que podemos tomarnos una vez mas a su función inversa  $F^{-1}(x)$ . Sin embargo, dado que  $F(a) = 0$ , se tiene que  $F^{-1}(0) = a > 0$ . Así que esta vez la función  $g(x) := F^{-1}(x)$  no satisface la condición  $g(0) = 0$ , la cual es requerida para aplicar el método general del cociente (ver Proposición 2.4). Por lo tanto, la hipótesis  $h[(0, \infty)] = (0, b)$  para algún  $b > 0$  es necesaria en la Proposición 2.7.

### Observación 2.21

En la prueba de la Proposición 2.7 tampoco se menciona explícitamente la utilidad de que el soporte de  $h(x)$  sea  $(0, \infty)$ , así que es conveniente resaltar el uso de esta hipótesis.

Desarrollemos escrupulosamente el cómo pasamos de (2.37) a (2.38) en la prueba de la Proposición 2.7. Veamos, para que la desigualdad

$$0 < F^{-1}(u) < h(v \cdot f(F^{-1}(u)))$$

se mantenga es necesario que  $h(v \cdot f(F^{-1}(u))) > 0$ , es decir, que  $v \cdot f(F^{-1}(u))$  este en el soporte de  $h(x)$ , el cual es  $(0, \infty)$ . De esta manera, como  $f(x)$  es no negativa, si  $v \cdot f(F^{-1}(u))$  está en el soporte de  $h(x)$ , entonces  $v > 0$  y  $f(F^{-1}(u)) > 0$ . En particular se tiene la implicación

$$0 < F^{-1}(u) < h(v \cdot f(F^{-1}(u))) \implies 0 < v.$$

Y de la monotonía de  $h(x)$  es claro que

$$0 < F^{-1}(u) < h(v \cdot f(F^{-1}(u))) \implies v \cdot f(F^{-1}(u)) < f(F^{-1}(u)).$$

Por lo tanto, tenemos que

$$0 < F^{-1}(u) < h(v \cdot f(F^{-1}(u))) \implies 0 < v \cdot f(F^{-1}(u)) < f(F^{-1}(u)). \quad (2.39)$$

La implicación inversa de (2.39) se sigue inmediatamente de la monotonía de  $f(x)$  y de la no negatividad de  $F^{-1}(x)$ , por lo que la implicación en (2.39) es de hecho una equivalencia y de dicha equivalencia se deduce inmediatamente la equivalencia entre (2.37) y (2.38).

### Observación 2.22

La Observación (2.21) exhibe la utilidad de que  $(0, \infty)$  sea el soporte de  $h(x)$ . Por lo que la Proposición 2.7 parece inútil para funciones de densidad con soporte  $(a, \infty)$  y  $a \neq 0$ , aun cuando dicha densidad cumpla con las restantes hipótesis de la proposición. Sin embargo, una densidad  $h(x)$  con tales características puede ser llevada a una densidad con soporte en  $(0, \infty)$  sin perder ninguna de ellas.

### Densidades crecientes en el intervalo $(-\infty, 0)$

Sea  $h(x)$  una función de densidad y  $g(u)$  una función real-valuada para la cual se satisfacen las hipótesis de la Proposición 2.4 para la distribución  $h$  y la constante  $c > 0$ . Supongamos que el soporte de la función  $h(x)$  es el intervalo  $(-\infty, 0)$ , en el que es creciente y sobre el cual se tiene que  $h[(-\infty, 0)] = (0, b)$  para algún  $b > 0$ . Entonces, bajo estas condiciones, la siguiente cadena de equivalencias es verdadera

$$\begin{aligned} 0 < u < g^{-1} \left( c \cdot h \left( \frac{v}{g'(u)} \right) \right) &\iff 0 < \frac{g(u)}{c} < h \left( \frac{v}{g'(u)} \right) \\ &\iff h^{-1} \left( \frac{g(u)}{c} \right) < \frac{v}{g'(u)} < 0 \\ &\iff g'(u)h^{-1} \left( \frac{g(u)}{c} \right) < v < 0. \end{aligned}$$

Por lo que, en este caso particular, la región  $S_g$  para la distribución  $h$  y la constante  $c$  esta dada por

$$S_g = \left\{ (u, v) : g'(u)h^{-1} \left( \frac{g(u)}{c} \right) < v < 0 \right\}.$$

Por lo tanto, una proposición similar a la Proposición 2.7 puede ser establecida para este tipo particular de distribuciones.

**Proposición 2.8** Sea  $h(x)$  una función de densidad con soporte  $(-\infty, 0)$  y estrictamente creciente en dicho intervalo tal que  $h[(-\infty, 0)] = (0, b)$  para algún  $b > 0$ . Defina

$$f(x) = \begin{cases} -h^{-1}(x) & \text{para } x \in (0, b), \\ 0 & \text{en otro caso.} \end{cases}$$

En donde  $h^{-1}(x)$  es la inversa de la función  $h(x)$  restringida al intervalo  $(-\infty, 0)$ . Entonces la función  $f(x)$  es una función de densidad y su respectiva función de distribución  $F(x)$  es estrictamente creciente en  $[0, b]$ . Además, si denotamos por  $F^{-1}(x)$  a su inversa en dicho intervalo, entonces la función  $g(x) = F^{-1}(x)$  satisface las hipótesis de la Proposición (2.4) para la distribución  $h$  y la constante  $c = 1$ , y la región  $S_g$  es exactamente el rectángulo  $(0, 1) \times (-1, 0)$ .

**Demostración.**

Debido a que la región bajo la gráfica de  $h(x)$  en  $(-\infty, 0)$  coincide con la región bajo la gráfica de  $-h^{-1}(x)$ , se tiene que la función  $f(x)$  es de densidad. Además, como el soporte de  $f(x)$  es el intervalo  $(0, b)$ , se tiene que su función de distribución es estrictamente creciente en  $[0, b]$  y, de hecho,  $F[[0, b]] = [0, 1]$ . Denotemos por  $F^{-1}(x)$  a la inversa de  $F(x)$  en este intervalo. Entonces, debido a que  $F(x)$  es estrictamente creciente, la función  $g(x) = F^{-1}(x)$  también es estrictamente creciente y  $g(x) = 0$  debido a que  $F(0) = 0$ .

Finalmente, dado que  $h[(-\infty, 0)] = (0, b)$  y  $(0, \infty)$  es el soporte de  $h(x)$ , para cada  $t \in \mathbb{R}$ , se tiene que

$$h(t) < b = \sup_x g(x).$$

De esta manera, la función  $g(x) = F^{-1}(x)$  satisface las hipótesis de la Proposición (2.4) para la distribución  $h$  y la constante  $c = 1$ . Además, debido a

que  $(F^{-1})'(u) = 1/f(F^{-1}(u))$ , se tiene que

$$\frac{v}{g'(u)} = v \cdot f(F^{-1}(u)).$$

Por lo tanto, para toda  $u \in (0, 1)$  se tiene la siguiente serie de equivalencias

$$0 < g(u) < h\left(\frac{v}{g'(u)}\right) \iff 0 < F^{-1}(u) < h(v \cdot f(F^{-1}(u))) \quad (2.40)$$

$$\iff h^{-1}(F^{-1}(u)) < v \cdot f(F^{-1}(u)) < 0 \quad (2.41)$$

$$\iff -1 < v < 0.$$

Así, pues, concluimos que  $S_g = (0, 1) \times (-1, 0)$ .  $\square$

La siguiente serie de observaciones es análoga a la llevada a cabo para la Proposición 2.7.

### Observación 2.23

En la Proposición 2.8 se pide como hipótesis que  $h[(-\infty, 0)] = (0, b)$  para alguna  $b > 0$ , lo que puede parecer una hipótesis bastante restrictiva, sin embargo no lo es como lo aparenta, pues como  $h(x)$  es estrictamente creciente en  $(-\infty, 0)$ , la imagen directa de este intervalo bajo  $h(x)$  necesariamente es un intervalo abierto. Además, como  $h(x)$  es no negativa, dicho intervalo abierto necesariamente debe estar contenido en  $(0, \infty)$ , es decir, de la monotonía estricta de  $h(x)$  se tiene que  $h[(-\infty, 0)] = (a, b)$  para algún  $0 \leq a$  y  $a < b$ . De esta manera, la hipótesis  $h[(-\infty, 0)] = (0, b)$  solo impone la condición de que  $a = 0$ .

### Observación 2.24

La demostración de la Proposición 2.8 no menciona explícitamente dónde es usada la hipótesis  $h[(-\infty, 0)] = (0, b)$ , por lo que puede cuestionarse su necesidad. Sin embargo, esta posible intuición está errada.

Veamos, sea  $h(x)$  una función de densidad con soporte  $(-\infty, 0)$  y estrictamente creciente en dicho intervalo, pero esta vez suponga que  $h[(-\infty, 0)] = (a, b)$  para algún  $a > 0$  y  $a < b$  (ver Observación 2.23). Entonces, tal como

se hizo en la prueba de la Proposición 2.8, podemos definir la función

$$f(x) = \begin{cases} -h^{-1}(x) & \text{para } x \in (a, b), \\ 0 & \text{en otro caso.} \end{cases}$$

En donde  $h^{-1}(x)$  es la inversa de la función  $h(x)$  restringida al intervalo  $(-\infty, 0)$ . Nuevamente, la función  $f(x)$  es de densidad, pero su soporte ahora será el intervalo  $(a, b)$ , en el cual es creciente. Esto provoca que su función de distribución  $F(x)$  sea estrictamente creciente en el intervalo  $[a, b]$ , por lo que podemos tomarnos una vez mas a su función inversa  $F^{-1}(x)$ . Sin embargo, dado que  $F(a) = 0$ , se tiene que  $F^{-1}(0) = a > 0$ . Así que esta vez la función  $g(x) := F^{-1}(x)$  no satisface la condición  $g(0) = 0$ , la cual es requerida para aplicar el método general del cociente (ver Proposición 2.4). Por lo tanto, la hipótesis  $h[(-\infty, 0)] = (0, b)$  para algún  $b > 0$  es necesaria en la Proposición 2.8.

### Observación 2.25

En la prueba de la Proposición 2.8 tampoco se menciona explícitamente la utilidad de que el soporte de  $h(x)$  sea  $(-\infty, 0)$ , así que es conveniente resaltar el uso de esta hipótesis.

Desarrollemos cuidadosamente el cómo pasamos de (2.40) a (2.41) en la prueba de la Proposición 2.8. Veamos, para que la desigualdad

$$0 < F^{-1}(u) < h(v \cdot f(F^{-1}(u))).$$

se mantenga es necesario que  $h(v \cdot f(F^{-1}(u))) < 0$ , es decir, que  $v \cdot f(F^{-1}(u))$  este en el soporte de  $h(x)$ , el cual es  $(-\infty, 0)$ . De esta manera, como  $f(x)$  es no negativa, si  $v \cdot f(F^{-1}(u))$  está en el soporte de  $h(x)$ , entonces  $v < 0$  y  $f(F^{-1}(u)) > 0$ . En particular se obtiene la implicación

$$0 < F^{-1}(u) < h(v \cdot f(F^{-1}(u))) \implies v < 0$$

Y de la monotonía de  $h(x)$  es claro que

$$0 < F^{-1}(u) < h(v \cdot f(F^{-1}(u))) \implies h^{-1}(F^{-1}(u)) < v \cdot f(F^{-1}(u)).$$

Por lo tanto, tenemos que

$$0 < F^{-1}(u) < h(v \cdot f(F^{-1}(u))) \implies h^{-1}(F^{-1}(u)) < -v \cdot h^{-1}(F^{-1}(u)) < 0. \quad (2.42)$$

La implicación inversa de (2.42) se sigue inmediatamente de la monotonía de  $f(x)$  y de la no negatividad de  $F^{-1}(x)$ , por lo que la implicación en (2.42) es de hecho una equivalencia y de dicha equivalencia se deduce inmediatamente la equivalencia entre (2.40) y (2.41).

### **Observación 2.26**

La Observación (2.25) exhibe la utilidad de que  $(-\infty, 0)$  sea el soporte de  $h(x)$ . Por lo que la Proposición 2.8 parece inútil para funciones de densidad con soporte  $(-\infty, a)$  y  $a \neq 0$ , aun cuando dicha densidad cumpla con las restantes hipótesis de la proposición. Sin embargo, una densidad  $h(x)$  con tales características puede ser llevada a una densidad con soporte en  $(-\infty, 0)$  sin perder ninguna de ellas.

### **Observación 2.27 (La inversa general de la función de distribución)**

Recordemos que el concepto de función inversa es extendido para funciones monótonas, entre las cuales se encuentran las funciones de distribuciones, por lo que es natural preguntarse si las Proposiciones 2.7 y 2.8 no pueden extenderse usando esta función inversa general. Sin embargo, para la demostración de las Proposiciones 2.7 y 2.8 resulta fundamental la igualdad

$$(F^{-1})'(u) = \frac{1}{f(F(u))},$$

pero dicha igualdad no se mantiene para la inversa generalizada  $F^{-1}(x)$ . Por ende, las Proposiciones 2.7 y 2.8 no pueden ser extendidas o al menos no usando la inversa extendida.

### **Notas y comentarios**

Las condiciones adicionales de la versión débil del método presente en la Proposición 2.5 son un generalización de las impuestas en las versiones previas del método para que las regiones resultantes sean acotadas. Sin embargo, como se menciono en su momento, estas condiciones son suficientes, mas no necesarias.

Al derivar y usar el teorema de la función inversa se puede comprobar fácilmente que la función  $T_g(u, v_1, \dots, v_n)$  es un difeomorfismo.

Las Proposiciones 2.7 y 2.8 son establecidas, aunque no enunciadas, en el trabajo de Martino, Luengo y Miguez en [10], por lo que al enunciarlas y comentarlas en el presente trabajo se busca esclarecer la forma en que actúa cada una de las hipótesis. Esta es la razón de las extensas observaciones que en este trabajo se derivan de dichas proposiciones.

La idea de una optimización de  $S_g$  a través de una ecuación diferencial sobre  $g(x)$  puede parecer encaminada al trabajo de Jones y Lunn en [6]; pero, aunque en cierta medida lo esta, se tiene una diferencia intrínseca.

El trabajo de Jones y Lunn asume que una cierta transformación  $\mathcal{D}(u, v)$  es dada y busca el correspondiente dominio  $\mathcal{A}$  de un vector aleatorio  $(U, V)$  de tal manera que  $D(U, V) \sim h$  para determinada distribución objetivo  $h$ . Por otro lado, la optimización de  $S_g$  fija la transformación  $\mathcal{D}(u, v) = v/g'(u)$  al igual que la región  $\mathcal{A} = S_g$  de tal manera que si  $(U, V) \sim \text{unif}(\mathcal{A})$ , entonces  $D(U, V) \sim h$  para la distribución objetivo  $h$  y lo que busca encontrar la función  $g(x)$  para la cual  $S_g$  sea una región simple, tal como un rectángulo.

# Conclusiones y discusión general

*A continuación se establecen las conclusiones del presente trabajo, pero también se discuten de forma general posibles líneas de investigación que pueden seguirse respecto al método general del cociente. Estas discusiones se desarrollan en un contexto de meras suposiciones e intuiciones, por lo que el lector debe criticar con severidad las ideas y nociones que se presentan.*

## Las regiones del método del cociente

*En la Sección 1.3 se caracterizó por completo a las regiones que resultan de aplicar el método del cociente a distribuciones univariadas. Esta caracterización se establece en la Proposición 1.6 y en ella se deja en claro la utilidad de describir a las regiones en coordenadas polares. Además, podemos apreciar que la idea fundamental detrás de la descripción en coordenadas polares yace fundamentalmente en el hecho de que el ángulo que forma el vector  $(u, v)$  con el eje  $X$  es igual a  $\arctan(v/u)$ . De esta manera, una descripción similar puede darse en varias dimensiones al considerar que si  $(u, v_1, \dots, v_n) \in \mathbb{R}^n$  con  $0 < u$ , entonces  $\arctan(v_k/u)$  es igual al ángulo que forma el vector  $(u, 0, \dots, v_k, \dots, 0)$  con el eje  $X$ .*

*La Proposición 1.6 nos permite definir distribuciones con ciertas características particulares, pues si nos tomamos una región  $A$  que satisface las hipótesis de la Proposición 1.6, entonces existe una distribución cuya región  $S$  coincide exactamente con  $A$  y, además, la función de densidad de esta distribución queda totalmente determinada por la curva en coordenadas polares*

que delimita a  $A$ .

### **Observación 2.28**

*Al caracterizar a las regiones del método del cociente en la Sección 1.3, se estableció la forma en que ciertas características de una distribución pueden determinarse geoméricamente a partir de la región que resulta de aplicarles el método del cociente. Sin embargo, ciertas propiedades de la distribución parecen mas difíciles de interpretar geoméricamente en estas regiones, por lo que es interesante examinar en qué radica esta dificultad y si el considerar las regiones del método general nos proporcionan una alternativa.*

## **Las regiones del método general del cociente**

*En la Sección 1.3 del Capítulo 1, la descripción en coordenadas polares de las regiones del método del cociente nos llevan a plantear y demostrar la Proposición 1.6, por lo que es natural pensar que una proposición semejante debe seguirse de la descripción en coordenadas polares de la región  $S_g$  del método general del cociente. Sin embargo, debido a que la región  $S_g$  depende de  $g(x)$ , para cada distribución se tiene una región  $S_g$  para cada  $g(x)$ , por lo que no se puede establecer una relación biunívoca entre una distribución y su regiones  $S_g$ . Aunque, al fijar  $g(x)$ , sí se puede obtener una proposición análoga a la Proposición 1.6 para esta función  $g(x)$ .*

*La descripción en coordenadas polares de la región  $S_g$  del método general del cociente puede extenderse naturalmente a varias dimensiones, tal como ocurrió para la región  $S$  del método del cociente. Dicha generalización proviene de que si  $(u, v_1, \dots, v_n) \in \mathbb{R}^{n+1}$ , entonces  $\arctan\left(v_k / \sqrt[n]{g'(u)}\right)$  es igual al ángulo que forma el vector  $(g'(u), 0, \dots, v_k, \dots, 0)$  con el eje  $X$ .*

*Al final de la Sección 2.9 se exhorta a plantear ecuaciones diferenciales sobre la función  $g(x)$  con el propósito de que la región  $S_g$  sea un rectángulo, pero también se advierte que dicho planteamiento en algunos casos puede llegar a ser una tarea titánica sino que imposible. Sin embargo, ¿existirán otras regiones fáciles de simular con distribución uniforme de tal manera que siempre sea relativamente sencillo plantear una ecuación diferencial*

sobre  $g(x)$ ? Esta es una pregunta difícil de responder, ya que existe una cantidad infinita de regiones, pero quizá la siguiente serie de observaciones ayuden a responder esta pregunta.

**Observación 2.29 (Secciones circulares)**

El transformar continua o diferenciablemente una región en otra puede ser difícil sino que imposible en algunas ocasiones y es intuitivamente claro que la dificultad dependerá de las regiones. Por ejemplo, intuitivamente, es más sencillo deformar un círculo en una elipse que un círculo en un cuadrado. Por lo tanto, para hacer de  $\mathcal{B}$  (ver Sección 2.8) una región sencilla, las condiciones que se deben imponer sobre  $g(x)$  dependerán fuertemente de la región objetivo para  $\mathcal{B}$ . Sin embargo, dado que la región está delimitada por una función en coordenadas polares, una excelente alternativa son las secciones circulares.

Veamos, para que  $\mathcal{B}$  sea una región de secciones circulares del mismo radio basta con que su función delimitadora sea constante en los puntos en donde no se anula, lo cual impone una ecuación diferencial relativamente sencilla sobre  $g(x)$ .

La función delimitadora (en coordenadas polares) de la región  $\mathcal{B}$  está dada por

$$\rho(\theta) := \frac{g'(g^{-1}(h(\tan(\theta))))}{\cos(\theta)}.$$

Por lo tanto, para que dicha función sea constante, es necesario que se satisfaga la siguiente ecuación diferencial sobre el soporte de  $\rho(\theta)$ .

$$\frac{g'(g^{-1}(h(\tan(\theta))))}{\cos(\theta)} = \text{constante}. \tag{2.43}$$

Sin pérdida de generalidad, trabajemos la ecuación diferencial en (2.43) con la constante igual a 1. De esta manera, la ecuación diferencial a resolver es la siguiente

$$g'(g^{-1}(h(\tan(\theta)))) = \cos(\theta). \tag{2.44}$$

Luego, al suponer que  $h(x)$  es diferenciable, por la regla de la cadena y

completando derivadas, la ecuación en (2.44) se torna en la siguiente

$$\frac{d}{d\theta}(g^{-1}(h(\tan(\theta)))) = \sec^3(\theta)h'(\tan(\theta)). \quad (2.45)$$

La ecuación (2.45) difícilmente puede ser resuelta explícitamente, aun para distribuciones sencillas como la exponencial. Sin embargo, usando métodos de aproximación, se pueden dar una buena función solución a la ecuación (2.45). Pero el hecho de que podamos darle solución a esta ecuación diferencial no implica que hallamos simplificado a la región  $S_g$ , pues recordemos que  $S_g = \mathcal{T}^{-1}[\mathcal{B}]$  y al simplificar a la región  $\mathcal{B}$  puede que  $\mathcal{T}(u, v)$  se halla complicado bastante.

### Observación 2.30

Es tentador pensar que si  $X \sim \text{unif}(\mathcal{B})$ , entonces  $\mathcal{T}^{-1}(X) \sim \text{unif}(\mathcal{T}^{-1}(\mathcal{B}))$ , lo cual haría que la Observación 2.29 tuviera importantes consecuencias. Sin embargo, esta tentadora intuición es correcta sólo en complejas circunstancias. Examinemos cuales son estas circunstancias.

Bajo la hipótesis de convexidad estricta de  $g(x)$ , la cual es requerida para la descripción polar de  $S_g$ , la función  $\mathcal{T}^{-1}(u, v)$  es un difeomorfismo. Por lo tanto, si  $X \sim \text{unif}(\mathcal{B})$ , entonces la función de densidad  $f_Y(u, v)$  para la variable  $Y := \mathcal{T}^{-1}(X)$  esta dada por

$$f_Y(u, v) = \frac{|\det(D\mathcal{T}(u, v))|}{\text{vol}(\mathcal{B})} \cdot \mathbb{1}_{\mathcal{B}}(\mathcal{T}(u, v)). \quad (2.46)$$

Ahora, debido a que  $\mathbb{1}_{\mathcal{B}} \circ \mathcal{T}(u, v) = \mathbb{1}_{\mathcal{T}^{-1}[\mathcal{B}]}$ , la expresión en (2.46) puede refinarse un poco mas, obteniendo

$$f_Y(u, v) = \frac{|\det(D\mathcal{T}(u, v))|}{\text{vol}(\mathcal{B})} \cdot \mathbb{1}_{\mathcal{T}^{-1}[\mathcal{B}]}(u, v).$$

En consecuencia,  $Y \sim \text{unif}(\mathcal{T}^{-1}[\mathcal{B}])$  si y solo si para todo  $(u, v) \in \mathcal{T}^{-1}[\mathcal{B}]$ ,

$$\frac{|\det(D\mathcal{T}(u, v))|}{\text{vol}(\mathcal{T}^{-1}[\mathcal{B}])} = \frac{1}{\text{vol}(\mathcal{B})},$$

, es decir,

$$|\det(D\mathcal{T}(u, v))| = \frac{\text{vol}(\mathcal{T}^{-1}[\mathcal{B}])}{\text{vol}(\mathcal{B})}. \quad (2.47)$$

Luego, como  $\mathcal{T}(u, v) = (g'(u), v)$ , se tiene  $|\det(D\mathcal{T}(u, v))| = g''(u)$ . De esta manera, la igualdad en (2.47) se mantendrá si y solo si para todo  $(u, v) \in \mathcal{T}^{-1}[\mathcal{B}]$ ,

$$g''(u) = \frac{\text{vol}(\mathcal{T}^{-1}[\mathcal{B}])}{\text{vol}(\mathcal{B})}. \quad (2.48)$$

En consecuencia, las circunstancias de las que hablábamos al comienzo de la observación son las impuestas en (2.48).

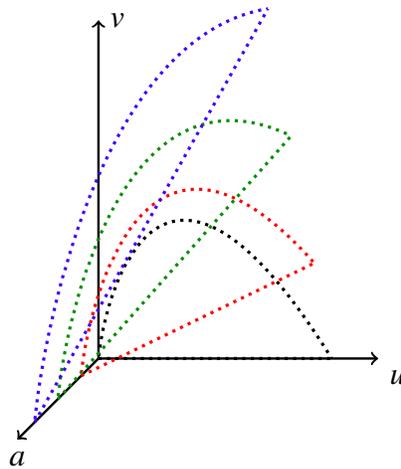
### **Observación 2.31**

Debido a que la región  $\mathcal{B}$  está determinada por  $g(x)$ , podemos pensar que la igualdad planteada en (2.48) plantea una ecuación diferencial sobre  $g(x)$ , cuya solución satisface que si  $X \sim \text{unif}(\mathcal{B})$ , entonces  $\mathcal{T}^{-1}[\mathcal{B}] \sim \text{unif}(\mathcal{T}^{-1}[\mathcal{B}])$ . Sin embargo, la igualdad en (2.48) difícilmente puede ser aterrizada en una ecuación diferencial, ya que, aun cuando  $\mathcal{B}$  esté expresado a través de  $g(x)$ , el volumen de  $\mathcal{B}$  y el de  $\mathcal{T}^{-1}[\mathcal{B}]$  difícilmente serán expresados a través de la función  $g(x)$ .

## **La región de una variable aleatoria al sumarle una constante**

En las Sección 1.5 se establece que si una distribución satisface las hipótesis del método del cociente (ver Proposición 1.3), entonces también las satisfacen las distribuciones que resultan de sumar una constante. Además, se establece la relación que tiene las regiones de estas distribuciones con la región correspondiente a la distribución original, de hecho, se demuestra que existe un isomorfismo lineal de por medio, el cual lleva una región en la otra.

Dado que el isomorfismo lineal que lleva a  $S_X$  en  $S_{X+a}$  depende continuamente de  $a$  (ver (1.13)), la deformación de  $S_X$  en  $S_{X+a}$  es continua. Además, note que si graficamos a la región  $S_{X+a}$  en función de  $a$ , entonces obtendremos una gráfica similar a la de la Figura 2.17, en la cual se ilustra a la región  $S_{X+a}$  de la distribución exponencial en función de  $a$ .



**Figura 2.17** . La región  $S_{X+a}$  para la distribución exponencial.

**Observación 2.32**

En la Figura 2.17 también se puede inferir que al graficar a  $S_{X+a}$  en función de  $a$  se obtiene un sólido no acotado en  $\mathbb{R}^3$ . Este sólido contiene en forma geométrica y compacta toda la información correspondiente a una distribución cuando se le suma una constante.

**Observación 2.33**

El entender a la región  $S_{X+a}$  en función de  $a$  puede ayudar a entender la región de la suma de variables aleatorias independientes, pues recordemos que si  $X$  y  $Y$  son variables aleatorias independientes y  $f_Y(y)$  es la función de densidad de  $Y$ , entonces, para cada  $a \in \mathbb{R}$  con  $f_Y(a) \neq 0$ , se tiene que

$$X + a = X + Y | (Y = a).$$

Así, pues, el sólido descrito en la Observación 2.32 puede ser de utilidad para determinar la región de la suma de variables aleatorias independientes, pero no hay un indicio claro de cómo dicho sólido debe ser utilizado.

**Generalización en tres partes**

La generalización del método del cociente se llevó a cabo a lo largo del Capítulo 2 y fue dividida en tres partes. La primera generalización consiste

*de una simple extensión a varias dimensiones, en la segunda se agrega un parámetro  $r$  que determina a las regiones resultantes y en la tercera se introduce una función  $g(x)$ , la cual abstrae en gran medida al parámetro  $r$  de la segunda generalización y determina activamente a las regiones resultantes.*

*La generalización en tres partes del método del cociente realizada en este trabajo goza principalmente de dos beneficios: una extensión natural y una vasta asimilación de los mecanismos que permite funcionar al método. Por lo tanto, el presente trabajo puede ser utilizado como punto de partida para el estudio del método del cociente.*

## **Descripción de la transformación $T_r$**

*El análisis global que se hace de la transformación  $\mathcal{T}$  de la Sección 2.5 proviene del análisis local de dicha transformación, por lo que la descripción global que se hace de  $\mathcal{T}$  es inherentemente incompleta, independientemente de cuan minuciosos seamos en dicha descripción.*

*La incompletitud inherente de la transformación  $\mathcal{T}$  provoca el cauteloso uso que debemos hacer de él en busca de valores de  $r$  para los cuales las regiones resultantes se ajusten lo mejor posible a su rectángulo mínimo.*

*Finalmente, la continuidad de  $T_r$  y mas generalmente la de  $T_g$ , implica que la deformación de  $S$  en  $S_r$  (o en su caso  $S_g$ ) sea continua.*

# Apéndice 1

Sea  $(U, V_1, \dots, V_n)$  un vector aleatorio de dimensión  $n + 1$ . Si dicho vector posee función de densidad, entonces el vector aleatorio  $(V_1/U, \dots, V_n/U)$  también posee función de densidad y puede ser calculada explícitamente a partir del teorema de cambio de variable.

## Proposición 2.9 (Densidad del cociente de variables aleatorias)

Sea  $(U, V_1, \dots, V_n)$  un vector aleatorio con densidad  $f(x_1, \dots, x_{n+1})$ . Suponga que  $U \neq 0$ , entonces la función de densidad del vector  $(V_1/U, \dots, V_n/U)$  está dada por

$$g(y_1, \dots, y_n) = \int_{-\infty}^{\infty} |z|^n f(z, zy_1, \dots, zy_n) dz.$$

### **Demostración.**

Defina sobre el conjunto  $D := \{(u, v_1, \dots, v_n) \in \mathbb{R}^{n+1} : u \neq 0\}$  a la función

$$\phi(u, v_1, \dots, v_n) = (u, v_1/u, \dots, v_n/u).$$

Por el teorema de la función inversa,  $\phi(u, v_1, \dots, v_n) = (u, v_1/u, \dots, v_n/u)$  es un difeomorfismo y, de hecho, su función inversa está dada por

$$\phi^{-1}(z, y_1, \dots, y_n) = (z, zy_1, \dots, zy_n).$$

Debido a que  $U \neq 0$ , la imagen del vector  $(U, V_1, \dots, V_n)$  se encuentra contenida en  $D$ . Por lo tanto, al aplicar el teorema de cambio de variable, obtenemos que la función de densidad del vector aleatorio

$$(U, V_1/U, \dots, V_n/U) = \phi(U, V_1, \dots, V_n)$$

está dada por

$$h(z, y_1, \dots, y_n) = f(\phi^{-1}(z, y_1, \dots, y_n)) \cdot |\det(D\phi^{-1}(z, y_1, \dots, y_n))|.$$

De esta manera, debido a que

$$D\phi^{-1}(z, y_1, \dots, y_n) = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ y_1 & z & 0 & \cdots & 0 & 0 & 0 \\ y_2 & 0 & z & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ y_{n-2} & 0 & 0 & \cdots & z & 0 & 0 \\ y_{n-1} & 0 & 0 & \cdots & 0 & z & 0 \\ y_n & 0 & 0 & \cdots & 0 & 0 & z \end{bmatrix}$$

se tiene que  $|\det(D\phi^{-1}(z, y_1, \dots, y_n))| = |z|^n$ . Consecuentemente,

$$h(z, y_1, \dots, y_n) = |z|^n f(z, zy_1, \dots, zy_n).$$

Por ende, la función de densidad del vector  $(V_1/U, \dots, V_n/U)$  está dada por

$$g(y_1, \dots, y_n) = \int_{-\infty}^{\infty} |z|^n f(z, zy_1, \dots, zy_n) dz.$$

□

### Observación 2.34

Se puede prescindir de la hipótesis  $U \neq 0$  en la Proposición 2.9, pues el conjunto de puntos  $(u, v_1, \dots, v_n)$  tales que  $u = 0$  es de medida de Lebesgue 0 en  $\mathbb{R}^{n+1}$ .

Consideremos una vez más  $(U, V_1, \dots, V_n)$  un vector aleatorio de dimensión  $n + 1$  y suponga que  $r \geq 0$ . Al tomar el vector aleatorio  $(V/U^r, \dots, V/U^r)$ , es evidente que el vector  $(V_1/U, \dots, V_n/U)$  es el caso particular  $r = 1$ . Por lo que una natural extensión de la Proposición 2.9 se logra al calcular la función de densidad del vector  $(V/U^r, \dots, V/U^r)$  a través del teorema de cambio de variable.

**Proposición 2.10**

Sea  $(U, V_1, \dots, V_n)$  un vector aleatorio con función de densidad  $f(x_1, \dots, x_{n+1})$  y sea  $r \geq 0$ . Suponga que  $U \neq 0$ , entonces la función de densidad del vector  $(V_1/uU^r, \dots, V_n/U^r)$  está dada por

$$g(y_1, \dots, y_n) := \int_{-\infty}^{\infty} |z|^{nr} f(z, z^r y_1, \dots, z^r y_n) dz.$$

**Demostración.**

Defina sobre el conjunto  $D := \{(u, v_1, \dots, v_n) \in \mathbb{R}^{n+1} : u \neq 0\}$  a la función

$$\phi(u, v_1, \dots, v_n) = (u, v_1/u^r, \dots, v_n/u^r).$$

Por el teorema de la función inversa,  $\phi(u, v_1, \dots, v_n) = (u, v_1/u, \dots, v_n/u)$  es un difeomorfismo y, de hecho, su función inversa está dada por

$$\phi^{-1}(z, y_1, \dots, y_n) = (z, z^r y_1, \dots, z^r y_n).$$

Debido a que  $U \neq 0$ , la imagen del vector  $(U, V_1, \dots, V_n)$  se encuentra contenida en  $D$ . Por lo tanto, al aplicar el teorema de cambio de variable, obtenemos que la función de densidad del vector aleatorio

$$(U, V_1/U, \dots, V_n/U) = \phi(U, V_1, \dots, V_n)$$

está dada por

$$h(z, y_1, \dots, y_n) = f(\phi^{-1}(z, y_1, \dots, y_n)) \cdot |\det(D\phi^{-1}(z, y_1, \dots, y_n))|.$$

De esta manera, debido a que

$$D\phi^{-1}(z, y_1, \dots, y_n) = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ rz^{r-1}y_1 & z^r & 0 & \cdots & 0 & 0 & 0 \\ rz^{r-1}y_2 & 0 & z^r & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ rz^{r-1}y_{n-2} & 0 & 0 & \cdots & z^r & 0 & 0 \\ rz^{r-1}y_{n-1} & 0 & 0 & \cdots & 0 & z^r & 0 \\ rz^{r-1}y_n & 0 & 0 & \cdots & 0 & 0 & z^r \end{bmatrix}$$

se tiene que  $|\det(D\phi^{-1}(z, y_1, \dots, y_n))| = |z|^{nr}$ . Consecuentemente,

$$h(z, y_1, \dots, y_n) = |z|^{nr} f(z, z^r y_1, \dots, z^r y_n).$$

Por ende, la función de densidad del vector  $(V_1/U, \dots, V_n/U)$  está dada por

$$g(y_1, \dots, y_n) = \int_{-\infty}^{\infty} |z|^{nr} f(z, z^r y_1, \dots, z^r y_n) dz.$$

□

## Apéndice 2

Las transformaciones de Box-Muller pueden probarse de distintas maneras, ver, por ejemplo, [10] y [2]. En este apéndice las probaremos usando el difeomorfismo de cambio de coordenadas polares a cartesianas.

**Lema 2.2** Si  $U \sim \text{unif}(0, 1)$ , entonces  $Y := -2\ln(U) \sim \chi^2(2)$ .

### **Demostración.**

Por el teorema de cambio de variable, tenemos que la función de densidad de la variable  $Y$  esta dada por

$$f_Y(y) = \frac{1}{2} e^{-\frac{y}{2}} \mathbb{1}_{(0,1)}\left(e^{-\frac{y}{2}}\right). \quad (2.49)$$

Finalmente, como

$$e^{-\frac{y}{2}} \in (0, 1) \iff y > 0,$$

se concluye inmediatamente de (2.49) que  $Y \sim \chi^2(2)$ .  $\square$

**Proposición 2.11** Si  $U$  y  $V$  son variables *i.i.d* con distribución uniforme sobre el intervalo  $(0, 1)$ , entonces las variables

$$\cos(2\pi V) \sqrt{-2\ln(U)} \quad \text{y} \quad \text{sen}(2\pi V) \sqrt{-2\ln(U)} \quad (2.50)$$

son *i.i.d* con distribución normal estándar.

### **Demostración.**

Dadas las hipótesis, defina las variables aleatorias

$$R^2 := -2\ln(U) \quad \text{y} \quad \Theta := 2\pi V.$$

Claramente  $R^2$  y  $\Theta$  son independientes.

Consideremos el difeomorfismo  $\Psi(r, \theta)$  definido sobre  $(0, 2\pi] \times (0, \infty)$  por

$$\Psi(r, \theta) := (\sqrt{r} \cos(\theta), \sqrt{r} \sin(\theta)).$$

Al derivar parcialmente se puede comprobar con facilidad que

$$\begin{bmatrix} \frac{1}{2\sqrt{r}} \cos(\theta) & -\sqrt{r} \sin(\theta) \\ \frac{1}{2\sqrt{r}} \sin(\theta) & \sqrt{r} \cos(\theta) \end{bmatrix}.$$

Por lo tanto,

$$|\det(D\Psi(r, \theta))| = \frac{1}{2}.$$

De esta manera, por el teorema de cambio de variable, la función de densidad de  $(X, Y) := \Psi(R^2, \Theta)$  está dada por

$$f(x, y) = f_{(R^2, \Theta)}(\Psi^{-1}(x, y)) \cdot |\det(D\Psi^{-1}(x, y))|.$$

Ahora, por teorema de la función inversa, se tiene que  $|\det(D\Psi^{-1}(x, y))| = 2$ . En consecuencia, tenemos que

$$\begin{aligned} f(x, y) &= 2f_{(R^2, \Theta)}(\Psi^{-1}(x, y)) \\ &= 2f_{(R^2, \Theta)}(x^2 + y^2, \theta(x, y)) \\ &= 2f_{R^2}(x^2 + y^2) \cdot f_{\Theta}(\theta(x, y)) \\ &= 2 \cdot \left( \frac{1}{2} e^{-\frac{x^2+y^2}{2}} \right) \cdot \left( \frac{1}{2\pi} \mathbb{1}_{(0, 2\pi)}(\theta(x, y)) \right) \\ &= \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}}. \end{aligned}$$

De donde se sigue inmediatamente que las variables definidas en (2.50) son *i.i.d* con distribución normal estándar.  $\square$

# Bibliografía

- [1] Athreya K. B. y Lahiri S. N. *Measure theory and probability theory*. Springer (2006).
- [2] Box G. E. P. y Muller M. E. A note on the generation of random normal deviates. *The Annals of Mathematical Statistics*. **29**(2), 610–611 (1958).
- [3] Dudley R. M. *Real analysis and probability*. 2nd ed. Cambridge University Press (2002).
- [4] Duistermaat J. J. y Kolk J. A. C. *Multidimensional real analysis I: Differentiation*. Number 86 in Cambridge Studies in Advanced Mathematics. Cambridge University Press (2004).
- [5] Jones M. C. On Khintchine’s theorem and its place in random variate generation. *Am. Stat.* **56**(4), 304–307 (2002).
- [6] Jones M. C. y Lunn A. D. Transformations and random variate generation: generalised ratio-of-uniforms methods. *J. Stat. Comput. Simul.* **55**(1), 49–55 (1996).
- [7] Kinderman A. J. y Monahan J. F. Computer generation of random variables using the ratio of uniform deviates. *ACM Trans. Math. Softw.* **3** 257–260 (1977).
- [8] Liang F., Liu C. y Carrol R.J. *Advanced Markov chain Monte Carlo methods*. Wiley (2010).

- [9] Marsaglia G. Ratios of normal variables and ratios of sums of uniforms variables. *Am. Stat. Assoc.* **60**(309), 193-204 (1965).
- [10] Martino L., Luengo D. y Míguez J. *Independent random sampling methods*. Springer (2018).
- [11] Sheldon M. R. *A first course in probability*. 8th ed. Pearson (2010).
- [12] Stefanescu S. y Vaduva I. On computer generation of random vectors by transformations of uniformly distributed vectors. *Computing* **39**, 141-153 (1987).
- [13] Wakefield J. C., Gelfand A. E. y Smith A. F. M. Efficient generation of random variates via the ratio-of-uniforms method. *Stat. Comput.* **1**(2), 129–133 (1991).