“PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ ”

ESCUELA DE POSGRADO

“Distribución Uniforme sobre la Intersección de un Simplex y una Esfera en Dimensiones Altas

TESIS PARA OPTAR EL GRADO ACADÉMICO DE MAGISTER EN MATEMÁTICAS

AUTOR:

Wilson Alberto Cabanillas Banda

ASESOR:

Dr. Johel Victorino Beltrán Ramírez

Octubre, 2017

PERÚ

Resumen

La presente tesis es acerca de deducir propiedades asintóticas acerca de la distribución uniforme sobre la intersección de una esfera y un simplex en $\mathbb{R}^n$ cuando la dimensión del espacio euclideano tiende a infinito. Claramente, para que tal intersección sea no vacía es necesario que los tamaños de la esfera y el simplex, que también haremos crecer al infinito, sean configurados de modo adecuado (esto es discutido con detalle en el Lema 2.1). El resultado importante de este trabajo es que, de acuerdo a la "*razón asintótica*" entre los tamaños de la esfera y el simplex, la distribución uniforme sobre la intersección de ellos se comportará de modos absolutamente distintos.

Para dar una idea aproximada del resultado que conseguiremos podemos explicarlo del siguiente modo: Si n es muy grande y

$$(X_1, \ldots, X_n)$$

es un punto elegido uniformemente sobre la intersección de una esfera (euclideana) de radio $\sqrt{nb}$ y un simplex de radio n (respecto a la norma de la suma) en $\mathbb{R}^n$ entonces

> (i) Para $1 < b < 2$ el tamaño de cada componente $|X_j|$ es de orden menor o igual a $\sqrt{\log(n)}$ *(en particular, no existe una componente notablemente mayor que las demás).*
>
> (ii) Para $b > 2$, existe una componente del vector cuyo tamaño es de orden $\sqrt{n}$ mientras que el tamaño del resto de componentes es de orden estrictamente menor.

Los enunciados precisos de estas afirmaciones son los Teoremas 2.3 y 2.4 de la Sección 2.2. Estos teoremas incluyen también el resultado de lo que sucede en el *valor crítico* $b = 2$.

Agradecimiento

Agradezco a Dios el Creador del universo, por la vida, la salud y por haberme permitido hacer esta bella y apasionante maestría en matemáticas en la PUCP.

Agradezco a mis padres y a mi hermana por su respaldo tan imprescindible y esencial en todo aspecto de mi vida.

Agradezco a mi asesor Johel Beltrán por su inmenso apoyo incondicional, paciencia y por todo el tiempo que me brindó durante este largo trayecto hacia la graduación.

Índice general

Capítulo 1

Introducción

El objetivo de este trabajo es deducir propiedades asintóticas acerca de la distribución uniforme sobre la intersección de una esfera y un simplex en $\mathbb{R}^n$ cuando la dimensión del espacio euclideano tiende a infinito. Claramente, para que tal intersección sea no vacía es necesario que los tamaños de la esfera y el simplex, que también haremos crecer al infinito, sean configurados de modo adecuado (esto es discutido con detalle en el Lema 2.1). El resultado importante de este trabajo es que, de acuerdo a la "*razón asintótica*" entre los tamaños de la esfera y el simplex, la distribución uniforme sobre la intersección de ellos se comportará de modos absolutamente distintos.

Para dar una idea aproximada del resultado que conseguiremos podemos explicarlo del siguiente modo: Si n es muy grande y

$$(X_1, \ldots, X_n)$$

es un punto elegido uniformemente sobre la intersección de una esfera (euclideana) de radio $\sqrt{nb}$ y un simplex de radio n (respecto a la norma de la suma) en $\mathbb{R}^n$ entonces

(*i*) Para $1 < b < 2$ el tamaño de cada componente $|X_j|$ es de orden menor o igual a $\sqrt{\log(n)}$ *(en particular, no existe una componente notablemente mayor que las demás)*.

(*ii*) Para $b > 2$, existe una componente del vector cuyo tamaño es de orden $\sqrt{n}$ mientras que el tamaño del resto de componentes es de orden estrictamente menor.

Los enunciados precisos de estas afirmaciones son los Teoremas 2.3 y 2.4 de la Sección 2.2. Estos teoremas incluyen también el resultado de lo que sucede en el *valor crítico* $b = 2$.

La principal motivación para este estudio proviene de versiones discretas no lineales de la ecuación de Schrödinger donde las componentes de $(X_1, \ldots, X_n)$ son interpretadas como cantidades de energía. El resultado presentado en este trabajo es entonces entendido como una *transición de fase* donde, por encima del valor crítico $b = 2$, la energía es localizada en una sóla componente (ver [8] para más detalles sobre la interpretación física del resultado).

Por otro lado, el comportamiento asintótico de la distribución uniforme sobre un simplex adecuadamente rescalado es un resultado bastante conocido en probabilidad: las coordenadas son asintóticamente independientes y tienen distribución aproximadamente exponencial. En el caso de una esfera, es también conocido que para dimensiones grandes, las coordenadas de un punto elegido uniformemente son asintóticamente independientes y tienen distribución aproximadamente gaussiana. Desde un punto de vista matemático, el resultado expuesto en este trabajo es entonces la sucesión natural a estos resultados obtenidos para la esfera y el simplex de forma individual. Para una revisión de estos resultados clásicos puede ver [4].

Este trabajo de tesis es basado en el artículo [3] que será próximamente publicado en *Journal of Topology and Analysis*. Durante la elobaración de la tesis mantuvimos comunicación con el autor, Sourav Chatterjee, a quien agradecemos mucho su atención a nuestras observaciones sobre su artículo.

Capítulo 2

Resultados Principales

El propósito de este capítulo es brindar los conceptos básicos necesarios para presentar con rigor, al final del capítulo, los resultados anunciados en la introducción. En la primera sección examinaremos la configuración correcta de los tamaños entre el simplex y la esfera de modo que la intersección no sea trivial. En beneficio de la exposición, hemos preferido postergar la explicación detallada de la distribución uniforme sobre la intersección al siguiente capítulo para enunciar los resultados principales del trabajo ya en este punto y motivar así la lectura.

2.1. Intersección de una esfera y un simplex

Escojamos un número real $b > 0$, un entero $n \geq 1$ y consideremos el conjunto

$$\Big\{(x_1, ..., x_n) \in \mathbb{R}^n \ : \ \sum_{i=1}^{n} |x_i| = n, \sum_{i=1}^{n} x_i^2 = nb\Big\}. \tag{2.1}$$

Este conjunto es la intersección de una esfera de radio n de acuerdo a la norma de la suma en $\mathbb{R}^n$ y de una esfera de radio $\sqrt{nb}$ de acuerdo a la norma euclideana en $\mathbb{R}^n$.

Por la simetría con respecto a los signos en la expresión que define al conjunto (2.1), para estudiar el comportamiento de la distribución uniforme sobre este conjunto, será suficiente considerar su restricción sobre el conjunto de vectores con coordenadas estrictamente positivas, i.e.

$$\mathcal{K} \ := \ \left\{(x_1, ..., x_n) \ : \ \sum_{i=1}^{n} x_i = n \,, \ \sum_{i=1}^{n} x_i^2 = nb \,, \ x_i > 0 \,, \ \forall i\right\} . \tag{2.2}$$

Antes de encontrar los valores de b para los cuales $\mathcal{K}$ es no vacío, consideremos el conjunto

$$\overline{\mathcal{K}} \ := \ \left\{(x_1, ..., x_n) \in \mathbb{R}_+^n \ : \ \sum_{i=1}^{n} x_i = n, \sum_{i=1}^{n} x_i^2 = nb\right\}$$

donde usamos la notación

$$\mathbb{R}_+^n \ := \ \{(x_1, \ldots, x_n) : x_i \geq 0 \,, \ \forall i\} \,. \tag{2.3}$$

El siguiente lema nos da una idea del aspecto geométrico de $\overline{\mathcal{K}}$ según los valores de b.

Lema 2.1. Se cumple lo siguiente.

(i) Si $\overline{\mathcal{K}}$ no es vacío entonces $b \in [1, n]$.

(ii) Si $b = 1$ entonces $\overline{\mathcal{K}} = \{\mathbf{1}\}$, donde $\mathbf{1} = (1, 1, ..., 1) \in \mathbb{R}^n$.

(iii) Si $b = n$ entonces $\overline{\mathcal{K}} = \{ne_1, ne_2, ..., ne_n\}$, donde $e_1, e_2, ..., e_n$ son los vectores canónicos del $\mathbb{R}^n$.

Demostración. Probemos (i). Ya que la aplicación $t \mapsto t^2$ es convexa tenemos la desigualdad

$$\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)^2 \leq \frac{1}{n}\sum_{i=1}^{n} x_i^2 \,.$$

Para que $\overline{\mathcal{K}}$ sea no vacío entonces, debido a esta desigualdad, resulta necesario que $1 \leq b$. Por otro lado, si $x_i \geq 0$ para todo $i = 1, \ldots, n$ entonces

$$\sum_{i=1}^{n} x_i^2 \leq \left(\sum_{i=1}^{n} x_i\right)^2 .$$

Ya que las componentes de los vectores en $\overline{\mathcal{K}}$ son no negativos, esta desigualdad implica que $b \leq n$ es necesario para que $\overline{\mathcal{K}}$ sea no vacío.

Para probar (ii) supongamos que $b = 1$ y sea $(x_1, ..., x_n) \in \overline{\mathcal{K}}$. Tenemos,

$$\begin{aligned}(x_1 - 1)^2 + \cdots + (x_n - 1)^2 &= x_1^2 + \cdots + x_n^2 - 2(x_1 + \cdots + x_n) + n \\ &= 0.\end{aligned}$$

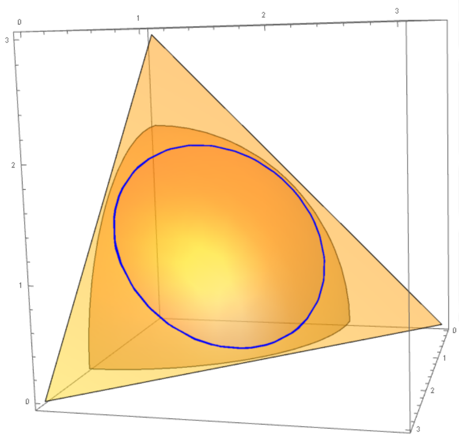
Eso implica que $x_i = 1$ para todo i y así concluimos que $\overline{\mathcal{K}} = \{\mathbf{1}\}$. Supongamos ahora que $b = n$ para probar (iii). Si $(x_1, ..., x_n) \in \overline{\mathcal{K}}$ tenemos

$$\begin{aligned}n^2 = (x_1 + \cdots + x_n)^2 &= x_1^2 + \cdots x_n^2 + 2 \sum_{1\leq i<j\leq n} x_i x_j \\ &= n^2 + 2 \sum_{1\leq i<j\leq n} x_i x_j.\end{aligned}$$

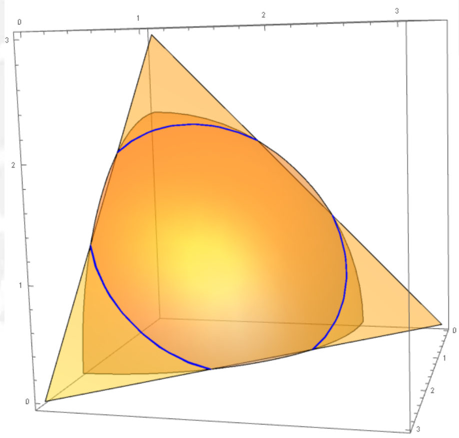
y por tanto $\sum_{1\leq i<j\leq n} x_i x_j = 0$. Siendo todas las componentes mayores o iguales a cero concluimos que

$$x_i x_j = 0, \quad \forall i \neq j$$

eso prueba que no pueden haber dos componentes estrictamente mayores que cero y por lo tanto $(x_1, ..., x_n)$ es múltiplo de algún vector canónico de $\mathbb{R}^n$. A partir de aquí es simple concluir que $\overline{\mathcal{K}} = \{ne_j : j = 1, \ldots, n\}$. $\square$



(a)



(b)

Figura 2.1: El conjunto $\mathcal{K} \subseteq \mathbb{R}^3$ para distintos valores de b. Observe que en (a) $\mathcal{K}$ resulta conexo mientras que en (b) resulta disconexo.

Lema 2.2. $\mathcal{K}$ es no vacío si y sólo si $1 \leq b < n$.

Demostración. Gracias al Lema 2.1, para que $\mathcal{K}$ no sea vacío es necesario que $1 \leq b < n$ y cuando $b = 1$ tendremos $\mathcal{K} = \{\mathbf{1}\}$. Para probar la recíproca asumamos ahora que $1 \leq b < n$. Definamos $\gamma : [0,1] \to \mathbb{R}^n_+$ como

$$\gamma(t) = (1-t)ne_1 + t\mathbf{1} \ , \quad t \in [0,1] \ , \tag{2.4}$$

donde, recordemos, e_1 es el primer vector canónico de $\mathbb{R}^n$ y $\mathbf{1} = (1,1,...,1) \in \mathbb{R}^n$. Es simple ver que

$$\gamma(t) \in \left\{(x_1,...,x_n) \in \mathbb{R}^n_+ \ : \ \sum_{i=1}^{n} x_i = n\right\} \tag{2.5}$$

para todo $t \in [0,1]$. Por otro lado, si

$$\gamma(t) = (\gamma_1(t), \gamma_2(t), \ldots, \gamma_n(t))$$

entonces

$$\sum_{i=1}^{n} \gamma_i(0)^2 = n^2 \quad \text{y} \quad \sum_{i=1}^{n} \gamma_i(1)^2 = n \ .$$

Ya que la aplicación $t \mapsto \sum_{i=1}^{n} \gamma_i(t)^2$ es continua, el teorema del valor intermediario nos permite asegurar la existencia de $t_0 \in [0,1]$ de modo que

$$\sum_{i=1}^{n} \gamma_i(t_0)^2 = bn \ . \tag{2.6}$$

Como $b < n$, esta última exresión es menor que n^2 y entonces $t_0 \neq 0$. Usando este hecho en (2.4) aseguramos que

$$\gamma_i(t_0) \ \geq \ t_0 \ > \ 0 \ , \quad \forall i = 1, \ldots, n \ . \tag{2.7}$$

De (2.5), (2.6) y (2.7) concluimos que $\gamma(t_0) \in \mathcal{K}$ probando que $\mathcal{K}$ es no vacío. $\square$

En este trabajo se fijará $b > 1$ para evitar el caso degenerado en que $\mathcal{K} = \{\mathbf{1}\}$. En los resultados que enunciaremos en la siguiente sección haremos que la dimensión $n \uparrow \infty$ mientras que $b > 1$ permanece fijo. Entonces, por el Lema 2.2, sabremos que $\mathcal{K} \neq \varnothing$ para todos los valores de n a partir de una cierta dimensión.

2.2. Teoremas Principales

Estamos interesados en observar el comportamiento asintótico de la distribución uniforme sobre $\mathcal{K}$ cuando fijamos el valor de $b > 1$ y hacemos $n \to \infty$. La definición precisa de *distribución uniforme* sobre $\mathcal{K}$ no resulta evidente debido a que $\mathcal{K}$ tiene medida de Lebesgue igual a cero. Para no distraernos con el tecnicismo de tal definición, la postergaremos al Capítulo 3 donde además probaremos propiedades importantes de simetría de la distribución. Enunciaremos más bien ahora los resultados principales del trabajo.

Durante el resto del texto,

$$X \ = \ (X_1, ..., X_n)$$

denotará un vector aleatorio siguiendo la distribución uniforme sobre $\mathcal{K}$. El primer teorema cubre el rango de valores menores o iguales al valor crítico.

Teorema 2.3. *Suponga $1 < b \leq 2$. Entonces existe un único par $(r, s) \in \mathbb{R}^2$ de modo que la densidad de probabilidad que es proporcional a*

$$\begin{cases} \exp(-rx^2 - sx)\ , & x \geq 0\ , \\ 0\ , & x < 0\ , \end{cases}$$

tiene primer momento 1 y segundo momento b. Sean $Y_1, Y_2, \ldots$ variables aleatorias i.i.d siguiendo esta densidad. Entonces se cumple :

(*a*) *Para todo k fijo, el vector aleatorio $(X_1, \ldots, X_k)$ de las primeras k coordenadas de X converge en distribución a $(Y_1, \ldots, Y_k)$ cuando la dimensión $n \to \infty$.*

(*b*) *Además, todos los momentos de $(X_1, \ldots, X_k)$ convergen a los correspondientes momentos de $(Y_1, \ldots, Y_k)$.*

(*c*) *Si $b < 2$, existe una constante C, dependiendo de b, tal que*

$$\lim_{n\to\infty} \mathbb{P}\Big(\max_{1\leq i\leq n} X_i > C\sqrt{\log n}\Big) = 0.$$

(*d*) *Cuando $b = 2$, la parte (c) se cumple pero con $\log n$ en vez de $\sqrt{\log n}$*

El siguiente teorema describe la situación cuando b está por encima del valor crítico. Un interesante fenómeno de localización ocurre en este rango de valores.

Teorema 2.4. *Supongamos que $b > 2$. Sean $Y_1, Y_2, \ldots$ variables aleatorias independientes tales que*

$$Y_j \;\sim\; exp(1)\;, \quad \forall j \geq 1\;.$$

Entonces se cumple lo siguiente.

(a) *Para cualquier k fijo, el vector aleatorio $(X_1, \ldots, X_k)$ converge en distribución a $(Y_1, \ldots, Y_k)$ cuando la dimensión $n \to \infty$.*

(b) *Si $M := \max_{1\leq i\leq n} X_i$ entonces*

$$\frac{M^2}{(b-2)n} \to 1 \textit{ en probabilidad.}$$

(c) *Sea M_2 el valor de la segunda coordenada más grande en $X = (X_1, \ldots, X_n)$. Entonces*

$$\frac{M_2^2}{n} \to 0 \textit{ en probabilidad.}$$

A diferencia de lo enunciado en el Teorema 2.3, cuando $b > 2$ no ocurre la convergencia de momentos. En efecto, probaremos en la Proposición 3.11 del siguiente capítulo que las coordenadas de $X = (X_1, X_2, \ldots, X_n)$ tienen la misma distribución. Por lo tanto

$$\mathbb{E}\left(|X_1|^2 + |X_2|^2 + \cdots + |X_n|^2\right) \;=\; nb$$

implica que $\mathbb{E}\big(|X_1|^2\big) = b$ para cualquier n mientras que $\mathbb{E}\left(|Y_1|^2\right) = 2$.

El teorema final 2.5 provee cotas de error para los resultados de convergencia distribucional enunciados en los Teoremas 2.3 y 2.4.

Teorema 2.5. *En las hipótesis del Teorema 2.3,*

$$\sup_{t_1,\ldots,t_k} \big|\mathbb{P}(X_1 \leq t_1, \ldots, X_k \leq t_k) - \mathbb{P}(Y_1 \leq t_1, \ldots, Y_k \leq t_k)\big| \leq Ck\sqrt{\frac{\log n}{n}}.$$

donde C es una constante que depende solamente de b. En las hipótesis del Teorema 2.4, la cota del lado derecho se convierte en $Ckn^{-1/4}$.

Capítulo 3

Distribución Uniforme sobre $\mathcal{K}$

En este capítulo presentaremos la definición de distribución uniforme sobre $\mathcal{K}$ que estaremos usando durante el texto. Probaremos también, al final del capítulo, algunas propiedades de simetría sobre esta distribución.

Consideraremos sobre $\mathbb{R}^n$ el producto interno euclideano

$$\langle x\,, y\,\rangle = \sum_{j=1}^{n} x_j y_j\ , \quad \text{para } x = (x_1, \ldots, x_n) \text{ e } y = (y_1, \ldots, y_n) \tag{3.1}$$

que determina la norma $\|x\| := \langle x, x\rangle^{1/2}$. Definamos para cada $x = (x_1, \ldots, x_n)$ la cantidad

$$\mu(x) := \frac{1}{n}\sum_{i=1}^{n} x_i \ = \ \frac{1}{n}\langle x, \mathbf{1}\rangle. \tag{3.2}$$

Recuerde que $\mathbf{1}$ representa el vector de coordenadas iguales a 1. Consideremos el hiperplano

$$V := \{x \in \mathbb{R}^n : \mu(x) = 0\}. \tag{3.3}$$

Dotaremos al espacio vectorial V del producto interno $\langle\cdot\,,\cdot\rangle$ inducido por $\mathbb{R}^n$ (haremos abuso de notación usando el mismo símbolo para el producto interno sobre V). Estaremos interesados en la esfera de radio $\sqrt{n}$ del espacio V:

$$\mathbb{S}_V := \{x \in V : \langle x, x\rangle = n\}\ . \tag{3.4}$$

Más adelante, en (3.17), identificaremos a $\mathcal{K}$ con el subconjunto $T^{-1}(\mathcal{K})$ de la esfera $\mathbb{S}_V$ a través de una homotecia T. Así, definiremos finalmente la distribución uniforme sobre

$\mathcal{K}$ por medio de la distribución uniforme sobre $\mathbb{S}_V$ condicionada a $T^{-1}(\mathcal{K})$. De acuerdo a este plan, comenzaremos examinando la distribución uniforme sobre $\mathbb{S}_V$ en la siguiente sección.

3.1. Distribución uniforme sobre $\mathbb{S}_V$

Acabamos de definir el espacio vectorial con producto interno $(V, \langle\cdot,\cdot\rangle)$ usando (3.3) y el producto interno euclideano en $\mathbb{R}^n$. Recordemos que una aplicación $A : V \to V$ es una *isometría lineal* cuando A es lineal y

$$\langle Ax, Ay\rangle = \langle x, y\rangle \ , \quad \text{para todo } x, y \in V .$$

Es claro que toda isometría lineal A es no singular, es decir, $Ax = 0$ si y sólo si $x = 0$, y su inversa A^{-1} es también una isometría lineal. De forma evidente, la misma definición y propiedades valen para isometrías lineales definidas sobre el espacio $(\mathbb{R}^n, \langle\cdot,\cdot\rangle)$.

Denotaremos por $\mathcal{B}_{\mathbb{R}^n}$ (resp. $\mathcal{B}_V$) el σ-álgebra de borelianos en $\mathbb{R}^n$ (resp. V). Es sabido que $\mathcal{B}_V$ coincide con la colección de borelianos de $\mathbb{R}^n$ contenidos en V. La distribución uniforme sobre $\mathbb{S}_V$ es definida como la única probabilidad sobre $(V, \mathcal{B}_V)$ con soporte en $\mathbb{S}_V$ y que es invariante bajo todas las isometrías lineales de V sobre sí mismo.

Definición 3.1. *La distribución uniforme sobre $\mathbb{S}_V$, que denotaremos por λ, es definida como la única medida de probabilidad sobre $(V, \mathcal{B}_V)$ tal que $\lambda(\mathbb{S}_V) = 1$ y*

$$\lambda = \lambda \circ A^{-1} \quad \textit{para toda } A : V \to V \textit{ isometría lineal.} \tag{3.5}$$

En la definición estamos denotando por $\lambda \circ A^{-1}$ la medida inducida por λ y A:

$$\lambda \circ A^{-1}(B) = \lambda\big(A^{-1}(B)\big) \ , \quad B \in \mathcal{B}_V \ .$$

La prueba de la existencia y de la unicidad aludidas en la Definición 3.1 se puede encontrar en el Teorema 9.2 en el Capítulo de Anexos.

Recuerde la definición de $\mu(\cdot)$ en (3.2) y que $\mathbf{1} \in \mathbb{R}^n$ es el vector de coordenadas iguales a 1. El siguiente lema nos será útil posteriormente, en las pruebas de la Proposición 3.3 y del Lema 3.4.

Lema 3.2. Si $A : \mathbb{R}^n \to \mathbb{R}^n$ es una isometría lineal tal que $A\mathbf{1} = \mathbf{1}$ entonces

$$\mu(Ax) = \mu(x) \ , \quad \forall x \in \mathbb{R}^n \ .$$

En particular, $AV = V$.

Demostración. Para probar la primera afirmación basta observar que

$$\begin{aligned} \mu(Ax) &= \frac{1}{n}\langle \mathbf{1}, Ax\rangle \\ &= \frac{1}{n}\langle A\mathbf{1}, Ax\rangle \\ &= \frac{1}{n}\langle \mathbf{1}, x\rangle \\ &= \mu(x) \ . \end{aligned}$$

Ahora, si $y \in AV$ y por tanto $y = Ax$ para algún $x \in V$ tendremos que $\mu(y) = \mu(Ax) = \mu(x) = 0$ probando que $y \in V$. Eso prueba que $AV \subseteq V$. Siendo A no singular, $dim(AV) = dim(V)$ y por tanto $AV = V$. □

La siguiente proposición nos permite caracterizar la distribución uniforme sobre $\mathbb{S}_V$ mediante la distribución inducida sobre $\mathbb{R}^n$ por la inclusión.

Proposición 3.3. *Sea ν una medida de probabilidad sobre $(V, \mathcal{B}_V)$ y sea $\hat{\nu}$ la medida de probabilidad sobre $(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n})$ inducida por ν y la función inclusión $i : V \to \mathbb{R}^n$, i.e.*

$$\hat{\nu}(B) := (\nu \circ i^{-1})(B) = \nu(B \cap V), \quad B \in \mathcal{B}_{\mathbb{R}^n}.$$

Entonces ν es la distribución uniforme sobre $\mathbb{S}_V$ si y sólo si $\hat{\nu}$ satisface las siguientes dos condiciones:

(i) $\hat{\nu}(\mathbb{S}_V) = 1$ *y*

(ii) *$\hat{\nu}$ es invariante por todas las isometrías lineales $A : \mathbb{R}^n \to \mathbb{R}^n$ tales que $A\mathbf{1} = \mathbf{1}$, es decir, $\hat{\nu} \circ A^{-1} = \hat{\nu}$.*

Demostración. Para comenzar supongamos que ν es la distribución uniforme sobre $\mathbb{S}_V$. Es claro que

$$\hat{\nu}(\mathbb{S}_V) \ = \ \nu(\mathbb{S}_V \cap V) \ = \ \nu(\mathbb{S}_V) \ = \ 1$$

probando que $\hat{\nu}$ satisface (i). Para probar que $\hat{\nu}$ satisface (ii) fijemos una isometría lineal $\hat{A} : \mathbb{R}^n \to \mathbb{R}^n$ tal que $\hat{A}\mathbf{1} = \mathbf{1}$. Gracias al Lema 3.2 podemos definir $A : V \to V$ como

$$Ax := \hat{A}x \ , \quad x \in V \ .$$

Es claro que A resulta una isometría y que, por definición, $\hat{A} \circ i = i \circ A$, es decir el siguiente diagrama conmuta

$$\begin{tikzcd} V \arrow[r, "i"] \arrow[d, "A"'] \arrow[dr, phantom, "\circlearrowleft"] & \mathbb{R}^n \arrow[d, "\hat{A}"] \\ V \arrow[r, "i"] & \mathbb{R}^n \end{tikzcd}$$

Así,

$$\begin{aligned} \hat{\nu} \circ \hat{A}^{-1} &= (\nu \circ i^{-1}) \circ \hat{A}^{-1} \\ &= (\nu \circ A^{-1}) \circ i^{-1} \\ &= \nu \circ i^{-1} \quad \text{(pues } \nu \text{ es, por definición, invariante por } A) \\ &= \hat{\nu}, \end{aligned}$$

lo cual prueba que $\hat{\nu}$ satisface (ii).

Ahora supongamos que $\hat{\nu}$ cumple las condiciones *(i)* y *(ii)*. Probaremos primeramente que $\nu(\mathbb{S}_V) = 1$. En efecto,

$$\begin{aligned} \nu(\mathbb{S}_V) &= \nu(\mathbb{S}_V \cap V) \\ &= \hat{\nu}(\mathbb{S}_V) \\ &= 1. \end{aligned}$$

Fijemos ahora una isometría lineal $A : V \to V$ arbitraria y probemos que ν es invariante por A. Si $\{v_1, v_2, ..., v_{n-1}\}$ es una base para V entonces $\{v_1, v_2, ..., v_{n-1}, \mathbf{1}\}$ es una base para $\mathbb{R}^n$ porque $\mathbf{1} = (1, 1, ..., 1) \notin V$. Así cada $x \in \mathbb{R}^n$ se escribe como $x = \sum_{i=1}^{n-1} \alpha_i v_i + \alpha_n \mathbf{1}$ y definimos $\hat{A} : \mathbb{R}^n \to \mathbb{R}^n$ como

$$\hat{A}x := \begin{cases} Ax & \text{si } x \in V \ , \\ \sum_{i=1}^{n-1} \alpha_i A v_i + \alpha_n \mathbf{1} & \text{si } x \notin V \ . \end{cases}$$

Obviamente $\hat{A}\mathbf{1} = \mathbf{1}$. Usando el hecho de que el vector $\mathbf{1}$ es ortogonal a V y que A es isometría lineal, no es difícil verificar que $\hat{A}$ resulta una isometría lineal. De (ii) se sigue que

$$\hat{\nu} = \hat{\nu} \circ \hat{A}^{-1} \tag{3.6}$$

Ya que $\nu(D) = \hat{\nu}(D)$ para cualquier $D \in \mathcal{B}_V$ entonces, para cualquier $B \in \mathcal{B}_V$, tenemos

$$\nu\big(A^{-1}(B)\big) \;=\; \hat{\nu}(A^{-1}(B)) \;=\; \hat{\nu}(\hat{A}^{-1}(B)) \tag{3.7}$$

donde usamos que $AV = V$ para la segunda igualdad. Por (3.6), el último término en (3.7) coincide con

$$\hat{\nu}(B) \;=\; \nu(B \cap V) \;=\; \nu(B)\;. \tag{3.8}$$

De (3.7) y (3.8) se sigue que $\nu \circ A^{-1} = \nu$. Ya podemos concluir que ν es la distribución uniforme sobre V y la prueba está completa. □

3.2. Aplicación de normalización

Para cada vector $x \in \mathbb{R}^n$ definimos

$$\mu_2(x) := \frac{1}{n}\langle x, x\rangle \quad \text{y} \quad \sigma(x) := \sqrt{\mu_2(x) - \mu(x)^2}\;.$$

Denotaremos por $\mathsf{L} := \{\alpha\mathbf{1},\ \alpha \in \mathbb{R}\}$ la *recta diagonal* del $\mathbb{R}^n$. Observe que

$$x \notin \mathsf{L} \iff \sigma(x) > 0 \tag{3.9}$$

y que

$$x \in \mathbb{S}_V \iff \mu(x) = 0 \text{ y } \sigma(x) = 1\;. \tag{3.10}$$

Consideramos la aplicación de normalización $\phi : \mathbb{R}^n \to V$ definida como

$$\phi(x) := \begin{cases} 0\;, & \text{si } x \in \mathsf{L}\;, \\ \frac{1}{\sigma(x)}(x - \mu(x)\mathbf{1})\;, & \text{si } x \notin \mathsf{L}\;. \end{cases} \tag{3.11}$$

Esta función es definida para tener

$$\mu(\phi x) = 0 \quad \text{y} \quad \sigma(\phi x) = 1 \quad \text{para todo } x \in \mathbb{R}^n \setminus \mathsf{L}\;. \tag{3.12}$$

Usaremos las siguientes propiedades acerca de ϕ.

Lema 3.4. Se cumple lo siguiente.

(i) Para toda isometría lineal $A : \mathbb{R}^n \to \mathbb{R}^n$ tal que $A\mathbf{1} = \mathbf{1}$ se tiene $\mu_2(Ax) = \mu_2(x)$ y

$$\phi(Ax) \;=\; A\phi(x)\;, \quad \forall x \in \mathbb{R}^n\;.$$

(ii) $\phi(x) = x$ para todo $x \in \mathbb{S}_V$.

(iii) $\phi(\mathbb{R}^n \setminus \mathsf{L}) = \mathbb{S}_V$.

Demostración. Para probar (i) recordemos del Lemma 3.2 que $\mu(Ax) = \mu(x)$. Luego, como $\mu_2(Ax) = \mu_2(x)$ es fácil verificar, tenemos $\sigma(Ax) = \sigma(x)$. Finalmente

$$\begin{aligned}\phi(Ax) &= \frac{1}{\sigma(Ax)}\left(Ax - \mu(Ax)\mathbf{1}\right)\\ &= \frac{1}{\sigma(x)}A\left(x - \mu(x)\mathbf{1}\right)\\ &= A\phi(x).\end{aligned}$$

El ítem (ii) es consecuencia directa de (3.10). De (3.12) y (3.10) se sigue que $\phi(x) \in \mathbb{S}_V$ para todo $x \in \mathbb{R}^n \setminus \mathsf{L}$. Entonces el ítem ($iii$) es consecuencia de este hecho y el ítem (ii). $\square$

El siguiente lema nos proveerá una forma útil de generar una distribución uniforme sobre $\mathbb{S}_V$.

Lema 3.5. Sea Z un vector aleatorio n-dimensional gaussiano estándar, entonces $\phi \circ Z$ estÃ¡ uniformemente distribuido sobre $\mathbb{S}_V$.

Demostración. Denotemos por θ la distribución de $\phi \circ Z$ y denotemos por γ la distribución gaussiana estándar n-dimensional, i.e.

$$\gamma := \mathbb{P} \circ Z^{-1}, \quad \theta := \mathbb{P} \circ (\phi \circ Z)^{-1} = \gamma \circ \phi^{-1}$$

Denotemos por $\hat{\theta}$ la medida inducida por θ y la inclusión $i : V \to \mathbb{R}^n$. Gracias a la Proposición 3.3, resta probar que $\hat{\theta}$ satisface (i) y (ii) de esa proposición.

Tenemos primero que

$$
\begin{aligned}
\hat{\theta}(\mathbb{S}_V) &= \theta(\mathbb{S}_V) \\
&= \gamma \circ \phi^{-1}(\mathbb{S}_V) \\
&= \gamma(\mathbb{R}^n \setminus \mathsf{L}) \,.
\end{aligned}
$$

En la tercera igualdad hemos usado el ítem (iii) del Lema 3.4. Ya que L tiene medida de Lebesgue cero y γ es absolutamente continua entonces el último término es igual a 1. Eso prueba que $\hat{\theta}$ satisface (i) de la Proposición 3.3.

Consideremos ahora una isometría lineal $A : \mathbb{R}^n \to \mathbb{R}^n$ tal que $A\mathbf{1} = \mathbf{1}$. Denotemos $\hat{\phi} = i \circ \phi$:

$$
\begin{tikzcd}
(\Omega, \mathbb{P}) \arrow[r, "Z"] & (\mathbb{R}^n, \gamma) \arrow[r, "\phi"] \arrow[dr, dotted, "\hat{\phi}:=i\circ\phi"'] & (V, \theta) \arrow[d, "i"] \\
& & (\mathbb{R}^n, \hat{\theta})
\end{tikzcd}
$$

Vemos que

$$
\begin{aligned}
\hat{\theta} \circ A^{-1} &= (\gamma \circ \phi^{-1} \circ i^{-1}) \circ A^{-1} \\
&= \gamma \circ (A \circ \hat{\phi})^{-1} \\
&= \gamma \circ (\hat{\phi} \circ A)^{-1} \qquad \text{(por } (i) \text{ de Lema 3.4)} \\
&= (\gamma \circ A^{-1}) \circ \hat{\phi}^{-1} \\
&= \gamma \circ \hat{\phi}^{-1} \qquad \text{(pues } \gamma \text{ es invariante por isometrías lineales)} \\
&= \hat{\theta}.
\end{aligned}
$$

Concluimos pues, gracias a la Proposición 3.3, que θ es la distribución uniforme λ sobre $\mathbb{S}_V$. □

3.3. La distribución uniforme sobre $\mathcal{K}$

En esta sección finalmente definimos la distribución uniforme sobre $\mathcal{K}$. Nos será necesario considerar los subconjuntos de $\mathbb{S}_V$ del siguiente tipo:

$$
\mathbb{S}_V(r) \;:=\; \{x \in \mathbb{S}_V \;:\; m(x) > r\} \,, \quad r \in \mathbb{R} \,,
$$

donde

$$m(x) \ := \ \min_{1\leq i\leq n} x_i \ .$$

El siguiente lema nos permitirá definir la distribución uniforme sobre estos subconjuntos. Recuerde que λ denota la distribución uniforme sobre $\mathbb{S}_V$.

Lema 3.6. Si $\mathbb{S}_V(r) \neq \varnothing$, entonces $\lambda\left(\mathbb{S}_V(r)\right) > 0$.

Demostración. Si $\mathbb{S}_V(r) \neq \varnothing$ entonces existe un $x_0 \in \mathbb{S}_V$ tal que $m(x_0) > r$. En particular, $\phi(x_0) = x_0$ por el ítem (ii) del Lema 3.4. Por tanto

$$m \circ \phi(x_0) \ = \ m(x_0) \ > \ r \ .$$

Siendo $m \circ \phi$ una aplicación continua en una vecindad de $\mathbb{S}_V$ existe entonces una bola B centrada en x_0 tal que

$$m(\phi(y)) > r \ , \quad \text{para todo } y \in B. \tag{3.13}$$

Usando este hecho tenemos

$$\{Z \in B\} \subseteq \{m(\phi(Z)) > r\}$$

para un vector aleatorio Z con distribución n-dimensional Gaussiana estándar. Entonces

$$\mathbb{P}\Big(m\big(\phi(Z)\big) > r\Big) \ \geq \ \mathbb{P}(Z \in B) \ > \ 0 \tag{3.14}$$

Por otro lado, como $\mathbb{P}(Z \in \mathsf{L}) = 0$ entonces, por el ítem (iii) del Lema 3.4 tenemos

$$\begin{aligned} \mathbb{P}\Big(m(\phi(Z)) > r\Big) \ &= \ \mathbb{P}\Big(m(\phi(Z)) > r, \ \phi(Z) \in \mathbb{S}_V\Big) \\ &= \ \mathbb{P}\Big(\phi(Z) \in \mathbb{S}_V(r)\Big) \ . \end{aligned}$$

Del Lema 3.5, de esta última observación y de (3.14) obtenemos

$$\lambda(\mathbb{S}_V(r)) \ = \ \mathbb{P}\Big(\phi(Z) \in \mathbb{S}_V(r)\Big) \ = \ \mathbb{P}\Big(m(\phi(Z)) > r\Big) \ > \ 0 \ ,$$

que es lo que queríamos probar. $\square$

Ahora estamos listos para definir la distribución uniforme sobre $\mathcal{K}$. Recuerde que estamos asumiendo $b > 1$. Denotaremos durante todo el resto del texto:

$$b' := \sqrt{b-1}\ . \tag{3.15}$$

Gracias a (3.10), tenemos

$$\begin{aligned}\mathbb{S}_V(-1/b') &= \{x \in \mathbb{S}_V \ :\ m(x) > -1/b'\} \\ &= \big\{x \in \mathbb{R}^n \ :\ \mu(x) = 0\ ,\ \ \sigma(x) = 1\ ,\ \ m(x) > -1/b'\big\}\end{aligned}$$

Observe por otro lado que $\mathcal{K}$ puede ser descrito del siguiente modo

$$\mathcal{K} \ := \ \big\{x \in \mathbb{R}^n \ :\ \mu(x) = 1\ ,\ \ \sigma(x) = b'\ ,\ \ m(x) > 0\big\}\ .$$

Entonces, si consideraremos la homotecia

$$T : V \to \mathbb{R}^n\ , \quad T(x) := b'x + \mathbf{1} \tag{3.16}$$

y denotamos $\mathbb{S}_V^* := \mathbb{S}_V(-1/b')$ tendremos

$$\mathcal{K} \ = \ T\big(\,\mathbb{S}_V^*\,\big)\ . \tag{3.17}$$

En particular esta relación prueba que $\mathbb{S}_V^* \neq \varnothing$ (a partir de una cierta dimensión, como fue comentado al final de la Sección 2.1). Gracias al Lema 3.6, podemos entonces definir la *distribución uniforme* sobre $\mathbb{S}_V^*$ de forma estándar.

Definición 3.7. *Definimos λ^*, la distribución uniforme sobre $\mathbb{S}_V^*$, como la probabilidad sobre $(V, \mathcal{B}_V)$ dada por*

$$\lambda^*(A) = \frac{\lambda(A \cap \mathbb{S}_V^*)}{\lambda(\mathbb{S}_V^*)}, \quad A \in \mathcal{B}_V.$$

Finalmente, debido a que T es una homotecia, resulta natural definir la distribución uniforme sobre $\mathcal{K}$ como la probabilidad inducida sobre ella por T y la distribución uniforme sobre $\mathbb{S}_V^*$.

Definición 3.8. *La distribución uniforme sobre $\mathcal{K}$, que denotaremos por $\lambda_{\mathcal{K}}$, es definida como*

$$\lambda_{\mathcal{K}} \ := \ \lambda^* \circ T^{-1} \tag{3.18}$$

3.4. Simetrías de la distribución uniforme sobre $\mathcal{K}$

Para terminar el capítulo, demostraremos la invarianza de $\lambda_{\mathcal{K}}$ por permutaciones de coordenadas. Una permutación de coordenadas es una aplicación $\Sigma : \mathbb{R}^n \to \mathbb{R}^n$ de la forma

$$\Sigma(x_1, x_2, ..., x_n) := (x_{\sigma(1)}, x_{\sigma(2)}, ..., x_{\sigma(n)}),$$

para alguna biyección $\sigma : \{1, 2, ..., n\} \to \{1, 2, ..., n\}$.

Lema 3.9. Para toda permutación de coordenadas $\Sigma : \mathbb{R}^n \to \mathbb{R}^n$ se cumple lo siguiente.

(i) Σ es una isometría lineal, $\Sigma(\mathbf{1}) = \mathbf{1}$ y $\Sigma(V) = V$.

(ii) $x \in \mathbb{S}^*_V$ si y sólo si $\Sigma(x) \in \mathbb{S}^*_V$.

Demostración. Las dos primera afirmaciones del ítem (i) son evidentes y de ellas se sigue que $\Sigma(V) = V$ gracias a la conclusión del Lema 3.2. Para (ii), es muy sencillo verificar que

$$x \in \mathbb{S}^*_V \implies \Sigma(x) \in \mathbb{S}^*_V \ .$$

Para probar la recíproca basta usar la función inversa Σ^{-1} que es también una permutación de coordenadas. □

Gracias al ítem (i) del lema anterior, cada permutación de coordenadas Σ se puede restringir a V:

$$\Sigma_V : V \to V \ . \tag{3.19}$$

Σ_V es claramente una isometría lineal y entonces

$$\lambda \circ \Sigma_V^{-1} = \lambda$$

por definición de λ. El ítem (ii) asegura que $\mathbb{S}^*_V$ resulta estable por Σ_V. Esto nos permite concluir que λ^*, la distribución uniforme sobre $\mathbb{S}^*_V$, es también invariante por la permutación de coordenadas Σ_V.

Proposición 3.10. *La probabilidad λ^* resulta invariante por Σ_V, i.e.*

$$\lambda^* \circ \Sigma_V^{-1} \ = \ \lambda^* \ ,$$

para toda permutación de coordenadas Σ.

Demostración. El ítem (ii) del Lema 3.9 nos dice que $\Sigma_V\left(\mathbb{S}_V^*\right)=\mathbb{S}_V^*$. Aplicando el Lema 9.4 del Anexo concluimos que $\lambda\mid_{\mathbb{S}_V^*}$ definido como

$$\lambda\mid_{\mathbb{S}_V^*}(A) \;=\; \lambda\big(A\cap\mathbb{S}_V^*\big)\;, \quad A\in\mathcal{B}_V$$

resulta una medida invariante por Σ_V. Siendo λ^* un múltiplo de $\lambda\mid_{\mathbb{S}_V^*}$ resulta también invariante por Σ_V. □

Teorema 3.11. *Sea $X=(X_1,X_2,...,X_n)$ un vector aleatorio con distribución $\lambda_{\mathcal{K}}$. Entonces el vector $\Sigma(X)$ tiene la misma distribución $\lambda_{\mathcal{K}}$ para cualquier permutación de coordenadas $\Sigma:\mathbb{R}^n\to\mathbb{R}^n$. En particular, las coordenadas de X son idénticamente distribuidas.*

Demostración. Fijemos una permutación de coordenadas $\Sigma:\mathbb{R}^n\to\mathbb{R}^n$. Es evidente que

$$\Sigma\circ T \;=\; T\circ\Sigma_V\;. \tag{3.20}$$

donde $T:V\to\mathbb{R}^n$ es la homotecia definida en (3.16). Luego

$$\begin{aligned}\lambda_{\mathcal{K}}\circ\Sigma^{-1} &= (\lambda^*\circ T^{-1})\circ\Sigma^{-1}\\ &= \lambda^*\circ(T\circ\Sigma_V)^{-1} \qquad \text{(usando (3.20))}\\ &= \lambda^*\circ T^{-1} \qquad \text{(usando Proposición 3.10)}\\ &= \lambda_{\mathcal{K}}.\end{aligned}$$

Para probar la última conclusión basta ver que X_j es la primera coordenada de $\Sigma(X)$ para una permutación conveniente. Así

$$\Sigma(X)\sim X \implies X_j\sim X_1\;.$$

□

Capítulo 4

Del conjunto $\mathcal{K}$ al abierto $\mathcal{K}^\epsilon$

Recuerde que X denota un vector aleatorio elegido uniformemente sobre

$$\mathcal{K} \ := \ \left\{ x \in \mathbb{R}^n \ : \ \mu(x) = 1 \ , \ \ \mu_2(x) = b \ , \ \ m(x) > 0 \right\}$$

donde hemos fijado el parámetro $b > 1$. Asumiremos siempre además en nuestras afirmaciones, sin hacer mención, que la dimensión n es suficientemente grande (que $n > b$) para que $\mathcal{K}$ resulte no vacío (ver Lema 2.2). Para cada $\epsilon > 0$, denotaremos ahora por X^ϵ un vector aleatorio elegido uniformemente sobre el conjunto

$$\mathcal{K}^\epsilon \ := \ \left\{ x \in \mathbb{R}^n : \epsilon < \mu(x) - 1 < 2\epsilon \ , \ \ \epsilon < \mu_2(x) - b < b\epsilon \ , \ \ m(x) > 0 \right\}.$$

En este capítulo veremos cómo podemos reemplazar X por X^ϵ en ciertas estimativas que nos serán importantes. Antes de eso, veamos que X^ϵ está bien definido. Observemos primero que $\mathcal{K}^\epsilon$ es no vacío para todo $\epsilon > 0$ suficientemente pequeño.

Lema 4.1. Existe una constante $c \in (0,1)$, que depende sólo de $b > 1$, de modo que $\mathcal{K}^\epsilon \neq \varnothing$ para todo $\epsilon \in (0, c)$.

Demostración. Ya que

$$\frac{b + \epsilon}{(1 + (3/2)\epsilon)^2} \ \to \ b \ , \quad \text{cuando } \epsilon \downarrow 0$$

existe $c > 0$ tal que

$$1 < \frac{b + \epsilon}{(1 + (3/2)\epsilon)^2} \ , \quad \forall \epsilon \in (0, c) \ . \tag{4.1}$$

Probaremos que c es la constante que buscamos. Ya que $b > 1$ tenemos $b + \epsilon < b(1 + \epsilon)$ y así

$$\frac{b+\epsilon}{(1+(3/2)\epsilon)^2} \;<\; \frac{b(1+\epsilon)}{(1+(3/2)\epsilon)^2} \;<\; b \;<\; n\,. \tag{4.2}$$

Gracias a las estimativas en (4.1) y (4.2) y gracias al Lema 2.2 podemos concluir que existe $z \in \mathbb{R}^n$ tal que $m(z) > 0$,

$$1 \;<\; \frac{b+\epsilon}{(1+(3/2)\epsilon)^2} \;<\; \frac{1}{n}\langle z, z\rangle \;=:\; \mu_2(z) \;<\; \frac{b(1+\epsilon)}{(1+(3/2)\epsilon)^2} \;<\; n$$

y $\mu(z) = 1$. Finalmente, haciendo $x = (1 + (3/2)\epsilon)z$ tendremos $x \in \mathcal{K}^\epsilon$ probando que $\mathcal{K}^\epsilon$ es no vacío. □

Llamaremos $c_0 > 0$ a la constante del Lema 4.1

Observemos que $\mathcal{K}^\epsilon$ resulta abierto porque $\mu(\cdot)$, $\mu_2(\cdot)$ y $m(\cdot)$ son funciones continuas. Además, la condición sobre $\mu_2(\cdot)$ en la definición de $\mathcal{K}^\epsilon$ asegura que $\mathcal{K}^\epsilon$ sea limitado. Concluimos así que

Lema 4.2. Para todo $\epsilon \in (0, c_0)$ la medida de Lebesgue de $\mathcal{K}^\epsilon$ es estrictamente positiva y finita.

Gracias a este hecho, la distribución uniforme sobre $\mathcal{K}^\epsilon$ es definida de forma usual como la restricción sobre $\mathcal{K}^\epsilon$ de la medida de Lebesgue en $\mathbb{R}^n$ y normalizada para ser probabilidad. Hasta aquí hemos probado que la siguiente definición tiene sentido.

Definición 4.3. *Para cada $\epsilon \in (0, c_0)$ vamos a denotar por X^ϵ un vector aleatorio con distribución uniforme sobre $\mathcal{K}^\epsilon$.*

Para valores de $\epsilon > 0$ próximos de cero, el conjunto $\mathcal{K}^\epsilon$ está de cierto modo próximo de $\mathcal{K}$. El conjunto $\mathcal{K}^\epsilon$ puede, por eso, ser visto como un "engrosamiento" de $\mathcal{K}$ y el vector X^ϵ como una perturbación de X. En la siguiente proposición vemos cómo podemos estimar la esperanza de funciones de X por expresiones en el vector X^ϵ. Recuerde la aplicación de normalización ϕ y la homotecia T definidas en el capítulo anterior. Nos resulta útil definir

$$\psi \;:=\; T \circ \phi$$

Proposición 4.4. *Existen constantes $c \in (0, c_0)$ y $C > 0$, que sólo dependen de b, de modo que para todo $\epsilon \in (0, c)$ y cualquier $f : \mathbb{R}^n \to [0, \infty)$ medible se tiene*

$$\mathbb{E}\, f(X) \;\leq\; \frac{\mathbb{E}\, f(\psi(X^\epsilon))}{\mathbb{P}(m(X^\epsilon) > C\epsilon)},$$

donde el lado derecho es interpretado como infinito si el denominador es cero.

Antes de probar esta proposición vamos a introducir algunos elementos y resultados previos. Para cada $\epsilon \in (0, c_0)$ definimos

$$\hat{\mathcal{K}}^\epsilon \;:=\; \big\{x \in \mathbb{R}^n : \epsilon < \mu(x) - 1 < 2\epsilon\,,\; \epsilon < \mu_2(x) - b < b\epsilon\big\},$$

de manera que $\mathcal{K}^\epsilon = \{x \in \hat{\mathcal{K}}^\epsilon : m(x) > 0\}$. Observe que $\hat{\mathcal{K}}^\epsilon$ es también un abierto, limitado no vacío. Consideremos la aplicación lineal afin $\hat{T} : \mathbb{R}^n \to \mathbb{R}^n$ definida como

$$\hat{T}(x) := b'x + \mathbf{1}\,, \quad x \in \mathbb{R}^n$$

de modo que la homotecia T es la restricción de $\hat{T}$ sobre V. Ya que $\hat{T}$ es una biyección, concluimos de (3.17) que

$$\mathbb{S}_V^* \;=\; \hat{T}^{-1}(\mathcal{K})\,.$$

Definamos

$$\hat{\mathbb{S}}^\epsilon := \hat{T}^{-1}(\hat{\mathcal{K}}^\epsilon) \quad \text{y} \quad \mathbb{S}^\epsilon := \hat{T}^{-1}(\mathcal{K}^\epsilon)$$

de modo que

$$\mathbb{S}^\epsilon \;=\; \{x \in \hat{\mathbb{S}}^\epsilon \,:\, m(x) > -1/b'\}\,.$$

Para una idea visual de estos conjuntos, $\mathbb{S}^\epsilon$ y $\hat{\mathbb{S}}^\epsilon$ pueden ser vistos como "engrosamientos"de $\mathbb{S}_V^*$ y $\mathbb{S}_V$, respectivamente.

Observe que $\hat{\mathbb{S}}^\epsilon$ es abierto, no vacío y limitado porque $\hat{\mathcal{K}}^\epsilon$ lo es y $\hat{T}$ conserva tales propiedades. Entonces la distribución uniforme sobre $\hat{\mathbb{S}}^\epsilon$ es la medida de Lebesgue de $\mathbb{R}^n$ restricta a $\hat{\mathbb{S}}^\epsilon$ y normalizada para ser probabilidad.

Lema 4.5. Para $\epsilon \in (0, c_0)$, si Y^ϵ es un vector aleatorio uniformemente distribuido sobre $\hat{\mathbb{S}}^\epsilon$ entonces $\phi(Y^\epsilon)$ está uniformemente distribuido sobre $\mathbb{S}_V$.

Demostración. Denotemos $\hat{\phi} := i \circ \phi$ donde $i : V \to \mathbb{R}^n$ es la inclusión de modo que

$$\hat{\phi} : \mathbb{R}^n \to \mathbb{R}^n \quad \text{con} \quad \hat{\phi}(x) = \phi(x)\,,\ \forall x\,.$$

Para probar que $\phi(Y^\epsilon)$ tiene distribución uniforme sobre $\mathbb{S}_V$ bastará probar que la distribución de $\hat{\phi}(Y^\epsilon)$ satisface las condiciones (i) y (ii) de la Proposición 3.3. Como L tiene medida de Lebesgue cero entonces $\mathbb{P}\{Y^\epsilon \in \mathsf{L}\} = 0$. Eso implica, por el tercer ítem del Lema 3.4, que

$$\mathbb{P}\big\{\hat{\phi}(Y^\epsilon) \in \mathbb{S}_V\big\} \;=\; 1\,.$$

Por lo tanto, la distribución de $\hat{\phi}(Y^\epsilon)$ satisface (i). Para probar que la distribución de $\hat{\phi}(Y^\epsilon)$ satisface (ii) fijemos una isometría lineal

$$A : \mathbb{R}^n \to \mathbb{R}^n \quad \text{tal que} \quad A\mathbf{1} = \mathbf{1}\,.$$

Del Lema 3.2 y del ítem (i) del Lema 3.4 tenemos

$$\mu(Ax) \;=\; \mu(x) \quad \text{y} \quad \mu_2(Ax) \;=\; \mu_2(x)\,, \quad \forall x \in \mathbb{R}^n\,.$$

Se sigue de inmediato que $A(\hat{\mathcal{K}}^\epsilon) \subseteq \hat{\mathcal{K}}^\epsilon$. Ya que A^{-1} es también una isometría lineal que cumple con $A^{-1}\mathbf{1} = \mathbf{1}$, podemos usar para ella el mismo argumento y concluir la inclusión contraria probando así que

$$\hat{\mathcal{K}}^\epsilon = A(\hat{\mathcal{K}}^\epsilon)\,. \tag{4.3}$$

Por otro lado, dado que $A\mathbf{1} = \mathbf{1}$, se sigue que A y $\hat{T}^{-1}$ conmutan:

$$A \circ \hat{T}^{-1}(x) \;=\; A((1/b')(x - \mathbf{1})) \;=\; (1/b')(Ax - \mathbf{1}) \;=\; \hat{T}^{-1} \circ A(x)\,.$$

Usando esa observación y (4.3) probamos que

$$A(\hat{\mathbb{S}}^\epsilon) \;=\; A \circ \hat{T}^{-1}(\hat{\mathcal{K}}^\epsilon) \;=\; \hat{T}^{-1} \circ A\,(\hat{\mathcal{K}}^\epsilon) \;=\; \hat{T}^{-1}(\hat{\mathcal{K}}^\epsilon) \;=\; \hat{\mathbb{S}}^\epsilon. \tag{4.4}$$

Gracias al Lema 9.4 del Anexo y a que la medida de Lebesgue es invariante por la isometría A, podemos concluir de (4.4) que la distribución uniforme sobre $\hat{\mathbb{S}}^\epsilon$ es invariante por A, es decir

$$A(Y^\epsilon) \;\sim\; Y^\epsilon\,. \tag{4.5}$$

Usando el primer ítem del Lema 3.4 tenemos

$$A \circ \hat{\phi}(Y^\epsilon) \;=\; \hat{\phi} \circ A(Y^\epsilon) \;\sim\; \hat{\phi}(Y^\epsilon)\,.$$

Eso prueba que la distribución de $\hat{\phi}(Y^\epsilon)$ satisface (ii) de la Proposición 3.3 y eso termina la prueba. □

Prueba de la Proposición 4.4. Hacemos

$$c \;:=\; \text{mín}\left\{(b-1)/7\,,\; c_0\,\right\}$$

de modo que si $\epsilon \in (0, c)$ tendremos en particular que $\mathcal{K}^\epsilon$ es no vacío y

$$\epsilon \;<\; \text{mín}\{1, (b-1)/7\} \tag{4.6}$$

porque $c_0 < 1$. Note que por definición de $\hat{\mathcal{K}}^\epsilon$, $x \in \hat{\mathcal{K}}^\epsilon$ implica que

$$\sigma(x) \quad < \quad \sqrt{b + b\epsilon - (1+\epsilon)^2} \quad < \quad \sqrt{(b-1)(1+\epsilon)^2} \quad = \quad b'(1 \;+\; \epsilon) \tag{4.7}$$

donde recuerde de (3.15) que $b' := \sqrt{b-1}$. También por definición de $\hat{\mathcal{K}}^\epsilon$, $x \in \hat{\mathcal{K}}^\epsilon$ implica que

$$\sigma(x) \;>\; \sqrt{b + \epsilon - (1+2\epsilon)^2} \;=\; \sqrt{b - 1 - 3\epsilon - 4\epsilon^2}\,.$$

Gracias a (4.6), la última expresión es mayor que

$$\sqrt{(b-1)\Big(1 - \frac{7\epsilon}{b-1}\Big)} \;>\; b'\Big(1 - \frac{7\epsilon}{b-1}\Big)\,.$$

Es decir, $x \in \hat{\mathcal{K}}^\epsilon$ implica también que

$$\sigma(x) \;>\; b'\Big(1 - \frac{7\epsilon}{b-1}\Big)\,. \tag{4.8}$$

Considere ahora un vector U uniformemente distribuido sobre $\mathbb{S}^*_V$ y Y^ϵ uniformemente distribuido sobre $\hat{\mathbb{S}}^\epsilon$. Se sigue del Lema 4.5 que para cualquier función $h : \mathbb{R}^n \to [0, \infty[$ medible

$$\mathbb{E}\, h(U) = \mathbb{E}\left[h(\phi(Y^\epsilon)) \mid m(\phi(Y^\epsilon)) > -1/b'\right]. \tag{4.9}$$

Es claro que

$$m(\phi(Y^\epsilon)) = \sigma(Y^\epsilon)^{-1}\big(m(Y^\epsilon) - \mu(Y^\epsilon)\big).$$

Así que

$$\{m(\phi(Y^\epsilon)) > -1/b'\} = \{m(Y^\epsilon) > \mu(Y^\epsilon) - (1/b')\sigma(Y^\epsilon)\}\,.$$

Usando esta observación en (4.9) tenemos

$$\mathbb{E}\,h(U) = \mathbb{E}\left[h(\phi(Y^\epsilon))\,|\,m(Y^\epsilon) > \mu(Y^\epsilon) - (1/b')\sigma(Y^\epsilon)\right]. \tag{4.10}$$

Ahora tome cualquier $y \in \hat{\mathbb{S}}^\epsilon$ y sea $x = \hat{T}(y)$. En particular, $x \in \hat{\mathcal{K}}^\epsilon$ y podemos usar las desigualdades (4.7) y (4.8) para $\sigma(x)$. Denotando

$$d \;=\; 1/b'$$

tenemos

$$\begin{aligned}
\mu(y) - d\sigma(y) &= \mu(\hat{T}^{-1}(x)) - d\sigma(\hat{T}^{-1}(x)) \\
&= \mu(d(x-\mathbf{1})) - d\sigma(d(x-\mathbf{1})) \\
&> d\epsilon - d^2 b'(1+\epsilon) \\
&= -d\,,
\end{aligned}$$

donde hemos usado (4.7) en la desigualdad. De lo anterior concluimos que

$$\{m(Y^\epsilon) > \mu(Y^\epsilon) - d\sigma(Y^\epsilon)\} \subset \{m(Y^\epsilon) > -d\}.$$

Usando este hecho en (4.10) obtenemos

$$\mathbb{E}\,h(U) \;=\; \mathbb{E}\left[h(\phi(Y^\epsilon))\,|\,m(Y^\epsilon) > \mu(Y^\epsilon) - d\sigma(Y^\epsilon),\; m(Y^\epsilon) > -d\right]\,. \tag{4.11}$$

Ahora, sea Z^ϵ uniformemente distribuida sobre $\mathbb{S}^\epsilon$. Luego, la ley de Z^ϵ es la ley de Y^ϵ condicionada al evento $\{m(Y^\epsilon) > -d\}$. Entonces de (4.11) se sigue que

$$\mathbb{E}\,h(U) \;=\; \mathbb{E}\left[h(\phi(Z^\epsilon))\,|\,m(Z^\epsilon) > \mu(Z^\epsilon) - d\sigma(Z^\epsilon)\right].$$

Dado que h es una función no negativa, se tiene que

$$\mathbb{E}\left[h(\phi(Z^\epsilon))\,|E\right] \;\leq\; \frac{1}{\mathbb{P}(E)}\mathbb{E}\left[h(\phi(Z^\epsilon))\right],$$

donde $E = \{m(Z^\epsilon) > \mu(Z^\epsilon) - d\sigma(Z^\epsilon)\}$. Así

$$\mathbb{E}\, h(U) \;\leq\; \frac{\mathbb{E}\, h(\phi(Z^\epsilon))}{\mathbb{P}\big(m(Z^\epsilon) > \mu(Z^\epsilon) - d\sigma(Z^\epsilon)\big)}, \tag{4.12}$$

(si el denominador es cero, interpretamos el lado derecho como infinito). Sin embargo, para cualquier $y = \hat{T}^{-1}(x) \in \hat{\mathbb{S}}^\epsilon$ se tiene que

$$\begin{aligned}
\mu(y) - d\sigma(y) &= d(\mu(x) - 1) - d^2\sigma(x) \\
&< 2d\epsilon - d^2 b'\Big(1 - \frac{7\epsilon}{b-1}\Big) \\
&= 2d\epsilon - d\Big(1 - \frac{7\epsilon}{(b')^2}\Big) \\
&= 2d\epsilon - d\Big(1 - \frac{7\epsilon}{(1/d)^2}\Big) \\
&= -d + (2d + 7d^3)\epsilon,
\end{aligned}$$

donde hemos usado (4.8) en la desigualdad. De esta estimativa tenemos que

$$\{m(Z^\epsilon) > -d + (2d + 7d^3)\epsilon\} \subset \{m(Z^\epsilon) > \mu(Z^\epsilon) - d\sigma(Z^\epsilon)\}. \tag{4.13}$$

Así de (4.12) y (4.13) se sigue que

$$\mathbb{E}\, h(U) \;\leq\; \frac{\mathbb{E}\, h(\phi(Z^\epsilon))}{\mathbb{P}\big(m(Z^\epsilon) > -d + (2d + 7d^3)\epsilon\big)}\,. \tag{4.14}$$

A partir de la definición de la distribución uniforme sobre $\mathcal{K}$, es claro que

$$X \sim T(U)$$

Por lo tanto, haciendo $h = f \circ \hat{T}$ en (4.14) tenemos

$$\mathbb{E}\, f(X) \;\leq\; \frac{\mathbb{E}\,[f \circ T(\phi(Z^\epsilon))]}{\mathbb{P}\big(m(Z^\epsilon) > -d + (2d + 7d^3)\epsilon\big)}\,. \tag{4.15}$$

Siendo $\hat{T}$ una homotecia tal que $\hat{T}(\mathbb{S}^\epsilon) = \mathcal{K}^\epsilon$ tenemos

$$\hat{T}(Z^\epsilon) \sim X^\epsilon\,. \tag{4.16}$$

Además es fácil verificar que

$$\phi \circ \hat{T}(x) \;=\; \phi(x)\,, \quad \forall x \in \mathbb{R}^n\,. \tag{4.17}$$

Usando (4.17) y (4.16) en (4.15) obtenemos

$$\mathbb{E}\, f(X) \le \frac{\mathbb{E}\,[f \circ T \circ \phi(X^\epsilon)]}{\mathbb{P}\big(m(Z^\epsilon) > -d + (2d + 7d^3)\epsilon\big)} \tag{4.18}$$

Por otro lado, $m(\hat{T}(Z^\epsilon)) = d^{-1}m(Z^\epsilon) + 1$, así que

$$\begin{aligned} \{m(Z^\epsilon) > -d + (2d + 7d^3)\epsilon\} \;&=\; \{d^{-1}m(Z^\epsilon) + 1 > (2 + 7d^2)\epsilon\} \\ &=\; \{m(\hat{T}(Z^\epsilon)) > (2 + 7d^2)\epsilon\}. \end{aligned}$$

Entonces

$$\begin{aligned} \mathbb{P}\{m(Z^\epsilon) > -d + (2d + 7d^3)\epsilon\} \;&=\; \mathbb{P}\{m(\hat{T}(Z^\epsilon)) > (2 + 7d^2)\epsilon\} \\ &=\; \mathbb{P}\{m(X^\epsilon) > (2 + 7d^2)\epsilon\}. \end{aligned} \tag{4.19}$$

Usando (4.19) en (4.18), se concluye que

$$\mathbb{E}\, f(X) \;\le\; \frac{\mathbb{E}\,[f \circ T \circ \phi(X^\epsilon)]}{\mathbb{P}\big(m(X^\epsilon) > C\epsilon\big)} \;=\; \frac{\mathbb{E}\,[f \circ \psi(X^\epsilon)]}{\mathbb{P}\big(m(X^\epsilon) > C\epsilon\big)}\,,$$

donde

$$C \;:=\; 2 + 7d^2 \;=\; 2 + \frac{7}{b-1}\,\cdot$$

Eso concluye la prueba. □

Las constantes de la Proposición 4.4 serán usadas sistemáticamente en el trabajo, por eso, a partir de ahora denotaremos como

$$c(b) \;:=\; c \;=\; \text{mín}\,\big\{(b-1)/7\,,\; c_0\,\big\}$$

y como

$$C(b) \;:=\; C \;=\; 2 + \frac{7}{b-1}$$

a dichas constantes.

Proposición 4.6. *Existe una constante $C_3(b) > 0$ dependiendo solamente de b de modo que para todo $\epsilon \in (0, c(b))$ y $g : \mathbb{R}^n \to \mathbb{R}$ para la cual existe $L > 0$ satisfaciendo*

$$|g(x) - g(y)| \leq L \max_{1\leq i\leq n} |x_i - y_i| \,, \quad \textit{para todo } x, y \in \mathbb{R}^n$$

se tiene

$$\mathbb{P}\big(|g(X) - a| > t\big) \leq \frac{\mathbb{P}\big(|g(X^\epsilon) - a| > t - C_3(b)L\epsilon n\big)}{\mathbb{P}\big(m(X^\epsilon) > C(b)\epsilon\big)},$$

para cualquier $a, t \in \mathbb{R}$.

Demostración. Tome $x \in \mathcal{K}^\epsilon$. De $y = \psi(x)$, se sigue que

$$\begin{aligned} |x_i - y_i| &= \Big|x_i - \frac{b'}{\sigma(x)}(x_i - \mu(x)) - 1\Big| \\ &= \Big|\frac{\sigma(x) - b'}{\sigma(x)}x_i + \frac{b'\mu(x) - \sigma(x)}{\sigma(x)}\Big| \\ &\leq \frac{|(\sigma(x) - b')x_i|}{\sigma(x)} + \frac{b'|\mu(x) - 1|}{\sigma(x)} + \frac{|b' - \sigma(x)|}{\sigma(x)} \\ &\leq C\,\epsilon(1 + x_i). \end{aligned} \tag{4.20}$$

para una constante C que sólo depende de b. La última desigualdad en (4.20) se sigue de (4.7), (4.8) y de que $\epsilon \in (0, c(b))$. Por otro lado,

$$x_i \leq \sum_{j=1}^n x_j = n\mu(x) \leq n(1 + 2\epsilon) \leq 3n. \tag{4.21}$$

Luego, tomando $C_3(b) = 4C$, tenemos

$$\begin{aligned} \max_{1\leq i\leq n} |x_i - y_i| &\leq C\epsilon(1 + 3n) \quad (\text{ gracias a (4.20) y (4.21) }) \\ &\leq C_3(b)\epsilon n. \end{aligned}$$

Así, para todo $x \in \mathcal{K}^\epsilon$

$$\begin{aligned} |g(x) - g(\psi(x))| &\leq L \max_{1\leq i\leq n} |x_i - y_i| \\ &\leq LC_3(b)\epsilon n. \end{aligned}$$

En particular, $|g(\psi(x)) - a| > t$ implica $|g(x) - a| > t - LC_3(b)\epsilon n$. Tomando $f(x) := 1_{\{|g(x)-a|>t\}}$, obtenemos por la Proposición 4.4 que

$$\begin{aligned}
\mathbb{P}(|g(X) - a| > t) &\leq \frac{\mathbb{P}(|g(\psi(X^\epsilon)) - a| > t)}{\mathbb{P}(m(X^\epsilon) > C(b)\epsilon)} \\
&\leq \frac{\mathbb{P}(|g(X^\epsilon) - a| > t - C_3 L\epsilon n)}{\mathbb{P}(m(X^\epsilon) > C(b)\epsilon)}
\end{aligned}$$

que es el resultado que deseábamos. □

En lo que resta del trabajo serán usadas las constantes

$$c(b)\ , \quad C(b) \quad \text{y} \quad C_3(b)$$

que han sido introducidas en este capítulo.

Capítulo 5

De $\mathcal{K}^\epsilon$ al vector Y

En este capítulo adaptaremos la Proposición 4.6 para que la distribución uniforme sobre $\mathcal{K}^\epsilon$ que aparece en la estimativa pueda ser reemplazada por la distribución de un vector aleatorio con coordenadas independientes, que llamaremos Y, condicionada al conjunto $\mathcal{K}^\epsilon$. Al final del capítulo mostraremos, en el Lema 5.4, un resultado asintótico sobre el vector Y que usaremos en la prueba de nuestros teoremas principales.

5.1. El vector Y en la Proposición 4.6

Definición 5.1. *Sea* $\mathcal{R} := \big(\,]0,+\infty[\times\mathbb{R}\big) \cup \big(\{0\}\times]0,+\infty[\,\big)$. *Para cada* $(r,s) \in \mathcal{R}$, *sea* $G_{r,s}$ *la distribución de probabilidad con densidad de probabilidad proporcional a* $\exp(-rx^2 - sx)$ *sobre* $]0,\infty[$.

Note que $(r,s) \in \mathcal{R}$ si y sólo si $\exp(-rx^2 - sx)$ es integrable en $]0,\infty[$. De ahora en adelante para simplificar el texto usaremos el siguiente abuso de lenguaje:

> *cuando digamos "para cualquier r,s" significará para cualquier (r,s) que pertenece a esta región admisible $\mathcal{R}$.*

En lo que sigue, $G_{r,s}^{\otimes n}$ denotará la medida producto sobre $\mathbb{R}^n$ cuyos factores son todos iguales a $G_{r,s}$. En particular, escribiremos $Y \sim G_{r,s}^{\otimes n}$ para querer decir que $Y = (Y_1, Y_2, \ldots, Y_n)$ es un vector aleatorio cuyas n componentes son variables i.i.d. con distribución común $G_{r,s}$.

El siguiente lema nos permitirá conseguir estimativas acerca de la distribución uniforme sobre $\mathcal{K}^\epsilon$ usando un vector del tipo $Y \sim G_{r,s}^{\otimes n}$.

Lema 5.2. Sea c_0 como en el Lema 4.1 y considere $Y \sim G_{r,s}^{\otimes n}$ para algún $(r,s) \in \mathcal{R}$. Para todo $\epsilon \in (0, c_0)$ y cualquier función medible $f : \mathcal{K}^\epsilon \to [0, \infty[$ tenemos

$$e^{-B\epsilon n}\mathbb{E}\, f(X^\epsilon) \leq \mathbb{E}\left(f(Y)|Y \in \mathcal{K}^\epsilon\right) \leq e^{B\epsilon n}\mathbb{E}\, f(X^\epsilon),$$

donde $B = 2br + 4|s|$.

Demostración. Usando la expresión para la función de densidad correspondiente a $G_{r,s}^{\otimes n}$ tenemos

$$\mathbb{E}\left(f(Y) \,|\, Y \in \mathcal{K}^\epsilon\right) \;=\; \frac{\int_{\mathcal{K}^\epsilon} f(x) e^{-rn\mu_2(x) - sn\mu(x)} dx}{\int_{\mathcal{K}^\epsilon} e^{-rn\mu_2(x) - sn\mu(x)} dx}\,. \tag{5.1}$$

Fijemos un $x \in \mathcal{K}^\epsilon$ arbitrario. De acuerdo a la expresión en (5.1) debemos encontrar cotas para la función de densidad $exp\{-rn\mu_2(x) - sn\mu(x)\}$ que puede ser escrita como

$$exp\{-rnb - sn\}\, exp\{-rn(\mu_2(x) - b) - sn(\mu(x) - 1)\} \tag{5.2}$$

Como $x \in \mathcal{K}^\epsilon$ tenemos

$$\epsilon < \mu_2(x) - b < b\epsilon \quad \text{y} \quad \epsilon < \mu(x) - 1 < 2\epsilon$$

de modo que el segundo factor en (5.2) tiene las siguientes cotas

$$\begin{aligned} &exp\{-rnb\epsilon - |s|n(2\epsilon)\} \\ &\qquad < \; exp\{-rn(\mu_2(x) - b) - sn(\mu(x) - 1)\} \\ &\qquad\qquad < \; exp\{-rn\epsilon + |s|n(2\epsilon)\}\,. \end{aligned} \tag{5.3}$$

Usando la notación $B = 2br + 4|s|$, podemos obtener de (5.3) las cotas

$$exp\{-\frac{1}{2}Bn\epsilon\} \;<\; exp\{-rn(\mu_2(x) - b) - sn(\mu(x) - 1)\} \;<\; exp\{\frac{1}{2}Bn\epsilon\}\,.$$

Usando esta estimativa para la expresión obtenida en (5.2) tenemos la estimativa que buscábamos para la densidad:

$$\begin{aligned} &exp\{-rnb - sn - \frac{1}{2}Bn\epsilon\} \;<\; exp\{-rn\mu_2(x) - sn\mu(x)\} \\ &\qquad\qquad < \; exp\{-rnb - sn + \frac{1}{2}Bn\epsilon\}\,. \end{aligned} \tag{5.4}$$

Luego usando (5.4) en (5.1) se sigue que

$$\mathbb{E}\left(f(Y)\,|\,Y \in \mathcal{K}^\epsilon\right) \;\geq\; \frac{\int_{\mathcal{K}^\epsilon} f(x)e^{-B\epsilon n/2}e^{-rbn-sn}dx}{\int_{\mathcal{K}^\epsilon} e^{B\epsilon n/2}e^{-rbn-sn}dx} \;=\; e^{-B\epsilon n}\mathbb{E}\, f(X^\epsilon)$$

y que

$$\mathbb{E}\left(f(Y)\,|\,Y \in \mathcal{K}^\epsilon\right) \;\leq\; \frac{\int_{\mathcal{K}^\epsilon} f(x)e^{B\epsilon n/2}e^{-rbn-sn}dx}{\int_{\mathcal{K}^\epsilon} e^{-B\epsilon n/2}e^{-rbn-sn}dx} \;=\; e^{B\epsilon n}\mathbb{E}\, f(X^\epsilon)$$

lo cual completa la prueba. □

Recuerde las constantes $c(b)$, $C(b)$ y $C_3(b)$ que fueron presentadas en el capítulo anterior.

Proposición 5.3. *Sea $Y \sim G_{r,s}^{\otimes n}$ para algún r, s. Para todo $\epsilon \in (0, c(b))$ y toda función g como en la Proposición 4.6 tenemos*

$$\mathbb{P}\big(|g(X) - a| > t\big) \;\leq\; e^{2B\epsilon n}\frac{\mathbb{P}\big(|g(Y) - a| > t - C_3(b)L\epsilon n, Y \in \mathcal{K}^\epsilon\big)}{\mathbb{P}\big(m(Y) > C(b)\epsilon, Y \in \mathcal{K}^\epsilon\big)},$$

para cualquier $a, t \in \mathbb{R}$, donde $B = 2br + 4|s|$.

Demostración. Haciendo $f := 1_A$ en el Lema 5.2 con $A = \{x : x \in \mathcal{K}^\epsilon, |g(x) - a| > t - C_3(b)L\epsilon n\}$ obtenemos de la primera desigualdad que

$$\begin{aligned} &e^{-B\epsilon n}\mathbb{P}\left(|g(X^\epsilon) - a| > t - C_3(b)L\epsilon n\right) \\ &\qquad\qquad \leq\; \mathbb{P}\left(|g(Y) - a| > t - C_3(b)L\epsilon n\,|\,Y \in \mathcal{K}^\epsilon\right) \end{aligned} \tag{5.5}$$

Haciendo ahora $f(x) := 1_{\{m(x) > C(b)\epsilon\}}$ en el Lema 5.2, obtenemos de la segunda desigualdad que

$$\mathbb{P}\left(m(Y) > C(b)\epsilon\,|\,Y \in \mathcal{K}^\epsilon\right) \;\leq\; e^{B\epsilon n}\mathbb{P}\left(m(X^\epsilon) > C(b)\epsilon\right) \tag{5.6}$$

Luego, usando las desigualdades (5.5) y (5.6) en la desigualdad obtenida en la Proposición 4.6 se sigue que

$$\begin{aligned} \mathbb{P}\big(|g(X) - a| > t\big) \;&\leq\; \frac{\mathbb{P}\big(|g(X^\epsilon) - a| > t - C_3(b)L\epsilon n\big)}{\mathbb{P}\big(m(X^\epsilon) > C(b)\epsilon\big)} \\ &\leq\; \frac{e^{B\epsilon n}\mathbb{P}\big(|g(Y) - a| > t - C_3(b)L\epsilon n\,|\,Y \in \mathcal{K}^\epsilon\big)}{e^{-B\epsilon n}\mathbb{P}\big(m(Y) > C(b)\epsilon\,|\,Y \in \mathcal{K}^\epsilon\big)} \\ &=\; e^{2B\epsilon n}\frac{\mathbb{P}\big(|g(Y) - a| > t - C_3(b)L\epsilon n, Y \in \mathcal{K}^\epsilon\big)}{\mathbb{P}\big(m(Y) > C(b)\epsilon, Y \in \mathcal{K}^\epsilon\big)} \end{aligned}$$

que es lo que queríamos probar. □

5.2. Un teorema local del límite central

Observe que si $W \sim G_{r,s}$ entonces, para cualquier $\alpha > 0$, tendremos $\alpha W \sim G_{r',s'}$ con ciertos $(r',s') \in \mathcal{R}$. Es posible por eso elegir $(r,s) \in \mathcal{R}$ de modo que $\mathbb{E}(Y_1) = 1$. Consideraremos ahora una sucesión de variables i.i.d. $Y_1, Y_2, \ldots$ todas con distribución común $G_{r,s}$ y de modo que

$$\mathbb{E}(Y_1, Y_1^2) = (1, \beta) \tag{5.7}$$

para algún $\beta > 0$. Para terminar este capítulo probaremos el siguiente teorema local del límite central para la sucesión de vectores aleatorios

$$(Y_n, Y_n^2), \quad n \geq 1 .$$

Lema 5.4. Suponga que para cada $n \in \mathbb{N}$ tenemos números reales $a_n \leq b_n$, $a'_n \leq b'_n$ tal que existen números reales $x_0, y_0 \in \mathbb{R}$, con

$$\begin{aligned} \lim_{n\to\infty} \sqrt{n}(a_n - 1) &= \lim_{n\to\infty} \sqrt{n}(b_n - 1) = x_0, \\ \lim_{n\to\infty} \sqrt{n}(a'_n - \beta) &= \lim_{n\to\infty} \sqrt{n}(b'_n - \beta) = y_0. \end{aligned}$$

Entonces

$$\lim_{n\to\infty} \frac{\mathbb{P}(a_n \leq \frac{1}{n}\sum_{i=1}^n Y_i \leq b_n, a'_n \leq \frac{1}{n}\sum_{i=1}^n Y_i^2 \leq b'_n)}{n(b_n - a_n)(b'_n - a'_n)} = \rho(x_0, y_0).$$

donde ρ es una densidad Gaussiana no degenerada sobre $\mathbb{R}^2$.

Para probar este lema usaremos el siguiente resultado más general cuya prueba se puede encontrar en [2], página 190.

Proposición 5.5. *Sea $\{U_n : n \geq 1\}$ una sucesión de vectores aleatorios i.i.d con valores en $\mathbb{R}^k$ y*

$$\mathbb{E}\, U_1 = 0, \qquad Cov(U_1) = \Sigma, \tag{5.8}$$

donde Σ es una matriz simétrica definida positiva. Sea Q_n la distribución de

$$n^{-1/2}(U_1 + U_2 + \cdots + U_n), \quad n \in \mathbb{N} .$$

Las siguientes afirmaciones son equivalentes :

(A) *Para todo n suficientemente grande, Q_n tiene densidad q_n y*

$$\lim_{n\to\infty} \sup_{x\in\mathbb{R}^k} |q_n(x) - \phi_{0,\Sigma}(x)| = 0, \tag{5.9}$$

donde $\phi_{0,\Sigma}$ es la densidad de una distribución normal con media 0 y covarianza Σ.

(B) *Existe un entero m tal que Q_m tiene densidad acotada (en casi todo punto).*

Queremos aplicar este resultado para

$$U_n := (Y_n - 1, Y_n^2 - \beta), \quad n \in \mathbb{N}. \tag{5.10}$$

Para ello, comenzaremos probando primero en el siguiente lema que U_1 tiene matriz de covarianza definida positiva.

Lema 5.6. La matriz $Cov(U_1)$ es definida positiva.

Demostración. Equivalentemente probaremos que la matriz de covarianza de (Y_1, Y_1^2) es definida positiva. Ya que $Cov(Y_1, Y_1^2)$ es una matriz simétrica semidefinida positiva es suficiente probar que

$$\det[Cov(Y_1, Y_1^2)] \neq 0 .$$

Supongamos que $\det[Cov(Y_1, Y_1^2)] = 0$. En ese caso tendríamos que

$$\mathbb{P}(Y_1 = aY_1^2 + b) = 1 . \tag{5.11}$$

Fijemos un $\delta \in (0, \frac{1}{2})$. Debido a la forma que tiene la función de densidad de Y_1, para cada $k \in \{1, 2, 3\}$ tenemos

$$0 < \mathbb{P}\big(k - \delta \leq Y_1 \leq k + \delta\big) = \mathbb{P}(E_k) , \tag{5.12}$$

donde

$$\begin{aligned} E_k &:= \big\{k - \delta \leq Y_1 \leq k + \delta\big\} \\ &= \big\{k - \delta \leq Y_1 \leq k + \delta,\ (k-\delta)^2 \leq Y_1^2 \leq (k+\delta)^2\big\} \end{aligned}$$

Por el supuesto (5.11) y por (5.12) tenemos

$$\mathbb{P}\big(E_k \cap \big\{Y_1 = aY_1^2 + b\big\}\big) > 0 . \tag{5.13}$$

Pero sobre el evento $E_k \cap \{Y_1 = aY_1^2 + b\}$ vemos que

$$k - \delta \ \leq \ Y_1 \ = \ aY_1^2 + b \ \leq \ a(k+\delta)^2 + b$$
$$\text{y que} \quad a(k-\delta)^2 + b \ \leq \ aY_1^2 + b \ = \ Y_1 \ \leq \ k + \delta \, .$$

Para que sea estrictamente positiva la probabilidad de la afirmación (5.13) se tendrá que cumplir entonces que

$$k - \delta \ \leq \ a(k+\delta)^2 + b \qquad \text{y que} \qquad a(k-\delta)^2 + b \ \leq \ k + \delta \, .$$

Ya que δ es arbitrariamente pequeño, haciendo $\delta \downarrow 0$, concluimos que

$$k = ak^2 + b \, .$$

Pero no existen a y b que puedan satisfacer esta ecuación para $k = 1$, $k = 2$ y $k = 3$. Concluimos que $\det[Cov(Y_1, Y_1^2)] \neq 0$ y que entonces la matriz $Cov(U_1)$ es definida positiva. $\square$

Probaremos ahora que el ítem (B) de la Proposición 5.5 se cumple para $m = 3$ para la sucesión U_n definida en (5.10).

Lema 5.7. El par $(Y_1 + Y_2 + Y_3, Y_1^2 + Y_2^2 + Y_3^2)$ tiene una densidad acotada en $\mathbb{R}^2$.

Demostración. Fijemos un par de números $u, v \in \mathbb{R}$ y un $\delta > 0$. Consideremos

$$A \ := \ \left\{(y_1, y_2, y_3) : |y_1 + y_2 + y_3 - u| < \delta \, , \ |y_1^2 + y_2^2 + y_3^2 - v| < \delta \right\} \, .$$

Ya que la densidad de probabilidad de (Y_1, Y_2, Y_3) es uniformemente limitada entonces

$$\mathbb{P}\big((Y_1, Y_2, Y_3) \in A\big) \ \leq \ C_0 \boldsymbol{m}^3(A) \tag{5.14}$$

donde $\boldsymbol{m}^3$ denota a la medida de Lebesgue de $\mathbb{R}^3$ y C_0 es una constante que sólo depende de r, s y por tanto sólo de b. Usando cálculo simple podemos probar que

$$\boldsymbol{m}^3(A) \ \leq \ C' \delta^2 \tag{5.15}$$

donde $C' > 0$ es una constante universal. Si definimos el evento

$$E_\delta \ := \ \Big\{|Y_1 + Y_2 + Y_3 - u| < \delta \, , \ |Y_1^2 + Y_2^2 + Y_3^2 - v| < \delta \Big\}$$

entonces juntando (5.14) y (5.15) tenemos

$$\mathbb{P}(E_\delta) \;\leq\; C_0' \, \boldsymbol{m}^2 \left(B_\delta(u,v)\right) \tag{5.16}$$

donde C_0' es una constante que sólo depende de b, $\boldsymbol{m}^2$ es la medida de Lebesgue en $\mathbb{R}^2$ y

$$B_\delta(u,v) \;:=\; \left\{(x,y) \in \mathbb{R}^2,\, |x-u| < \delta,\, |y-v| < \delta\right\}.$$

Dado que $E_\delta = \{(Y_1 + Y_2 + Y_3, Y_1^2 + Y_2^2 + Y_3^2) \in B_\delta(u,v)\}$, entonces hemos probado en (5.16) que

$$\mu\left(B_\delta(u,v)\right) \leq C_0' \, \boldsymbol{m}^2 \left(B_\delta(u,v)\right),$$

donde μ es la ley del vector $(Y_1 + Y_2 + Y_3, Y_1^2 + Y_2^2 + Y_3^2)$ y C_0' no depende de u ni de v. Ahora podemos aplicar finalmente el Lema 9.5 del Anexo para concluir que μ es absolutamente continua con respecto a $\boldsymbol{m}^2$ y que la derivada de Radon Nikodym

$$f \;=\; \frac{d\mu}{d\boldsymbol{m}^2}$$

es uniformemente acotada. □

Recuerde que r, s fueron fijados de modo que $\mathbb{E}\left(U_1\right) = (0,0)$ (ver (5.7)).

Lema 5.8. Para la sucesión de vectores aleatorios en $\mathbb{R}^2$

$$V_n \;:=\; n^{-1/2}\Big(\sum_{i=1}^{n}(Y_i - 1), \sum_{i=1}^{n}(Y_i^2 - \beta)\Big)\,, \quad n \geq 1$$

existe una densidad Gaussiana no degenerada ρ sobre $\mathbb{R}^2$ tal que si ρ_n es la densidad de probabilidad de V_n, entonces

$$\lim_{n\to\infty} \sup_{(x,y)\in\mathbb{R}^2} |\rho_n(x,y) - \rho(x,y)| = 0. \tag{5.17}$$

Demostración. Observe que

$$V_n \;=\; n^{-1/2}\sum_{i=1}^{n} U_i\,, \quad n \geq 1\,,$$

donde U_n es la sucesión fijada en (5.10). Gracias al Lema 5.6 podemos aplicar la Proposición 5.5 para la sucesión (U_n). El Lema 5.7 garantiza que se cumple el ítem (B) de la Proposición 5.5 para $m = 3$ y el ítem (A) nos permite concluir (5.17) como queríamos. □

La prueba del Lema 5.4 es ahora una consecuencia inmediata de este último resultado.

Prueba del Lema 5.4. Sean ρ_n como en el Lema 5.8 y

$$\begin{aligned} u_n &:= \sqrt{n}(a_n - 1) \quad ; \quad v_n := \sqrt{n}(b_n - 1), \\ u'_n &:= \sqrt{n}(a'_n - \beta) \quad ; \quad v'_n := \sqrt{n}(b'_n - \beta). \end{aligned}$$

Entonces

$$\begin{aligned} \mathbb{P}\Big(a_n \leq \frac{1}{n}\sum_{i=1}^{n} Y_i \leq b_n, a'_n \leq \frac{1}{n}\sum_{i=1}^{n} Y_i^2 \leq b'_n\Big) &= \mathbb{P}\left(V_n \in [u_n, v_n] \times [u'_n, v'_n]\right) \\ &= \int_{u_n}^{v_n}\int_{u'_n}^{v'_n} \rho_n(x,y)dydx. \end{aligned} \tag{5.18}$$

Sean ahora

$$\delta_n := \sup_{(x,y)\in\mathbb{R}^2} |\rho_n(x,y) - \rho(x,y)| \tag{5.19}$$

y

$$\tau_n := \sup_{u_n \leq x \leq v_n,\, u'_n \leq y \leq v'_n} |\rho(x,y) - \rho(x_0,y_0)|. \tag{5.20}$$

Entonces $\delta_n \to 0$ por el Lema 5.8, y $\tau_n \to 0$ debido a la continuidad de ρ. Finalmente, observe que

$$\begin{aligned} \left|\int_{u_n}^{v_n}\int_{u'_n}^{v'_n} \rho(x,y)dydx - \underbrace{n(b_n - a_n)(b'_n - a'_n)}_{=(v_n - u_n)(v'_n - u'_n)}\rho(x_0,y_0)\right| &= \\ \left|\int_{u_n}^{v_n}\int_{u'_n}^{v'_n} (\rho(x,y) - \rho(x_0,y_0))dydx\right| &\leq \\ \int_{u_n}^{v_n}\int_{u'_n}^{v'_n} \underbrace{|\rho(x,y) - \rho(x_0,y_0)|}_{\leq \tau_n \quad \text{gracias a (5.20)}} dydx &\leq \tau_n n(b_n - a_n)(b'_n - a'_n), \end{aligned}$$

y

$$\begin{aligned} \left|\int_{u_n}^{v_n}\int_{u'_n}^{v'_n} (\rho_n(x,y) - \rho(x,y))dydx\right| &\leq \int_{u_n}^{v_n}\int_{u'_n}^{v'_n} \underbrace{|\rho_n(x,y) - \rho(x,y)|}_{\leq \delta_n \quad \text{gracias a (5.19)}} dydx \\ &\leq \delta_n n(b_n - a_n)(b'_n - a'_n) \end{aligned}$$

De las dos últimas desigualdades dadas concluimos que

$$\left| \frac{\int_{u_n}^{v_n} \int_{u_n'}^{v_n'} \rho_n(x,y)dydx}{n(b_n - a_n)(b_n' - a_n')} - \rho(x_0, y_0) \right| \to 0 \qquad \text{cuando } n \to \infty,$$

y gracias a (5.18) esto completa la prueba. $\square$

Capítulo 6

Prueba del Teorema 2.3

En este capítulo terminaremos la prueba del Teorema 2.3. Comenzaremos probando el siguiente resultado sobre la familia de distribuciones $G_{r,s}$ dada en la definición 5.1.

Proposición 6.1. *Si* $1 < b \leq 2$ *entonces existen* $r, s \in \mathbb{R}$ *tal que la distribución de probabilidad* $G_{r,s}$ *tiene media 1 y segundo momento* b.

Para probar este resultado, definamos $\theta : \mathcal{R} \to \mathbb{R}$ de la siguiente manera

$$\theta(r,s) := \frac{\mathbb{E}\,(W^2)}{(\mathbb{E}\,W)^2}, \tag{6.1}$$

donde $W \sim G_{r,s}$ y recuerde del capítulo anterior que $\mathcal{R} = \big(\,]0,+\infty[\times\mathbb{R}\big)\cup\big(\{0\}\times]0,+\infty[\,\big)$.

Lema 6.2. La función θ definida anteriormente es continua en todo su dominio.

Demostración. Para $W \sim G_{r,s}$ se tiene que su densidad es $\rho_W(x) := k(r,s)e^{-rx^2-sx}$, donde $k(r,s) = 1/\int_0^\infty e^{-rx^2-sx}dx$, y que

$$\begin{aligned}\mathbb{E}\,W &= \int_0^\infty x\rho_W(x)dx = k(r,s)\int_0^\infty xe^{-rx^2-sx}dx \\ \mathbb{E}\,W^2 &= \int_0^\infty x^2\rho_W(x)dx = k(r,s)\int_0^\infty x^2e^{-rx^2-sx}dx\ .\end{aligned}$$

Entonces

$$\theta(r,s) = \frac{1}{k(r,s)}\frac{\int_0^\infty x^2e^{-rx^2-sx}dx}{\big(\int_0^\infty xe^{-rx^2-sx}dx\big)^2}\ .$$

Para probar que θ es continua, tomemos una sucesión arbitraria (r_n, s_n) en $\mathcal{R}$, tal que $(r_n, s_n) \to (r, s)$. Mostraremos primero que $k(\cdot,\cdot)$ es continua, es decir que

$$\lim_{n\to\infty} \int_0^\infty e^{-r_n x^2 - s_n x} dx = \int_0^\infty e^{-rx^2-sx} dx \tag{6.2}$$

En efecto, se sigue de inmediato que para todo $x \in [0, \infty[$

$$\lim_{n\to\infty} \underbrace{e^{-r_n x^2 - s_n x}}_{f_n(x)} = \underbrace{e^{-rx^2-sx}}_{f(x)}, \tag{6.3}$$

pues e^x es continua. Ahora acotemos $f_n(x) = e^{-r_n x^2 - s_n x}$ para usar el teorema de la convergecia dominada. Para ello consideraremos los casos: $r > 0, s \in \mathbb{R}$ y $r = 0, s > 0$ por separado.
Para el primer caso, es posible obtener un $n_0 \in \mathbb{N}$, tal que $\forall\, n > n_0$ se cumpla

$$\frac{r}{2} < r_n, \quad -s_n < |s| + 1.$$

Luego de lo anterior se sigue que

$$f_n(x) = e^{-r_n x^2 - s_n x} \leq \underbrace{e^{-\frac{r}{2}x^2 + (|s|+1)x}}_{= g_1(x) \quad \text{integrable}} . \tag{6.4}$$

Para el segundo caso, también es posible hallar un $n_1 \in \mathbb{N}$ tal que $\forall\, n > n_1$ se cumpla que $s_n > \frac{s}{2}$ y como $r_n \geq 0$ entonces se sigue que

$$f_n(x) = e^{-r_n x^2 - s_n x} \leq \underbrace{e^{-\frac{s}{2}x}}_{= g_2(x) \quad \text{integrable}} . \tag{6.5}$$

Luego de (6.3), (6.4), (6.5) y aplicando el Teorema de la Convegencia Dominada, se sigue el resultado deseado en (6.2). De forma muy similar podemos probar que

$$\begin{aligned} \lim_{n\to\infty} \int_0^\infty x e^{-r_n x^2 - s_n x} dx &= \int_0^\infty x e^{-rx^2-sx} dx \\ \lim_{n\to\infty} \int_0^\infty x^2 e^{-r_n x^2 - s_n x} dx &= \int_0^\infty x^2 e^{-rx^2-sx} dx. \end{aligned} \tag{6.6}$$

Por lo tanto de las ecuaciones (6.2) y (6.6) se sigue de inmediato que

$$\begin{aligned}
\lim_{n\to\infty} \theta(r_n, s_n) &= \lim_{n\to\infty} \frac{1}{k(r_n, s_n)} \frac{\int_0^\infty x^2 e^{-r_n x^2 - s_n x} dx}{\left(\int_0^\infty x e^{-r_n x^2 - s_n x} dx\right)^2} \\
&= \frac{1}{k(r, s)} \frac{\int_0^\infty x^2 e^{-r x^2 - s x} dx}{\left(\int_0^\infty x e^{-r x^2 - s x} dx\right)^2} \\
&= \theta(r, s).
\end{aligned}$$

Lo cual prueba la continuidad de θ. □

Prueba de la Proposición 6.1. Como ya fue observado antes, si $W \sim G_{r,s}$ entonces para cualquier $\alpha > 0$, $\alpha W \sim G_{r',s'}$ para ciertos r', s'. Por lo tanto, es suficiente mostrar que para cualquier $b \in]1, 2]$, existen r, s tal que si $W \sim G_{r,s}$ entonces

$$\theta(r, s) := \frac{\mathbb{E}\,(W^2)}{(\mathbb{E}\,(W))^2} = b. \tag{6.7}$$

Por el Lema 6.2 se tiene que θ es una función continua de r, s. Es fácil ver que $\theta(0, 1) = 2$, pues si $W \sim G_{0,1}$ entonces la densidad de W es

$$\rho_W(x) = \frac{e^{-x}}{\int_0^\infty e^{-x} dx} = e^{-x} , \quad x > 0$$

y así $\mathbb{E}\, W = 1$, $\mathbb{E}\, W^2 = 2$. Para cada $r > 0$, sean $W_r \sim G_{r,1}$ y $Z_r := \sqrt{r} W_r$ de modo que

$$\theta(r, 1) = \frac{\mathbb{E}\,(W_r^2)}{(\mathbb{E}\,(W_r))^2} = \frac{\mathbb{E}\,(Z_r^2)}{(\mathbb{E}\,(Z_r))^2} , \quad r > 0 .$$

Probaremos ahora que

$$\lim_{r\to\infty} \theta(r, 1) = \frac{\pi}{2} \tag{6.8}$$

Para ello, probaremos que, cuando $r \to \infty$, los momentos de Z_r convergen a los momentos de una variable Z con densidad

$$\rho_Z(z) = \frac{e^{-z^2}}{\int_0^\infty e^{-z^2} dz} = \frac{2}{\sqrt{\pi}} e^{-z^2} , \quad z > 0$$

y por tanto

$$\mathbb{E}\, Z = \frac{1}{\sqrt{\pi}} \quad \text{y} \quad \mathbb{E}\, Z^2 = \frac{1}{2} \implies \frac{\mathbb{E}\,(Z^2)}{(\mathbb{E}\,(Z))^2} = \frac{\pi}{2}$$

probando así (6.8). La densidad de Z_r es

$$\rho_{Z_r}(z) = \frac{e^{-z^2-z/\sqrt{r}}}{\int_0^\infty e^{-z^2-z/\sqrt{r}}dz}\ , \quad z > 0\ .$$

Como $e^{-z^2-z/\sqrt{r}} \uparrow e^{-z^2}$ cuando $r \uparrow \infty$ entonces por el Teorema de la convergencia monótona se sigue que

$$\lim_{r\to\infty} \int_0^\infty e^{-z^2-z/\sqrt{r}}dz = \int_0^\infty e^{-z^2}dz$$

por lo cual obtenemos, para cada $k \geq 0$ y $z > 0$ que

$$\lim_{r\to\infty} z^k \rho_{Z_r}(z) = z^k \rho_Z(z)\ . \tag{6.9}$$

Por otro lado,

$$|z^k \rho_{Z_r}(z)| \leq z^k \frac{e^{-z^2}}{\int_0^\infty e^{-z^2-z}dz} =: g(z), \quad k \in \mathbb{N} \cup \{0\}, \tag{6.10}$$

y dado que g es integrable sobre $[0, \infty[$, concluimos de (6.9), (6.10) y del Teorema de la convergencia dominada que para cada $k \in \mathbb{N}$

$$\mathbb{E}\, Z_r^k = \int_0^\infty z^k \rho_{Z_r}(z)dz \xrightarrow{(r\uparrow\infty)} \int_0^\infty z^k \rho_Z(z)dz = \mathbb{E}\, Z^k.$$

Eso concluye la prueba de (6.8). Ya que θ es continua y $\theta(0,1) = 2$, por el Teorema del valor intermedio concluimos de (6.8) que

$$\{\theta(r,1),\ 0 \leq r < \infty\} \supseteq\ ]\pi/2, 2]. \tag{6.11}$$

Observe ahora que $Z \sim G_{1,0}$ y entonces

$$\theta\Big(\frac{1}{1-u}, \frac{-2u}{1-u}\Big) = \frac{\pi}{2}, \qquad \text{para } u = 0. \tag{6.12}$$

Para $u \in [0, 1[$, sea V_u con densidad

$$\rho_{V_u}(v) \propto \exp\Big(-\frac{(v-u)^2}{1-u}\Big), \qquad v \geq 0.$$

En otras palabras, $V_u \sim G_{1/(1-u),-2u/(1-u)}$ y

$$\theta\Big(\frac{1}{1-u}, \frac{-2u}{1-u}\Big) = \frac{\mathbb{E}\, V_u^2}{(\mathbb{E}\, V_u)^2}, \qquad \text{para } u \in [0, 1[\ .$$

Note que V_0 tiene la misma distribución que Z. A continuación probaremos que los momentos 1 y 2 de V_u convergen a 1 cuando $u \uparrow 1$. En efecto, no es difícil verificar que $\mathbb{E}\,|V_u|^k < \infty$. Ahora bien, sean $k \geq 1$ y $u_n \uparrow 1$. Afirmamos que

$$\mathbb{E}\,|V_{u_n} - u_n|^k \to 0. \tag{6.13}$$

En efecto, ya que la densidad de $W_n = \frac{V_{u_n} - u_n}{\sqrt{1-u_n}}$ es

$$\rho_{W_n}(w) = \frac{1}{c_n} e^{-w^2} 1_{\{w \geq -\frac{u_n}{\sqrt{1-u_n}}\}},$$

donde $c_n = \int_{\frac{-u_n}{\sqrt{1-u_n}}}^{\infty} e^{-w^2} dw$, se tiene que

$$\begin{aligned}
\mathbb{E}\,|V_{u_n} - u_n|^k &= \mathbb{E}\left[(1-u_n)^{k/2}\left|\frac{V_{u_n} - u_n}{\sqrt{1-u_n}}\right|^k\right] \\
&= (1-u_n)^{k/2}\frac{1}{c_n}\int_{\frac{-u_n}{\sqrt{1-u_n}}}^{\infty} |w|^k e^{-w^2} dw \\
&\leq (1-u_n)^{k/2}\frac{\int_{-\infty}^{\infty} |w|^k e^{-w^2} dw}{\int_0^{\infty} e^{-w^2} dw} \qquad (6.14)
\end{aligned}$$

Haciendo $n \to \infty$ en (6.14), se obtiene lo afirmado en (6.13). Luego si $k = 1$ en (6.13) se sigue que

$$\mathbb{E}\,V_{u_n} \to 1 \tag{6.15}$$

cuando $n \to \infty$. Para $k = 2$ en (6.13), $\mathbb{E}\,|V_{u_n} - u_n|^2 \to 0$, es decir,

$$\mathbb{E}\,V_{u_n}^2 - 2u_n \mathbb{E}\,V_{u_n} + u_n^2 \;\to\; 0\,.$$

Luego, gracias a (6.15) se sigue que

$$\mathbb{E}\,V_{u_n}^2 \to 1. \tag{6.16}$$

De (6.15) y (6.16) concluimos que

$$\lim_{u \to 1^-} \theta\Big(\frac{1}{1-u}, \frac{-2u}{1-u}\Big) = 1. \tag{6.17}$$

Por lo tanto de (6.12), (6.17) y el Teorema del valor intermedio tenemos que

$$\big\{\theta\big(1/(1-u), -2u/(1-u)\big),\ 0 \leq u < 1\big\} \supseteq\ ]1, \pi/2]. \tag{6.18}$$

Finalmente de (6.11) y (6.18) se obtiene la afirmación hecha en (6.7). Esto completa la prueba. □

6.1. Prueba del Teorema 2.3, parte (a)

Para la prueba de la parte (a) del Teorema 2.3, elija r, s tal que $G_{r,s}$ tenga primer momento 1 y segundo momento b (esto es posible por la Proposición 6.1). En esta prueba C, C_n denotarán cualquier constante positiva que puede depender solamente de b, r o s y no de otros parámetros. (A priori, no sabemos aún que r, s son únicamente determinados por b, así que los tratamos como parámetros independientes). Sean $c(b), C(b)$ y $C_3(b)$ como en la Proposición 5.3. Usaremos la notación

$$\epsilon = n^{-10}\,, \quad n \geq 1\,.$$

Será evidente de la prueba que el número 10 en el exponente no es esencial, cualquier número suficientemente grande serviría.

Usando el Lema 5.4, probaremos que para todo n suficientemente grande,

$$C_1^{-1} n\epsilon^2 \leq \mathbb{P}(Y \in \mathcal{K}^\epsilon) \leq C_1 n\epsilon^2. \tag{6.19}$$

En efecto,

$$\begin{aligned} \mathbb{P}(Y \in \mathcal{K}^\epsilon) &= \mathbb{P}(Y \in \hat{\mathcal{K}}^\epsilon, m(Y) > 0) \\ &= \mathbb{P}(Y \in \hat{\mathcal{K}}^\epsilon) \end{aligned}$$

pues $\mathbb{P}(m(Y) > 0) = 1$. La última expresión es igual a

$$\begin{aligned} & \mathbb{P}\big(\epsilon < \mu(Y) - 1 < 2\epsilon,\ \ \epsilon < \mu_2(Y) - b < b\epsilon\big) \\ = \ & \mathbb{P}\Big(\epsilon + 1 < \frac{1}{n}\sum_{i=1}^n Y_i < 2\epsilon + 1,\ \ b + \epsilon < \frac{1}{n}\sum_{i=1}^n Y_i^2 < b\epsilon + b\Big) \\ = \ & \mathbb{P}\Big(a_n < \frac{1}{n}\sum_{i=1}^n Y_i < b_n,\ \ a'_n < \frac{1}{n}\sum_{i=1}^n Y_i^2 < b'_n\Big), \end{aligned}$$

donde

$$a_n = \epsilon + 1 = \frac{1}{n^{10}} + 1\,, \quad b_n = 1 + 2\epsilon = 1 + \frac{2}{n^{10}}$$

y

$$a'_n = b + \epsilon = b + \frac{1}{n^{10}}\,, \quad b'_n = b\epsilon + b = \frac{b}{n^{10}} + b\,.$$

Es fácil ver que

$$\begin{aligned} a_n \leq b_n \quad &, \quad a'_n \leq b'_n, \\ \lim_{n\to\infty} \sqrt{n}(a_n - 1) &= \lim_{n\to\infty} \sqrt{n}(b_n - 1) = 0, \\ \lim_{n\to\infty} \sqrt{n}(a'_n - \beta) &= \lim_{n\to\infty} \sqrt{n}(b'_n - \beta) = 0, \end{aligned}$$

donde $\beta = \mathbb{E}\, Y_1^2 = b$. Ahora aplicando el Lema 5.4 se tiene

$$\lim_{n\to\infty} \frac{\mathbb{P}(Y \in \mathcal{K}^\epsilon)}{n\left(\frac{1}{n^{10}}\right)\left(\frac{b-1}{n^{10}}\right)} = \rho(0,0),$$

donde ρ es una densidad gaussiana no degenerada sobre $\mathbb{R}^2$, y escrito de otra forma

$$\lim_{n\to\infty} \frac{\mathbb{P}(Y \in \mathcal{K}^\epsilon)}{(b-1)n\epsilon^2} = \rho(0,0),$$

Debido a que $\rho(0,0) > 0$, $\exists\, n_0 \in \mathbb{N}$ tal que $\forall\, n \geq n_0$ se tiene

$$\frac{1}{C} \;\leq\; \frac{\mathbb{P}(Y \in \mathcal{K}^\epsilon)}{(b-1)n\epsilon^2} \;\leq\; C$$

para alguna constante $C > 0$, de lo cual

$$\frac{1}{C}(b-1)n\epsilon^2 \;\leq\; \mathbb{P}(Y \in \mathcal{K}^\epsilon) \;\leq\; C(b-1)n\epsilon^2\,.$$

Así

$$\frac{1}{C_1}n\epsilon^2 \;\leq\; \mathbb{P}(Y \in \mathcal{K}^\epsilon) \;\leq\; C_1 n\epsilon^2\,,$$

donde $C_1 = \max\{C(b-1), \frac{C}{b-1}\}$. Note que esto prueba, en particular, que $\mathcal{K}^\epsilon$ es no vacío. Ahora, si $Y_1 \leq C(b)\epsilon$ y $Y \in \mathcal{K}^\epsilon$ se tiene que $Y_1 > 0$ y

$$\begin{aligned} -C(b)n^{-1}\epsilon &\leq \frac{-Y_1}{n} < 0 \\ -C(b)^2 n^{-1}\epsilon^2 &\leq \frac{-Y_1^2}{n} < 0, \end{aligned}$$

de lo cual se sigue

$$\frac{1}{n}\sum_{i=2}^{n} Y_i = \mu(Y) - \frac{Y_1}{n} \in \,]1 + \epsilon - n^{-1}C(b)\epsilon, 1 + 2\epsilon[= I_1, \tag{6.20}$$

$$\frac{1}{n}\sum_{i=2}^{n} Y_i^2 = \mu_2(Y) - \frac{Y_1^2}{n} \in \,]b + \epsilon - n^{-1}C(b)^2\epsilon^2, b + \epsilon b[= I_2. \tag{6.21}$$

Sea E el evento tal que los dos eventos anteriores suceden, es decir:

$$E := \left\{ \mu(Y) - \frac{Y_1}{n} \in \ I_1, \ \ \mu_2(Y) - \frac{Y_1^2}{n} \in \ I_2 \right\}.$$

Note de lo anterior que el evento $\{Y_1 \leq C(b)\epsilon, \ Y \in \mathcal{K}^\epsilon\}$ implica el evento E. Luego utilizando de nuevo el Lema 5.4, probaremos que

$$\mathbb{P}(E) \leq C_2 n\epsilon^2, \tag{6.22}$$

donde C_2 es alguna constante positiva que depende sólo de b. En efecto,

$$\begin{aligned} E &= \left\{ \mu(Y) - \frac{Y_1}{n} \in \ I_1, \ \ \mu_2(Y) - \frac{Y_1^2}{n} \in \ I_2 \right\} \\ &= \left\{ a_n \leq \frac{\sum_{i=2}^n Y_i}{n-1} \leq b_n \ , \ a_n' \leq \frac{\sum_{i=2}^n Y_i^2}{n-1} \leq b_n' \right\}, \end{aligned}$$

donde $a_n = \frac{n+n\epsilon-C(b)\epsilon}{n-1}$, $b_n = \frac{n+2n\epsilon}{n-1}$, $a_n' = \frac{nb+n\epsilon-C(b)^2\epsilon^2}{n-1}$, $b_n' = \frac{nb(1+\epsilon)}{n-1}$. Es fácil ver que

$$\begin{aligned} a_n \leq b_n \quad &, \quad a_n' \leq b_n', \\ \lim_{n\to\infty} \sqrt{n}(a_n - 1) &= \lim_{n\to\infty} \sqrt{n}(b_n - 1) = 0, \\ \lim_{n\to\infty} \sqrt{n}(a_n' - \beta) &= \lim_{n\to\infty} \sqrt{n}(b_n' - \beta) = 0, \end{aligned}$$

donde $\beta = \mathbb{E}\, Y_1^2 = b$. Ahora aplicando el Lema 5.4 se tiene

$$\lim_{n\to\infty} \mathbb{P}(E) = \lim_{n\to\infty} \frac{\mathbb{P}\left(a_n \leq \frac{1}{n-1}\sum_{i=2}^n Y_i \leq b_n \ , \ a_n' \leq \frac{1}{n-1}\sum_{i=2}^n Y_i^2 \leq b_n'\right)}{n\left(\frac{n\epsilon+C(b)\epsilon}{n-1}\right)\left(\frac{n\epsilon(b-1)+C(b)^2\epsilon^2}{n-1}\right)} = \hat{\rho}(0,0)$$

donde $\hat{\rho}$ es como en el Lema 5.4. Luego, podemos hallar un $n_0 \in \mathbb{N}$ tal que $\forall\, n \geq n_0$

$$\frac{\mathbb{P}(E)}{n\tau_1\tau_2} \leq 2\hat{\rho}(0,0),$$

donde $\tau_1 = \epsilon\left(\frac{n}{n-1} + \frac{C(b)}{n-1}\right)$, $\tau_2 = \epsilon\left(\frac{n}{n-1}(b-1) + \frac{C(b)^2\epsilon}{n-1}\right)$. Luego $\mathbb{P}(E) \leq 2\hat{\rho}(0,0)n\tau_1\tau_2$, pero $\tau_1 \leq \epsilon(2 + C(b))$ y $\tau_2 \leq \epsilon(2(b-1) + C(b)^2)$ para $n \in \mathbb{N}$ adecuado, entonces

$$\begin{aligned} \mathbb{P}(E) &\leq 2\hat{\rho}(0,0)n\epsilon(2 + C(b))\epsilon(2(b-1) + C(b)^2) \\ &= 2\hat{\rho}(0,0)(2 + C(b))(2(b-1) + C(b)^2)n\epsilon^2 \\ &= C_2 n\epsilon^2. \end{aligned}$$

Dado que el evento E es independiente del evento $(Y_1 \leq C(b)\epsilon)$ se sigue que

$$\begin{aligned}\mathbb{P}(Y_1 \leq C(b)\epsilon, Y \in \mathcal{K}^\epsilon) &\leq \mathbb{P}(\{Y_1 \leq C(b)\epsilon\} \cap E) \\ &= \mathbb{P}(Y_1 \leq C(b)\epsilon)\mathbb{P}(E) \leq Cn\epsilon^3, \end{aligned} \tag{6.23}$$

donde en la última desigualdad se ha usado (6.22) y que $\mathbb{P}(Y_1 \leq C(b)\epsilon)$ tiene cota superior $C\epsilon$. Combinando con (6.19) y observando que $n^2\epsilon^3 \ll n\epsilon^2$, afirmamos que para n suficientemente grande,

$$\mathbb{P}(m(Y) > C(b)\epsilon, Y \in \mathcal{K}^\epsilon) \geq C^{-1}n\epsilon^2. \tag{6.24}$$

En efecto, como

$$\begin{aligned}(Y \in \mathcal{K}^\epsilon) &= (m(Y) > C(b)\epsilon, Y \in \mathcal{K}^\epsilon) \cup (m(Y) \leq C(b)\epsilon, Y \in \mathcal{K}^\epsilon) \\ &= (m(Y) > C(b)\epsilon, Y \in \mathcal{K}^\epsilon) \cup \bigcup_{i=1}^{n} (Y_i \leq C(b)\epsilon, Y \in \mathcal{K}^\epsilon),\end{aligned}$$

se sigue que

$$\begin{aligned}\mathbb{P}\left(Y \in \mathcal{K}^\epsilon\right) &\leq \mathbb{P}\left(m(Y) > C(b)\epsilon, Y \in \mathcal{K}^\epsilon\right) + \sum_{i=1}^{n} \mathbb{P}\left(Y_i \leq C(b)\epsilon, Y \in \mathcal{K}^\epsilon\right) \\ &= \mathbb{P}\left(m(Y) > C(b)\epsilon, Y \in \mathcal{K}^\epsilon\right) + n\mathbb{P}\left(Y_1 \leq C(b)\epsilon, Y \in \mathcal{K}^\epsilon\right),\end{aligned}$$

luego de (6.19) y (6.23)

$$\begin{aligned}\mathbb{P}\left(m(Y) > C(b)\epsilon, Y \in \mathcal{K}^\epsilon\right) &\geq \mathbb{P}\left(Y \in \mathcal{K}^\epsilon\right) - n\mathbb{P}\left(Y_1 \leq C(b)\epsilon, Y \in \mathcal{K}^\epsilon\right) \\ &\geq C_1^{-1}n\epsilon^2 - nCn\epsilon^3 = C_1^{-1}n\epsilon^2 - Cn^2\epsilon^3 \\ &= n\epsilon^2\left(\frac{1}{C_1} - Cn\epsilon\right) \\ &\geq n\epsilon^2\left(\frac{1}{C_1} - \frac{1}{2C_1}\right) \quad \text{(para } n \text{ suficientemente grande)} \\ &= n\epsilon^2\frac{1}{2C_1} = (2C_1)^{-1}n\epsilon^2.\end{aligned}$$

Sea $h : \mathbb{R} \to \mathbb{R}$ una función satisfaciendo $|h(x)| \leq 1$ y $|h(x) - h(y)| \leq L|x - y|$ para todo $x, y \in \mathbb{R}$, donde L es alguna constante positiva. Defina $g : \mathbb{R}^n \to \mathbb{R}$ como

$$g(x) := \frac{1}{n}\sum_{i=1}^{n} h(x_i).$$

Note que $|g(x) - g(y)| \leq L \max_i |x_i - y_i|$. Sea $a := \mathbb{E}\, h(Y_1)$. Luego por la desigualdad de Hoeffding (9.7) para sumas de variables aleatorias acotadas independientes, tenemos que para cualquier $t > 0$

$$\mathbb{P}(|g(Y) - a| > t) \leq 2e^{-nt^2/2}, \tag{6.25}$$

pues como $\mathbb{E}\, g(Y) = \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^n h(Y_i)\right) = \frac{1}{n}\sum_{i=1}^n \mathbb{E}\, h(Y_i) = a$, entonces se tiene

$$\begin{aligned}
\mathbb{P}\left(|g(Y) - a| > t\right) &= \mathbb{P}\left(|ng(Y) - na| > nt\right) \\
&= \mathbb{P}\Big(\Big|\sum_{i=1}^n h(Y_i) - \mathbb{E}\sum_{i=1}^n h(Y_i)\Big| > nt\Big) \\
&\leq 2e^{-\frac{2n^2t^2}{4n}} = 2e^{-\frac{nt^2}{2}}.
\end{aligned}$$

Por lo tanto se sigue de la Proposición 5.3, (6.24) y de (6.25) que para todo $t > C_3(b)L\epsilon n$,

$$\mathbb{P}(|g(X) - a| > t) \leq Cn^{-1}\epsilon^{-2}e^{-n(t - C_3(b)L\epsilon n)^2/2}, \tag{6.26}$$

pues

$$\begin{aligned}
\mathbb{P}(|g(X) - a| > t) &\leq \frac{e^{2B\epsilon n}\mathbb{P}(|g(Y) - a| > t - C_3(b)L\epsilon n, Y \in \mathcal{K}^\epsilon)}{\mathbb{P}(m(Y) > C(b)\epsilon, Y \in \mathcal{K}^\epsilon)} \\
&\leq \frac{e^{2B\epsilon n}2e^{-n(t - C_3(b)L\epsilon n)^2/2}}{C^{-1}n\epsilon^2} \\
&\leq 2e^{2B}Cn^{-1}\epsilon^{-2}e^{-n(t - C_3(b)L\epsilon n)^2/2} \\
&= C_3 n^{-1}\epsilon^{-2}e^{-n(t - C_3(b)L\epsilon n)^2/2}, \qquad C_3 = 2e^{2B}C
\end{aligned}$$

A continuación probaremos que la cota superior en (6.26) decae rápidamente en el régimen $t > A(n^{-1}\log n)^{1/2}$, donde A es una constante adecuada que depende de b, de lo cual es fácil deducir que

$$\mathbb{E}\,|g(X) - a| \leq C\sqrt{\frac{\log n}{n}} \tag{6.27}$$

En efecto, reemplazando $\epsilon = n^{-10}$ en (6.26) se sigue que

$$\mathbb{P}(|g(X) - a| > t) \leq Cn^{-1}\epsilon^{-2}e^{-n(t - C_3(b)L\epsilon n)^2/2} = Cn^{19}e^{-n(t - \frac{C_3(b)L}{n^9})^2/2}.$$

Luego, es fácil ver que $\exists\, l_0 \in \mathbb{N}$ tal que $\forall\, n \geq l_0$, $\frac{L}{n^9} \leq \sqrt{\frac{\log n}{n}}$. Luego si $A = \text{máx}\{2C_3(b), \sqrt{20 \times 8}\}$ tendremos que $\forall\, n \geq l_0$ y $\forall\, t > A\sqrt{\frac{\log n}{n}}$

$$\frac{C_3(b)L}{n^9} \leq \frac{A}{2}\sqrt{\frac{\log n}{n}} < t/2,$$

y por lo tanto $\mathbb{P}\big(|g(X) - a| > t\big) \leq Cn^{19}e^{-n(\frac{t}{2})^2/2}$. Ahora, para $n \geq l_0$ tenemos:

$$\begin{aligned}
\mathbb{E}\,|g(X) - a| &= \int_0^\infty \mathbb{P}\big(|g(X) - a| > t\big)dt \\
&\leq \int_0^{A\sqrt{\frac{\log n}{n}}} \mathbb{P}\big(|g(X) - a| > t\big)dt + \int_{A\sqrt{\frac{\log n}{n}}}^\infty \mathbb{P}\big(|g(X) - a| > t\big)dt \\
&\leq A\sqrt{\frac{\log n}{n}} + \int_{A\sqrt{\frac{\log n}{n}}}^\infty Cn^{19}e^{-\frac{nt^2}{8}}dt. \qquad (6.28)
\end{aligned}$$

Haciendo el cambio de variable $t = A\sqrt{\frac{\log n}{n}}s$, se tiene que :

$$\begin{aligned}
\int_{A\sqrt{\frac{\log n}{n}}}^\infty Cn^{19}e^{-\frac{nt^2}{8}}dt &= A\sqrt{\frac{\log n}{n}} \int_1^\infty Cn^{19}e^{-\frac{A^2}{8}(\log n)s^2}ds \\
&= AC\sqrt{\frac{\log n}{n}} \int_1^\infty e^{19\log n - \frac{A^2}{8}(\log n)s^2}ds \\
&\leq AC\sqrt{\frac{\log n}{n}} \int_1^\infty e^{19(\log n)s^2 - \frac{A^2}{8}(\log n)s^2}ds \\
&\leq AC\sqrt{\frac{\log n}{n}} \int_1^\infty e^{-(\log n)s^2}ds \\
&\leq AC\sqrt{\frac{\log n}{n}} \underbrace{\int_1^\infty e^{-(\log 2)s^2}ds}_{M\in\mathbb{R}} \\
&= ACM\sqrt{\frac{\log n}{n}}. \qquad (6.29)
\end{aligned}$$

Luego de (6.29) en (6.28) se sigue la desigualdad (6.27). Por otro lado, por la desigualdad de Jensen tenemos

$$\begin{aligned}
\mathbb{E}\,|g(X) - a| &\geq |\mathbb{E}\,(g(X) - a)| \\
&= |\mathbb{E}\,g(X) - a|. \qquad (6.30)
\end{aligned}$$

Además por el Teorema 3.11 tenemos que $h(X_i) \sim h(X_1)$ para todo i y por lo cual

$$\mathbb{E}\, g(X) = \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n} h(X_i)\right) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\, h(X_i) = \mathbb{E}\, h(X_1). \tag{6.31}$$

Luego de (6.27), (6.30) y (6.31) se sigue que

$$|\mathbb{E}\, h(X_1) - \mathbb{E}\, h(Y_1)| \leq C\sqrt{\frac{\log n}{n}}\cdot \tag{6.32}$$

Finalmente en la desigualdad anterior haciendo $n \to \infty$ tenemos que

$$\mathbb{E}\, h(X_1) \to \mathbb{E}\, h(Y_1),$$

y por la Proposición 9.9, para $k = 1$, concluimos que $X_1 \xrightarrow{ley} Y_1$.

Para mostrar la convergencia en ley conjunta de $X_1, X_2, ..., X_k$, procedemos de la siguiente manera. En vez de una sola función h, consideramos k funciones $h_1, ..., h_k$, cada una satisfaciendo $|h_i(x)| \leq 1$ y $|h_i(x) - h_i(y)| \leq L|x - y|$. Defina $g_1, ..., g_k$ y $a_1, ..., a_k$ como $g_k(x) := \frac{1}{n}\sum_{i=1}^{n} h_k(x_i)$, y $a_k = \mathbb{E}\, h_k(Y_k)$. Luego $|g_i| \leq 1$, y por lo tanto por un simple argumento telescópico obtenemos

$$\begin{aligned}\mathbb{E}\left|\prod_{i=1}^{k} g_i(X) - \prod_{i=1}^{k} a_i\right| &\leq \sum_{i=1}^{k} \mathbb{E}\, |g_i(X) - a_i| \\ &\leq k \max_{1\leq i\leq k} \mathbb{E}\, |g_i(X) - a_i| \leq Ck\sqrt{\frac{\log n}{n}}\cdot \end{aligned} \tag{6.33}$$

Ahora, colocando $A := \prod_{i=1}^{k} a_i$ y usando la desigualdad de Jensen, afirmamos que

$$\begin{aligned}\mathbb{E}\left|\prod_{i=1}^{k} g_i(X) - A\right| &\geq \left|\frac{1}{n^k}\sum_{1\leq i_1,\ldots,i_k\leq n} [\mathbb{E}\,(h_1(X_{i_1})\cdots h_k(X_{i_k})) - A]\right| \\ &\geq |\mathbb{E}\,(h_1(X_1)\cdots h_k(X_k)) - A| + \mathcal{O}(1/n). \end{aligned} \tag{6.34}$$

En efecto,

$$\begin{aligned}
\mathbb{E}\left|\prod_{i=1}^{k} g_i(X)-A\right| &\geq \left|\mathbb{E}\prod_{i=1}^{k} g_i(X)-\mathbb{E}\,A\right| \\
&= \left|\mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n} h_1(X_i)\times\cdots\times\frac{1}{n}\sum_{i=1}^{n} h_k(X_i)\right)-A\right| \\
&= \left|\frac{1}{n^k}\mathbb{E}\left(\sum_{1\leq i_1,\ldots,i_k\leq n} h_1(X_{i_1})\cdots h_k(X_{i_k})\right)-A\right| \\
&= \left|\frac{1}{n^k}\sum_{1\leq i_1,\ldots,i_k\leq n}\mathbb{E}\left[h_1(X_{i_1})\cdots h_k(X_{i_k})-A\right]\right|. \qquad (6.35)
\end{aligned}$$

Podemos dividir la sumatoria dada en (6.35) en dos sumatorias de la siguiente manera:

$$\sum_{1\leq i_1,\ldots,i_k\leq n}\mathbb{E}\left[h_1(X_{i_1})\cdots h_k(X_{i_k})-A\right]=S+R \qquad (6.36)$$

donde

$$\begin{aligned}
S &:= \sum_{\substack{1\leq i_1,\ldots,i_k\leq n\\ \text{sin repetición}}}\mathbb{E}\left[h_1(X_{i_1})\cdots h_k(X_{i_k})-A\right], \\
R &:= \sum_{\substack{1\leq i_1,\ldots,i_k\leq n\\ \text{con repetición}}}\mathbb{E}\left[h_1(X_{i_1})\cdots h_k(X_{i_k})-A\right].
\end{aligned}$$

A continuación hallaremos una cota inferior conveniente para S. En efecto, por el Teorema 3.11, $(X_1,...,X_k)\sim(X_{i_1},...,X_{i_k})$ y por lo cual

$$h_1(X_1)\times\cdots\times h_k(X_k)\sim h_1(X_{i_1})\times\cdots\times h_k(X_{i_k}),$$

donde $i_m\neq i_n$, y así

$$\begin{aligned}
S &= n(n-1)\cdots(n-k+1)\big(\mathbb{E}\left(h_1(X_1)\cdots h_k(X_k)\right)-A\big) \\
&= c_{n,k}\big(\mathbb{E}\left(h_1(X_1)\cdots h_k(X_k)\right)-A\big), \qquad (6.37)
\end{aligned}$$

donde $c_{n,k}=n(n-1)\cdots(n-k+1)$, pero

$$\begin{aligned}
|c_{n,k}(\mathbb{E}\left(h_1(X_1)\cdots h_k(X_k)\right)-A)| &= |(n^k+c_{n,k}-n^k)(\mathbb{E}\left(h_1(X_1)\cdots h_k(X_k)\right)-A)| \\
&\geq n^k|\mathbb{E}\left(h_1(X_1)\cdots h_k(X_k)\right)-A|- \\
&\quad |(c_{n,k}-n^k)(\mathbb{E}\left(h_1(X_1)\cdots h_k(X_k)\right)-A)|. \qquad (6.38)
\end{aligned}$$

Por lo tanto de (6.38) en (6.37) se tiene que

$$|S| \geq n^k|\mathbb{E}\left(h_1(X_1)\cdots h_k(X_k)\right) - A| - \\ |(c_{n,k} - n^k)(\mathbb{E}\left(h_1(X_1)\cdots h_k(X_k)\right) - A)|. \tag{6.39}$$

Ahora hallaremos una cota superior para R.

$$\begin{aligned} |R| &\leq \sum_{\substack{1\leq i_1,\ldots,i_k\leq n \\ \text{con repetición}}} |\mathbb{E}\left(h_1(X_{i_1})\cdots h_k(X_{i_k})\right) - A| \\ &\leq \sum_{\substack{1\leq i_1,\ldots,i_k\leq n \\ \text{con repetición}}} |\mathbb{E}\left(h_1(X_{i_1})\cdots h_k(X_{i_k})\right)| + |A| \\ &\leq \sum_{\substack{1\leq i_1,\ldots,i_k\leq n \\ \text{con repetición}}} \mathbb{E}\left|h_1(X_{i_1})\cdots h_k(X_{i_k})\right| + |A| \\ &\leq \sum_{\substack{1\leq i_1,\ldots,i_k\leq n \\ \text{con repetición}}} (1+|A|) = (n^k - c_{n,k})(1+|A|). \end{aligned} \tag{6.40}$$

entonces de las desigualdades (6.39), (6.40) en (6.36), se sigue de (6.35) que

$$\begin{aligned} \mathbb{E}\left|\prod_{i=1}^{k} g_i(X) - A\right| &\geq \frac{1}{n^k}\big(|S| - |R|\big) \\ &\geq |\mathbb{E}\left(h_1(X_1)\cdots h_k(X_k)\right) - A| - \\ &\quad \left|\left(\frac{c_{n,k}}{n^k} - 1\right)\left[\mathbb{E}\left(h_1(X_1)\cdots h_k(X_k)\right) - A\right]\right| - \left(1 - \frac{c_{n,k}}{n^k}\right)(1+|A|) \\ &= |\mathbb{E}\left(h_1(X_1)\cdots h_k(X_k)\right) - A| + \\ &\quad \left(\frac{c_{n,k}}{n^k} - 1\right)\big(1+|A| + |\mathbb{E}\left(h_1(X_1)\cdots h_k(X_k)\right) - A|\big) \\ &= |\mathbb{E}\left(h_1(X_1)\cdots h_k(X_k)\right) - A| + \mathcal{O}(1/n). \end{aligned} \tag{6.41}$$

Luego de las desigualdades (6.33), (6.34) y la Proposición 9.9 se concluye que $(X_1, ..., X_k)$ converge en ley a $(Y_1, ..., Y_k)$. Lo cual completa la prueba de la parte (a).

Finalmente, afirmamos que la convergencia en distribución automáticamente prueba la unicidad de r, s. En efecto, sean r, s de modo que $\mathbb{E}\, Y = 1$, $\mathbb{E}\, Y^2 = b$ tal que

$$X_1 \xrightarrow{ley} Y$$

y r', s' de modo que $\mathbb{E}\, Y' = 1$, $\mathbb{E}\,(Y')^2 = b$ tal que

$$X_1 \xrightarrow{ley} Y'$$

entonces $Y \sim Y'$ de lo cual Y e Y' tienen la misma función de densidad y así $r = r'$ y $s = s'$.

6.2. Prueba del Teorema 2.3, partes (c) y (d)

Probaremos en esta sección las afirmaciones hechas en (c) y (d). Considere

$$g(x) \;=\; \max_{1\le i\le n} x_i \;.$$

Para esta función se satisface que

$$|g(x) - g(y)| \le \max_{1\le i\le n} |x_i - y_i| \;, \quad x, y \in \mathbb{R}^n \;.$$

Comenzaremos probando la afirmación en (c). Asumimos entonces que $b < 2$. Eso implica que $r > 0$, dado que $\theta(0, s) = 2$. En esta situación no es difícil concluir que

$$\mathbb{P}\,(g(Y) > t) \le Cne^{-t^2/C}. \tag{6.42}$$

En efecto,

$$\begin{aligned} \mathbb{P}\,(g(Y) > t) \;&=\; \mathbb{P}\left(\max_{1\le i\le n} Y_i > t\right) \\ &\le\; n\mathbb{P}\,(Y_1 > t) \\ &=\; n\mathbb{P}\left(e^{Y_1^2/C} > e^{t^2/C}\right) \\ &\le\; n\mathbb{E}\,(e^{Y_1^2/C})e^{-t^2/C}. \end{aligned}$$

Es posible hallar una constante C adecuada tal que $\mathbb{E}\,(e^{Y_1^2/C}) < \infty$. Veamos,

$$\mathbb{E}\,(e^{Y_1^2/C}) \;=\; \frac{1}{C_{r,s}}\int_0^\infty e^{x^2/C}e^{-rx^2-sx}dx \;=\; \frac{1}{C_{r,s}}\int_0^\infty e^{x^2(1/C-r)-sx}dx \;.$$

Entonces podemos tomar $C > 1/r$ y así obtenemos (6.42).

Por otro lado, de la Proposición 5.3 y la estimativa (6.24) se sigue que, para todo $t > C_3(b)\epsilon n$,

$$\begin{aligned}\mathbb{P}(g(X) > t) &\leq \frac{e^{2B\epsilon n}\mathbb{P}(|g(Y)| > t - C_3(b)\epsilon n, Y \in \mathcal{K}^\epsilon)}{\mathbb{P}(m(Y) > C(b)\epsilon, Y \in \mathcal{K}^\epsilon)} \\ &\leq \frac{e^{2B}\mathbb{P}(g(Y) > t - C_3(b)\epsilon n)}{C^{-1}n\epsilon^2} \,. \end{aligned} \tag{6.43}$$

Usando ahora (6.42), esta última expresión resulta menor o igual que

$$\frac{e^{2B}Cne^{-(t-C_3(b)\epsilon n)^2/C}}{C^{-1}n\epsilon^2} = C\epsilon^{-2}e^{-(t-C_3(b)\epsilon n)^2/C} \tag{6.44}$$

Finalmente, si escogemos una constante A, que depende de b, tal que $A > \sqrt{80C}$, se sigue que

$$\lim_{n\to\infty} \mathbb{P}\big(\max_{1\leq i\leq n} X_i > A\sqrt{\log n}\big) = 0 \,, \tag{6.45}$$

pues

$$\begin{aligned}\mathbb{P}\big(\max_{1\leq i\leq n} X_i > A\sqrt{\log n}\big) &\leq C\epsilon^{-2}e^{-(A\sqrt{\log n}-C_3(b)\epsilon n)^2/C} \\ &\leq C\epsilon^{-2}e^{-\left(A\sqrt{\log n}-\frac{A\sqrt{\log n}}{2}\right)^2/C}\end{aligned}$$

para todo n suficientemente grande y la última expresión es igual a

$$C\epsilon^{-2}e^{-\left(A^2\log n\right)/4C} = Cn^{20}n^{-A^2/4C} \,.$$

Luego haciendo $n \to \infty$ se sigue (6.45) como queríamos.

Para probar ahora (d), supongamos que $b = 2$. En ese caso tenemos que $r = 0$, $s = 1$ por la unicidad de r, s y del hecho de que $\theta(0, 1) = 2$. Luego repetimos el mismo argumento anterior, excepto que en la estimativa obtenida en (6.42) debemos cambiar $e^{-t^2/C}$ por $e^{-t/C}$, lo cual nos da $\log n$ en vez de $\sqrt{\log n}$ en la estimativa enunciada en el teorema.

6.3. Prueba del Teorema 2.3, parte (b)

Probaremos ahora el ítem (b), primero para $k = 1$. Suponga que queremos probar la convergencia del p-ésimo momento. Fije n y para $x > 0$ sea

$$h(x) := \text{mín}\{x^p, (\log n)^{2p}\}.$$

Calculemos una constante de Lipschitz para h. En efecto,

- Si $x > (\log n)^2$ e $y > (\log n)^2$, entonces $h(x) - h(y) = 0$.
- Si $x \leq (\log n)^2$ e $y \leq (\log n)^2$, entonces

$$\begin{aligned} |h(x) - h(y)| &= |x^p - y^p| \\ &= |x - y||x^{p-1} + x^{p-2}y + \cdots + y^{p-1}| \\ &\leq p(\log n)^{2p-2}|x - y|. \end{aligned}$$

- Finalmente si $x \leq (\log n)^2$ pero $y > (\log n)^2$, entonces

$$\begin{aligned} |h(x) - h(y)| &= |h(x) - h((\log n)^2)| \\ &\leq p(\log n)^{2p-2}|x - (\log n)^2| \\ &\leq p(\log n)^{2p-2}|x - y|. \end{aligned}$$

Por lo tanto podemos tomar $L = p(\log n)^{2p-2}$ como la constante de Lipschitz. Luego definimos

$$H(x) := \frac{1}{(\log n)^{2p}} h(x).$$

Es fácil ver que $|H(x)| \leq 1$ y $|H(x) - H(y)| \leq L'|x - y|$, donde $L' = \frac{p}{(\log n)^2}$. Además definimos también $G : \mathbb{R}^n \to \mathbb{R}$ como

$$G(x) := \frac{1}{n} \sum_{i=1}^{n} H(x_i).$$

De nuevo es fácil ver que $|G(x) - G(y)| \leq L' \max_i |x_i - y_i|$. Ahora bien, sea $a = \mathbb{E}\, H(Y_1)$, luego usando la desigualdad de Hoeffding, dada en la Proposición 9.7, podemos proceder de manera similar a (6.25) para obtener

$$\mathbb{P}(|G(Y) - a| > t) \leq 2e^{-nt^2/2}.$$

Luego por la Proposición 5.3 y (6.24) obtenemos para todo $t > C_3(b)L'\epsilon n$

$$\mathbb{P}(|G(X) - a| > t) \leq Cn^{-1}\epsilon^{-2}e^{-n(t-C_3(b)L'\epsilon n)^2/2}. \tag{6.46}$$

De forma similar a (6.27) podemos probar que la estimativa (6.46) decae rápidamente en el régimen $t > A(n^{-1}\log n)^{1/2}$ (donde A es una constante adecuada que depende de b) y que entonces

$$\mathbb{E}\,|G(X) - a| \;\leq\; C\sqrt{\frac{\log n}{n}}.$$

Por la desigualdad de Jensen y ya que las coordenadas de X son idénticamente distribuidas (Teorema 3.11) se sigue que

$$|\mathbb{E}\,H(X_1) - \mathbb{E}\,H(Y_1)| \leq \mathbb{E}\,|G(X) - a|,$$

de lo cual,

$$|\mathbb{E}\,h(X_1) - \mathbb{E}\,h(Y_1)| \leq C(\log n)^{2p}\sqrt{\frac{\log n}{n}}. \tag{6.47}$$

Usando el Teorema de la convergencia monótona es fácil ver que

$$\mathbb{E}\,h(Y_1) \to \mathbb{E}\,Y_1^p. \tag{6.48}$$

Luego de (6.47) y (6.48) concluimos que

$$\mathbb{E}\ \text{mín}\{X_1^p, (\log n)^{2p}\} = \mathbb{E}\,h(X_1) \to \mathbb{E}\,Y_1^p \tag{6.49}$$

Ahora bien, usando la parte (c) y del hecho que $1 \leq X_1 \leq n$, probaremos que cuando $1 < b \leq 2$,

$$|\mathbb{E}\ \text{mín}\{X_1^p, (\log n)^{2p}\} - \mathbb{E}\,(X_1^p)| \to 0. \tag{6.50}$$

Veamos, sea $Z := \text{mín}\{X_1^p, (\log n)^{2p}\} - X_1^p$, luego

$$\begin{aligned} |\mathbb{E}\,Z| &= \left|\int_{\{X_1^p \leq (\log n)^{2p}\}} Z d\mathbb{P} + \int_{\{X_1^p > (\log n)^{2p}\}} Z d\mathbb{P}\right| \\ &\leq \int_{\{X_1^p > (\log n)^{2p}\}} |(\log n)^{2p} - X_1^p| d\mathbb{P} \\ &\leq \int_{\{X_1^p > (\log n)^{2p}\}} 2n^p d\mathbb{P} \\ &= 2n^p\mathbb{P}\left(X_1 > (\log n)^2\right) \end{aligned} \tag{6.51}$$

donde la última desigualdad se cumple para todo n suficientemente grande.

En lo que sigue acotaremos $\mathbb{P}(X_1 > (\log n)^2)$. Para eso utilizaremos la desigualdad que se sigue de (6.43) y (6.44) de la parte (c). Veamos,

$$\begin{aligned}\mathbb{P}\left(X_1 > (\log n)^2\right) &\leq \mathbb{P}\left(\max_{1\leq i\leq n} X_i > (\log n)^2\right)\\ &\leq C\epsilon^{-2}e^{-[(\log n)^2 - C_3(b)\epsilon n]^2/C}\\ &\leq C\epsilon^{-2}e^{-[(\log n)^2 - \frac{(\log n)^2}{2}]^2/C}\\ &= C\epsilon^{-2}e^{-\frac{(\log n)^4}{4C}} \qquad (6.52)\end{aligned}$$

donde la última desigualdad se cumple para todo n suficientemente grande.

Usando (6.52) en (6.51), tenemos $|\mathbb{E}\, Z| \leq 2n^p C\epsilon^{-2}e^{-\frac{(\log n)^4}{4C}}$. Luego haciendo $n \to \infty$ se sigue lo afirmado en (6.50). Finalmente de las ecuaciones (6.49) y (6.50) se completa la prueba para $k = 1$, es decir que $\mathbb{E}\, X_1^p \to \mathbb{E}\, Y_1^p$.

Suponemos ahora que $k > 1$. Definimos $b_i := \mathbb{E}\, H_i(Y_1)$, $H_i : \mathbb{R} \to \mathbb{R}$ como

$$H_i(x) := \frac{1}{(\log n)^{2p_i}} \min\{x^{p_i}, (\log n)^{2p_i}\}, \qquad i = 1, ..., k.$$

y $G_i : \mathbb{R}^n \to \mathbb{R}$ como

$$G_i(x) := \frac{1}{n}\sum_{j=1}^n H_i(x_j) = \frac{1}{n}\sum_{j=1}^n \frac{\min\{x_j^{p_i}, (\log n)^{2p_i}\}}{(\log n)^{2p_i}}, \qquad i = 1, ..., k.$$

Note que $|G_i(x)| \leq 1$. Luego procedemos como en (6.33) de la parte (a) para obtener

$$\mathbb{E}\left|\prod_{i=1}^k G_i(X) - \prod_{i=1}^k b_i\right| \leq k \max_{1\leq i\leq n} \mathbb{E}\,|G_i(X) - b_i| \leq Ck\sqrt{\frac{\log n}{n}}.$$

y

$$\begin{aligned}\mathbb{E}\left|\prod_{i=1}^k G_i(X) - B\right| &\geq \left|\frac{1}{n^k}\sum_{1\leq i_1,\dots,i_k\leq n}[\mathbb{E}\,(H_1(X_{i_1})\cdots H_k(X_{i_k})) - B]\right|\\ &\geq |\mathbb{E}\,(H_1(X_1)\cdots H_k(X_k)) - B| + \mathcal{O}(1/n),\end{aligned}$$

donde $B := \prod_{i=1}^k b_i$. Y así,

$$|\mathbb{E}\,(H_1(X_1)\cdots H_k(X_k)) - B| + \mathcal{O}(1/n) \leq Ck\sqrt{\frac{\log n}{n}}. \qquad (6.53)$$

Expresando (6.53) en función de h_i, obtenemos la siguiente desigualdad

$$|\mathbb{E}\,(h_1(X_1)\cdots h_k(X_k)) - A| + \prod_{i=1}^{k}(\log n)^{2p_i}\mathcal{O}(1/n)$$
$$\leq Ck\prod_{i=1}^{k}(\log n)^{2p_i}\sqrt{\frac{\log n}{n}}, \quad (6.54)$$

donde $A = \prod_{i=1}^{k}\mathbb{E}\,(h_i(Y_i))$. Por otro lado, por el Teorema de la convergencia monótona tenemos que

$$A = \prod_{i=1}^{k}\mathbb{E}\,(h_i(Y_i)) \to \prod_{i=1}^{k}\mathbb{E}\,(Y_i^{p_i}), \quad (6.55)$$

y así de (6.54) y (6.55) se sigue que

$$\mathbb{E}\,(h_1(X_1)\cdots h_k(X_k)) \to \prod_{i=1}^{k}\mathbb{E}\,(Y_i^{p_i}). \quad (6.56)$$

Ahora bien, sea $D_k = \bigcap_{i=1}^{k}\{X_i \leq (\log n)^2\}$. Luego,

$$\left|\mathbb{E}\,(h_1(X_1)\cdots h_k(X_k)) - \mathbb{E}\,(X_1^{p_1}\cdots X_k^{p_k})\right| =$$
$$\left|\int_{D_k}\prod_{i=1}^{k}h_i(X_i)d\mathbb{P} + \int_{D_k^c}\prod_{i=1}^{k}h_i(X_i)d\mathbb{P} - \int_{D_k}\prod_{i=1}^{k}X_i^{p_i}d\mathbb{P} - \int_{D_k^c}\prod_{i=1}^{k}X_i^{p_i}d\mathbb{P}\right|.$$

En la ecuación anterior la primera integral con la tercera se anulan y así

$$|\mathbb{E}\,(h_1(X_1)\cdots h_k(X_k)) - \mathbb{E}\,(X_1^{p_1}\cdots X_k^{p_k})|$$
$$= \left|\int_{D_k^c}\prod_{i=1}^{k}h_i(X_i)d\mathbb{P} - \int_{D_k^c}\prod_{i=1}^{k}X_i^{p_i}d\mathbb{P}\right| \quad (6.57)$$

Podemos acotar superiormente el lado derecho de la ecuación anterior del siguiente modo

$$\begin{aligned}\left|\int_{D_k^c}\prod_{i=1}^{k}h_i(X_i)d\mathbb{P} - \int_{D_k^c}\prod_{i=1}^{k}X_i^{p_i}d\mathbb{P}\right| &\leq \int_{D_k^c}\prod_{i=1}^{k}|h_i(X_i)|d\mathbb{P} + \int_{D_k^c}\prod_{i=1}^{k}|X_i^{p_i}|d\mathbb{P}\\ &\leq \int_{D_k^c}\prod_{i=1}^{k}(\log n)^{2p_i}d\mathbb{P} + \int_{D_k^c}\prod_{i=1}^{k}n^{p_i}d\mathbb{P}\end{aligned}$$

Para todo n suficientemente grande la última expresión es acotada superiormente por

$$2n^{p_1+\cdots p_k}\mathbb{P}(D_k^c) \;=\; 2n^{p_1+\cdots p_k}k\mathbb{P}(X_1 > (\log n)^2)\ ,$$

probando así que

$$\Big|\int_{D_k^c}\prod_{i=1}^{k} h_i(X_i)d\mathbb{P} - \int_{D_k^c}\prod_{i=1}^{k} X_i^{p_i}d\mathbb{P}\Big| \;\leq\; 2n^{p_1+\cdots p_k}k\,\mathbb{P}\big(X_1 > (\log n)^2\big) \tag{6.58}$$

Entonces de (6.57), (6.58) y (6.52) se sigue que

$$|\mathbb{E}\left(h_1(X_1)\cdots h_k(X_k)\right) - \mathbb{E}\left(X_1^{p_1}\cdots X_k^{p_k}\right)| \leq 2kn^{p_1+\cdots p_k}C\epsilon^{-2}e^{-\frac{(\log n)^4}{4C}}. \tag{6.59}$$

Por lo tanto de (6.56) y (6.59) se sigue que

$$\mathbb{E}\left(X_1^{p_1}\cdots X_k^{p_k}\right) \to \prod_{i=1}^{k}\mathbb{E}\left(Y_i^{p_i}\right),$$

lo cual completa la prueba.

Capítulo 7

Prueba del Teorema 2.4

En este capítulo trataremos el caso $b > 2$. Como en el capítulo anterior C y C_n denotarán ciertas constantes que depende solamente de b. También establecemos la notación

$$q := \sqrt{b-2}$$

La prueba de la localización es inspirada por teorema de localización para el modelo Hopfield p-spin.

7.1. Prueba del Teorema 2.4, partes (b) y (c)

Demostración. Sean $Y_1, Y_2, ...$ variables aleatorias i.i.d $\sim exp(1)$ y sea $Y = (Y_1, Y_2, ..., Y_n)$. Tome $\epsilon = n^{-10}$, y definamos

$$\begin{aligned} a &:= 1/100, \\ Z_i &:= Y_i 1_{\{Y_i \leq n^a\}}, \\ v &:= \mathbb{E}\left(Z_1^2\right). \end{aligned}$$

Por la desigualdad de Hoeffding (9.7), tenemos

$$\mathbb{P}\left(\left|\sum_{i=1}^{n}(Z_i^2 - v)\right| > t\right) \leq 2e^{-t^2/2n^{1+4a}}, \tag{7.1}$$

pues como $Y_1, ..., Y_n$ son i.i.d entonces $Z_1, ..., Z_n$ y $Z_1^2, ..., Z_n^2$ también lo son, luego

$$\begin{aligned}\sum_{i=1}^{n}(Z_i^2 - v) &= \sum_{i=1}^{n} Z_i^2 - nv = \sum_{i=1}^{n} Z_i^2 - n\mathbb{E}\, Z_1^2 \\ &= \sum_{i=1}^{n} Z_i^2 - \mathbb{E}\left(\sum_{i=1}^{n} Z_i^2\right),\end{aligned}$$

y así,

$$\begin{aligned}\mathbb{P}\left(\left|\sum_{i=1}^{n}(Z_i^2 - v)\right| > t\right) &= \mathbb{P}\left(\left|\sum_{i=1}^{n} Z_i^2 - \mathbb{E}\left(\sum_{i=1}^{n} Z_i^2\right)\right| > t\right) \\ &\leq 2e^{-2t^2/D}, \text{donde } D = \sum_{i=1}^{n}(n^{2a})^2 = n^{4a+1} \\ &\leq 2e^{-t^2/2n^{1+4a}}.\end{aligned}$$

Por lo tanto, si definimos

$$A := \left\{\left|\sum_{i=1}^{n}(Z_i^2 - v)\right| > n^{5/6}\right\},$$

entonces usando (7.1) obtenemos fácilmente que

$$\mathbb{P}(A) \leq 2\exp\left(-\frac{n^{\frac{2}{3}-4a}}{2}\right). \tag{7.2}$$

Ahora, sea B el evento tal que existe un conjunto $I \subseteq \{1, ..., n\}$ de tamaño $k := [n^{(1-a)/2}]$ tal que $Y_i > n^a$ para todo I, es decir:

$$B = \left\{\omega \in \Omega \;\middle|\; \begin{array}{c} \exists\, I(\omega) \subseteq \{1, ..., n\},\ \text{con tamaño } k = [n^{(1-a)/2}] \\ \text{y además } Y_i(\omega) > n^a,\ \forall\, i \in I(\omega) \end{array}\right\}.$$

El conjunto B también puede ser expresado de la siguiente manera: definiendo $\mathbb{I}_k^n := \{I \subset \{1, ..., n\},\ |I| = k\}$, se tiene

$$B = \bigcup_{I \in \mathbb{I}_k^n} \{Y_i > n^a,\ \forall\, i \in I\}.$$

Luego,

$$\begin{aligned}
\mathbb{P}(B) &\leq \sum_{I\in\mathbb{I}_k^n} \mathbb{P}(Y_i > n^a,\, \forall\, i \in I)\\
&= \binom{n}{k}(e^{-n^a})^k\\
&\leq n^k e^{-kn^a}\\
&\leq 2\exp\left(-\frac{n^{(1+a)/2}}{2}\right),
\end{aligned}$$

donde la última desigualdad es porque

$$\begin{aligned}
n^k e^{-kn^a} &= e^{k\log n} e^{-kn^a}\\
&= e^{k(\log n - n^a)}\\
&\leq e^{n^{(1-a)/2}(\log n - n^a)}, \qquad (7.3)
\end{aligned}$$

y dado que $\dfrac{\log n}{n^a} \to 0$, se tiene que $\exists\, l_0 \in \mathbb{N}$ tal que si $n \geq l_0$ entonces

$$\log n < \frac{n^a}{2}\cdot \qquad (7.4)$$

Luego de (7.4) en (7.3) se sigue que $\forall\, n \geq l_0$:

$$\begin{aligned}
n^k e^{-kn^a} &\leq e^{n^{(1-a)/2}(-n^a/2)}\\
&= e^{-\frac{1}{2}n^{(1+a)/2}}\\
&\leq 2e^{-\frac{1}{2}n^{(1+a)/2}}.
\end{aligned}$$

Ahora sea D el evento tal que $\sum_{i\in I} Y_i > qn^{1/2} + n^{(2-a)/4}$ para algún subconjunto $I \subseteq \{1, ..., n\}$ de tamaño menor que k, es decir:

$$D = \left\{\omega \in \Omega \;\middle|\; \begin{array}{c} \sum_{i\in I} Y_i(\omega) > qn^{1/2} + n^{(2-a)/4},\\ \text{para algún } I \subseteq \{1, ..., n\} \text{ con } |I| < k \end{array}\right\}.$$

El conjunto D también puede ser expresado de la siguiente manera:

$$D = \bigcup_{j=1}^{k-1} \bigcup_{I\in\mathbb{I}_j^n} \left\{\sum_{i\in I} Y_i(\omega) > qn^{1/2} + n^{(2-a)/4}\right\}.$$

Note que para cualquier j, $\sum_{i\in I} Y_i$ sigue una distribución $Gamma(j,1)$. Por lo tanto, para cualquier $j \geq 2$, $t > 2$,

$$
\begin{aligned}
\mathbb{P}\left(\sum_{i=i}^{j} Y_i > t\right) &= \int_t^{\infty} \frac{x^{j-1}}{(j-1)!} e^{-x} dx \\
&= \int_0^{\infty} \frac{(u+t)^{j-1}}{(j-1)!} e^{-(u+t)} du \quad \text{(cambio de variable } u = x - t\text{)} \\
&= e^{-t} \int_0^{\infty} \frac{(u+t)^{j-1}}{(j-1)!} e^{-u} du \\
&\leq e^{-t} \int_0^{\infty} \frac{2^{j-2}(u^{j-1} + t^{j-1})}{(j-1)!} e^{-u} du \\
&\leq e^{-t}(2^{j-2} + t^{j-1}) \\
&\leq 2t^{j-1} e^{-t},
\end{aligned}
$$

donde la primera desigualdad anterior es porque $f(x) = x^n$, $x \geq 0$ es convexa y en la segunda desigualdad anterior se utilizó el hecho que $\int_0^{+\infty} x^n e^{-x} dx = n!$.
Note que la desigualdad es también verdadera para $j = 1$. Por lo tanto, si n es suficientemente grande de manera que $k < n/2$ y tomamos $t := qn^{1/2} + n^{(2-a)/4}$, entonces

$$
\begin{aligned}
\mathbb{P}(D) &\leq \sum_{j=1}^{k-1} \binom{n}{j} 2t^{j-1} e^{-t} \\
&\leq Ckn^k t^k e^{-t}.
\end{aligned}
$$

Y dado que $k = [n^{(1-a)/2}]$ y $(1-a)/2 < (2-a)/4$, afirmamos que

$$
\mathbb{P}(D) \leq C \exp\left(-qn^{1/2} - \frac{n^{(2-a)/4}}{C}\right).
$$

En efecto, se tiene $t = qn^{1/2} + n^{(2-a)/4} = n^{1/2}(q + n^{-a/4})$, luego como $q + n^{-a/4}$ converge a q, entonces es posible hallar $l_1, l_2 \in \mathbb{N}$ tal que

- $qn^{1/2} < t < 2qn^{1/2} \quad \forall n \geq l_1$ y
- $\log t \leq \log n \quad \forall n \geq l_2$.

Como $k = \left[n^{(1-a)/2}\right]$, entonces $k \leq n^{(1-a)/2}$, además $\dfrac{1-a}{2} \leq n^{(1-a)/2}$ y dado que $\frac{\log n}{n^{(2-a)/4-(1-a)/2}} \to 0$ entonces $\exists l_3 \in \mathbb{N}$, tal que $\log n \leq an^{(2-a)/4-(1-a)/2}$ $\forall\ n \geq l_3$. Por lo tanto $\forall\, n \geq l_0 = \text{máx}\{l_1, l_2, l_3\}$ obtenemos

$$\begin{aligned}
Ckn^k t^k e^{-t} &= Ce^{\log k + k\log n + k\log t - t} \\
&\leq Ce^{\frac{1-a}{2}\log n + n^{(1-a)/2}\log n + n^{(1-a)/2}\log t - t} \\
&\leq Ce^{n^{(1-a)/2}\log n + n^{(1-a)/2}\log n + n^{(1-a)/2}\log n - qn^{1/2} - n^{(2-a)/4}} \\
&= Ce^{3n^{(1-a)/2}\log n - qn^{1/2} - n^{(2-a)/4}} \\
&\leq Ce^{3n^{(1-a)/2}\times an^{(2-a)/4-(1-a)/2} - qn^{1/2} - n^{(2-a)/4}} \\
&= Ce^{3an^{(2-a)/4} - qn^{1/2} - n^{(2-a)/4}} \\
&\leq Ce^{-\frac{1}{2}n^{(2-a)/4} - qn^{1/2}} \\
&\leq C_4 e^{-\frac{1}{C_4}n^{(2-a)/4} - qn^{1/2}}
\end{aligned}$$

donde $C_4 := \text{máx}\{2, C\}$. Ahora suponga que $A^c \cap B^c \cap D^c \cap \{Y \in \mathcal{K}^\epsilon\}$ sucede. Sea I el conjunto de i tal que $Y_i > n^a$. Dado que B^c ha sucedido, se tiene $|I| < k$ y dado que D^c ha ocurrido tenemos que

$$\sum_{i\in I} Y_i \leq qn^{1/2} + n^{(2-a)/4}. \tag{7.5}$$

De nuevo, dado que $Y \in \mathcal{K}^\epsilon$, tenemos

$$\left|\sum_{i=1}^{n} Y_i^2 - bn\right| < nb\epsilon = bn^{-9}, \tag{7.6}$$

pues si

$$\begin{aligned}
Y \in \mathcal{K}^\epsilon &\Rightarrow b + \epsilon < \mu_2(Y) < b + b\epsilon \\
&\Rightarrow nb + n\epsilon < \sum_{i=1}^{n} Y_i^2 < nb + nb\epsilon \\
&\Rightarrow n\epsilon < \sum_{i=1}^{n} Y_i^2 - nb < nb\epsilon \\
&\Rightarrow \left|\sum_{i=1}^{n} Y_i^2 - bn\right| < nb\epsilon.
\end{aligned}$$

Pero debido a A^c,

$$\left|\sum_{i\notin I} Y_i^2 - vn\right| \leq n^{5/6}, \tag{7.7}$$

pues

$$\begin{aligned}
A^c &= \left\{\left|\sum_{i=1}^{n}(Z_i^2 - v)\right| \leq n^{5/6}\right\} \\
&= \left\{\left|\sum_{i=1}^{n} Z_i^2 - nv\right| \leq n^{5/6}\right\} \\
&= \left\{\left|\sum_{\substack{i=1\\ i\in I}}^{n} Z_i^2 + \sum_{\substack{i=1\\ i\notin I}}^{n} Z_i^2 - nv\right| \leq n^{5/6}\right\} \\
&= \left\{\left|\sum_{\substack{i=1\\ i\notin I}}^{n} Z_i^2 - nv\right| \leq n^{5/6}\right\} = \left\{\left|\sum_{\substack{i=1\\ i\notin I}}^{n} Y_i^2 - vn\right| \leq n^{5/6}\right\}.
\end{aligned}$$

Combinando las desigualdades (7.6) y (7.7) obtenemos

$$\left|\sum_{i\in I} Y_i^2 - (b-v)n\right| \leq bn^{-9} + n^{5/6} \leq Cn^{5/6},$$

pues

$$\begin{aligned}
\left|\sum_{i\in I} Y_i^2 - (b-v)n\right| &= \left|\sum_{i=1}^{n} Y_i^2 - \sum_{\substack{i=1\\ i\notin I}}^{n} Y_i^2 - bn + vn\right| \\
&\leq bn^{-9} + n^{5/6} \\
&= n^{5/6}(bn^{-59/6} + 1) \\
&\leq n^{5/6}(b{,}1 + 1) = Cn^{5/6}.
\end{aligned}$$

Pero

$$v = \int_0^{n^a} x^2 e^{-x} dx = 2 - \int_{n^a}^{\infty} x^2 e^{-x} dx \qquad \left(\text{pues } \int_0^{\infty} x^2 e^{-x} dx = 2\right),$$

por lo cual para n suficientemente grande

$$\begin{aligned}
|v - 2| &\leq \int_{n^a}^{\infty} x^2 e^{-x} dx \\
&\leq \int_{n^a}^{\infty} e^{x/2} e^{-x} dx \\
&= \int_{n^a}^{\infty} e^{-x/2} dx = \frac{1}{2} e^{-n^a/2}.
\end{aligned} \tag{7.8}$$

Por lo tanto, bajo $A^c \cap B^c \cap D^c \cap \{Y \in \mathcal{K}^\epsilon\}$,

$$\left|\sum_{i\in I} Y_i^2 - (b-2)n\right| \leq Cn^{5/6}, \tag{7.9}$$

pues

$$\begin{aligned}\left|\sum_{i\in I} Y_i^2 - (b-2)n\right| &= \left|\sum_{i\in I} Y_i^2 - (b-v)n + n(2-v)\right| \\ &\leq Cn^{5/6} + n\frac{1}{2}e^{-n^a/2} \\ &\leq Cn^{5/6},\end{aligned}$$

donde la última desigualdad se cumple para n suficientemente grande. Sea ahora

$$M^Y := \max_{1\leq i\leq n} Y_i$$

y recuerde que $q := \sqrt{b-2}$. Luego la desigualdad anterior (7.9) combinada con (7.5), muestra que bajo $A^c \cap B^c \cap D^c \cap \{Y \in \mathcal{K}^\epsilon\}$ se tiene

$$q^2 n - Cn^{5/6} \leq M^Y(qn^{1/2} + n^{(2-a)/4}), \tag{7.10}$$

pues como $\left|\sum_{i\in I} Y_i^2 - (b-2)n\right| \leq Cn^{5/6}$ entonces

$$\begin{aligned}q^2 n - Cn^{5/6} &\leq \sum_{i\in I} Y_i^2 \\ &\leq M^Y \sum_{i\in I} Y_i \qquad (\text{ pues } Y_i^2 \leq M^Y Y_i \text{ }) \\ &\leq M^Y(qn^{1/2} + n^{(2-a)/4}),\end{aligned}$$

donde en la última desigualdad se usó (7.5). Luego, dado que $a/4 < 1/6$, afirmamos que

$$M^Y \geq qn^{1/2}(1 - Cn^{-a/4}). \tag{7.11}$$

En efecto, de (7.10) se sigue que

$$\begin{aligned} M^Y &\geq \frac{q^2 n - Cn^{5/6}}{qn^{1/2} + n^{(2-a)/4}} \\ &= \frac{q^2 n}{qn^{1/2}} \left(\frac{1 - (C/q^2) n^{-1/6}}{1 + \frac{n^{-a/4}}{q}} \right) \\ &= \frac{qn^{1/2}(1 - C_5 n^{-1/6})}{1 + C_6 n^{-a/4}}, \end{aligned}$$

donde $C_5 = \frac{C}{q^2}$ y $C_6 = \frac{1}{q}$. Luego probaremos que

$$\frac{1 - C_5 n^{-1/6}}{1 + C_6 n^{-a/4}} > 1 - Cn^{-a/4},$$

para alguna constante $C > 0$, lo cual es lo mismo que probar

$$Cn^{-a/4} \geq 1 - \frac{1 - C_5 n^{-1/6}}{1 + C_6 n^{-a/4}},$$

para alguna constante $C > 0$. En efecto,

$$\begin{aligned} 1 - \frac{1 - C_5 n^{-1/6}}{1 + C_6 n^{-a/4}} &= \frac{C_6 n^{-a/4} + C_5 n^{-1/6}}{1 + C_6 n^{-a/4}} \\ &= \frac{n^{-a/4}(C_6 + C_5 n^{-(1/6 - a/4)})}{1 + C_6 n^{-a/4}} \\ &\leq n^{-a/4} \left(\frac{C_6 + C_5}{1} \right) = Cn^{-a/4}. \end{aligned}$$

Pero bajo D^c tenemos

$$M^Y \leq qn^{1/2} + n^{(2-a)/4}, \tag{7.12}$$

pues D^c implica que

$$\sum_{i \in I} Y_i \leq qn^{1/2} + n^{(2-a)/4} \quad \Rightarrow \quad M^Y \leq \sum_{i \in I} Y_i \leq qn^{1/2} + n^{(2-a)/4}.$$

Luego, bajo $A^c \cap B^c \cap D^c \cap \{Y \in \mathcal{K}^\epsilon\}$, tenemos

$$|M^Y - qn^{1/2}| \leq Cn^{(2-a)/4}, \tag{7.13}$$

pues de las desigualdades (7.11) y (7.12) se sigue que

$$qn^{1/2} - Cqn^{(2-a)/4} \leq M^Y \leq qn^{1/2} + n^{(2-a)/4},$$

de lo cual

$$-Cqn^{(2-a)/4} \leq M^Y - qn^{1/2} \leq n^{(2-a)/4},$$

luego tomando $C_7 = \max\{1, Cq\}$ se tiene

$$\begin{aligned} -C_7 n^{(2-a)/4} \leq -Cqn^{(2-a)/4} &\leq M^Y - qn^{1/2} \\ &\leq C_7 n^{(2-a)/4}, \end{aligned}$$

y así $|M^Y - qn^{1/2}| \leq C_7 n^{(2-a)/4}$. Por lo tanto, de las cotas superiores para $\mathbb{P}(A), \mathbb{P}(B), \mathbb{P}(D)$ obtenidas anteriormente y observando que la cota superior de $\mathbb{P}(D)$ domina a las otras dos , afirmamos que

$$\begin{aligned} \mathbb{P}\left(|M^Y - qn^{1/2}| > Cn^{(2-a)/4}, Y \in \mathcal{K}^\epsilon\right) &\leq \mathbb{P}(A \cup B \cup D) \\ &\leq C \exp\left(-qn^{1/2} - \frac{n^{(2-a)/4}}{C}\right). \end{aligned} \tag{7.14}$$

En efecto, de (7.13) tenemos que

$$\{|M^Y - qn^{1/2}| > Cn^{(2-a)/4}\} \subset A \cup B \cup D \cup \{Y \in \mathcal{K}^\epsilon\}^c,$$

de lo cual

$$\{|M^Y - qn^{1/2}| > Cn^{(2-a)/4}\} \setminus \{Y \in \mathcal{K}^\epsilon\}^c \subset A \cup B \cup D$$

y así

$$\{|M^Y - qn^{1/2}| > Cn^{(2-a)/4}, Y \in \mathcal{K}^\epsilon\} \subset A \cup B \cup D.$$

Luego para n suficientemente grande

$$\begin{aligned}
\mathbb{P}(| \quad M^Y & - qn^{1/2}| > Cn^{(2-a)/4}, Y \in \mathcal{K}^\epsilon) \\
& \leq \mathbb{P}(A \cup B \cup D) \\
& \leq \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(D) \\
& \leq e^{-qn^{1/2} - \frac{n^{(2-a)/4}}{C}} \left(2e^{-\frac{n^{2/3-4a}}{2} + qn^{1/2} + \frac{n^{(2-a)/4}}{C}} + Ce^{\frac{-n^{(1+a)/2}}{C} + qn^{1/2} + \frac{n^{(2-a)/4}}{C}} + C \right) \\
& \leq e^{-qn^{1/2} - \frac{n^{(2-a)/4}}{C}} (C + C) \text{ (pues los exponentes de } e \text{ tienden a } -\infty \text{)} \\
& \leq C \exp\left(-qn^{1/2} - \frac{n^{(2-a)/4}}{C} \right).
\end{aligned}$$

Sea M_2^Y la segunda coordenada más grande entre las Y_i, entonces bajo $A^c \cap B^c \cap D^c \cap \{Y \in \mathcal{K}^\epsilon\}$

$$\begin{aligned}
M_2^Y & \leq \sum_{i \in I} Y_i - M^Y \\
& \leq qn^{1/2} + n^{(2-a)/4} - (qn^{1/2} - Cn^{(2-a)/4}) \\
& = Cn^{(2-a)/4},
\end{aligned}$$

donde en la segunda desigualdad se ha utilizado (7.5) y (7.13). Lo anterior implica

$$\{M_2^Y > Cn^{(2-a)/4}\} \subseteq A \cup B \cup D \cup \{Y \in \mathcal{K}^\epsilon\}^c,$$

y así

$$\{M_2^Y > Cn^{(2-a)/4}, Y \in \mathcal{K}^\epsilon\} \subseteq A \cup B \cup D.$$

Por lo tanto, de nuevo, tenemos

$$\mathbb{P}\left(M_2^Y > Cn^{(2-a)/4}, Y \in \mathcal{K}^\epsilon\right) \leq C \exp\left(-qn^{1/2} - \frac{n^{(2-a)/4}}{C} \right). \tag{7.15}$$

Esto nos das cotas para el numerador en la Proposición 5.3, excepto que tenemos que evaluar la constante de Lipschitz L para M y M_2. Para un vector $x \in \mathbb{R}^n$, denoten $g_1(x)$ y $g_2(x)$ las componentes primera y segunda más grandes de x respectivamente. Dado que

$$|g_1(x) - g_1(y)| = |\max_i x_i - \max_i y_i| \leq \max_i |x_i - y_i|, \tag{7.16}$$

se sigue que podemos tomar como constante de Lipschitz para g_1 a $L = 1$. Usando la misma lógica que prueba la desigualdad en (7.16) tenemos que

$$|\max_{i<j}(x_i + x_j) - \max_{i<j}(y_i + y_j)| \;\leq\; \max_{i<j}|(x_i + x_j) - (y_i + y_j)|$$

La última expresión coincide con

$$\begin{aligned}
& \max_{i<j}|(x_i - y_i) + (x_j - y_j)| \\
\leq \;& \max_{i<j}\Big(|x_i - y_i| + |x_j - y_j|\Big) \\
\leq \;& 2\max_{i}|x_i - y_i|.
\end{aligned}$$

Sin embargo, $\max_{i<j}(x_i + x_j) = g_1(x) + g_2(x)$, entonces

$$\begin{aligned}
|g_2(x) - g_2(y)| &= |(\max_{i<j}(x_i + x_j) - g_1(x)) - (\max_{i<j}(y_i + y_j) - g_1(y))| \\
&= |\max_{i<j}(x_i + x_j) - \max_{i<j}(y_i + y_j) + g_1(y) - g_1(x)| \\
&\leq |\max_{i<j}(x_i + x_j) - \max_{i<j}(y_i + y_j)| + |g_1(y) - g_1(x)| \\
&\leq 2\max_{i}|x_i - y_i| + \max_{i}|x_i - y_i| \\
&= 3\max_{i}|x_i - y_i|.
\end{aligned}$$

Por lo tanto, podemos tomar $L = 3$ para g_2. Ahora bien, por la Proposición 5.3 obtenemos que

$$\mathbb{P}(|M - qn^{1/2}| > Cn^{\frac{2-a}{4}}) \leq \frac{e^{2B\epsilon n}\mathbb{P}(|M^Y - qn^{1/2}| > Cn^{\frac{2-a}{4}} - C_3(b)\epsilon n,\, Y \in \mathcal{K}^\epsilon)}{\mathbb{P}(m(Y) > C(b)\epsilon, Y \in \mathcal{K}^\epsilon)} \tag{7.17}$$

y

$$\mathbb{P}(M_2 > Cn^{\frac{2-a}{4}}) \leq \frac{e^{2B\epsilon n}\mathbb{P}(M_2^Y > Cn^{\frac{2-a}{4}} - C_3(b)3\epsilon n,\, Y \in \mathcal{K}^\epsilon)}{\mathbb{P}(m(Y) > C(b)\epsilon, Y \in \mathcal{K}^\epsilon)}. \tag{7.18}$$

Ahora bien, con el fin de poder utilizar (7.14), (7.15) y acotar superiormente (7.17) y (7.18), consideramos n lo suficientemente grande tal que $3C_3(b) < \frac{C}{2}n^{(2-a)/4}$ y así

$$\mathbb{P}(|M - qn^{1/2}| > Cn^{\frac{2-a}{4}}) \leq \frac{e^{2B\epsilon n}\mathbb{P}(|M^Y - qn^{1/2}| > \frac{C}{2}n^{(2-a)/4},\, Y \in \mathcal{K}^\epsilon)}{\mathbb{P}(m(Y) > C(b)\epsilon, Y \in \mathcal{K}^\epsilon)} \tag{7.19}$$

y

$$\mathbb{P}(M_2 > Cn^{\frac{2-a}{4}}) \leq \frac{e^{2B\epsilon n}\mathbb{P}(M_2^Y > \frac{C}{2}n^{(2-a)/4},\, Y \in \mathcal{K}^\epsilon)}{\mathbb{P}(m(Y) > C(b)\epsilon, Y \in \mathcal{K}^\epsilon)}. \tag{7.20}$$

Ahora aplicamos (7.14), (7.15) en las dos desigualdades anteriores lo que nos resulta

$$\mathbb{P}(|M - qn^{1/2}| > Cn^{(2-a)/4}) \leq \frac{C\exp(-qn^{1/2} - C^{-1}n^{(2-a)/4})}{\mathbb{P}(m(Y) > C(b)\epsilon, Y \in \mathcal{K}^\epsilon)} \tag{7.21}$$

y

$$\mathbb{P}(M_2 > Cn^{(2-a)/4}) \leq \frac{C\exp(-qn^{1/2} - C^{-1}n^{(2-a)/4})}{\mathbb{P}(m(Y) > C(b)\epsilon, Y \in \mathcal{K}^\epsilon)}, \tag{7.22}$$

Comencemos ahora a trabajar con el denominador en las expresiones anteriores. Sea $\delta := C(b)\epsilon$. Note que

$$\begin{aligned}\mathbb{P}(m(Y) > \delta, Y \in \mathcal{K}^\epsilon) &= \mathbb{P}(Y \in \mathcal{K}^\epsilon \,|\, m(Y) > \delta)\mathbb{P}(m(Y) > \delta) \\ &= \mathbb{P}(Y \in \mathcal{K}^\epsilon \,|\, m(Y) > \delta)e^{-\delta n}. \end{aligned} \tag{7.23}$$

Dado que $Y_1, Y_2, ..., Y_n$ son i.i.d $\sim exp(1)$, se sigue de la propiedad de falta de memoria de la distribución exponencial que la distribución condicional de Y dado $m(Y) > \delta$ es la misma que la distribución de $Y + \delta\mathbf{1}$. Es decir, que

$$\mathbb{P}(Y \in \mathcal{K}^\epsilon \,|\, m(Y) > \delta) = \mathbb{P}(Y + \delta\mathbf{1} \in \mathcal{K}^\epsilon). \tag{7.24}$$

En efecto, consideremos la colección

$$\mathcal{F} := \{I_1 \times I_2 \times \cdots \times I_n \text{ donde } I_k =]t_k, +\infty[,\ t_k \in \mathbb{R}\}.$$

Es fácil ver que $\mathcal{F}$ es un π-sistema y que $\mathcal{F}$ genera la colección de borelianos del $\mathbb{R}^n$. En lo que sigue probaremos que

$$\mathbb{P}(Y \in A \,|\, m(Y) > \delta) = \mathbb{P}(Y + \delta\mathbf{1} \in A),$$

para todo $A \in \mathcal{F}$. Veamos, para $A \in \mathcal{F}$ tenemos que

$$\begin{aligned}
\mathbb{P}(Y \in A \,|\, m(Y) > \delta) &= \frac{\mathbb{P}(Y \in A\ ,\ Y_1 > \delta, ..., Y_n > \delta)}{\mathbb{P}(Y_1 > \delta, ..., Y_n > \delta)} \\
&= \frac{\mathbb{P}(Y_1 \in I_1, ..., Y_n \in I_n\ ,\ Y_1 > \delta, ..., Y_n > \delta)}{\mathbb{P}(Y_1 > \delta) \cdots \mathbb{P}(Y_n > \delta)} \\
&= \frac{\mathbb{P}(Y_1 \in I_1, Y_1 > \delta)}{\mathbb{P}(Y_1 > \delta)} \times \cdots \times \frac{\mathbb{P}(Y_n \in I_n, Y_n > \delta)}{\mathbb{P}(Y_n > \delta)} \\
&= \frac{\mathbb{P}(Y_1 > t_1, Y_1 > \delta)}{\mathbb{P}(Y_1 > \delta)} \times \cdots \times \frac{\mathbb{P}(Y_n > t_n, Y_n > \delta)}{\mathbb{P}(Y_n > \delta)},
\end{aligned} \tag{7.25}$$

y que

$$\begin{aligned}
\mathbb{P}(Y + \delta\mathbf{1} \in A) &= \mathbb{P}(Y + \delta\mathbf{1} \in]t_1, +\infty[\times \cdots \times]t_n, +\infty[) \\
&= \mathbb{P}(Y_1 + \delta > t_1, ..., Y_n + \delta > t_n) \\
&= \mathbb{P}(Y_1 + \delta > t_1) \times \cdots \times \mathbb{P}(Y_n + \delta > t_n).
\end{aligned} \tag{7.26}$$

No es difícil verificar que si $\delta < t_i$ o si $\delta \geq t_i$ $\forall i$, entonces (7.25) y (7.26) son iguales. Luego usando el Teorema 1.2.4 de [1] se sigue la igualdad deseada (7.24). Y por consiguiente tenemos de (7.23) y (7.24) que

$$\mathbb{P}(m(Y) > \delta, Y \in \mathcal{K}^\epsilon) = \mathbb{P}(Y + \delta\mathbf{1} \in \mathcal{K}^\epsilon) e^{-\delta n}. \tag{7.27}$$

Ahora acotaremos inferiormente a $\mathbb{P}(Y + \delta\mathbf{1} \in \mathcal{K}^\epsilon)$. Para lo cual empecemos definiendo el conjunto E como sigue

$$E := \left\{ \left|\mu(Y) - \left(1 - \delta + \frac{3}{2}\epsilon\right)\right| < \epsilon^2,\ \left|\mu_2(Y) - \left(b - 2\delta + \frac{b+1}{2}\epsilon\right)\right| < \epsilon^2 \right\},$$

y probaremos que para todo n suficientemente grande

$$E \subseteq \{Y + \delta\mathbf{1} \in \mathcal{K}^\epsilon\}. \tag{7.28}$$

Comencemos primero notando que

$$\mu(Y + \delta\mathbf{1}) = \mu(Y) + \delta, \quad \mu_2(Y + \delta\mathbf{1}) = \mu_2(Y) + 2\delta\mu(Y) + \delta^2.$$

Ahora bajo E, se sigue que

$$1 + \frac{3}{2}\epsilon - \epsilon^2 < \mu(Y) + \delta < 1 + \frac{3}{2}\epsilon + \epsilon^2,$$

y por lo tanto, si n es suficientemente grande tal que $\epsilon^2 < \epsilon/2$, tenemos

$$1 + \epsilon < \mu(Y + \delta\mathbf{1}) < 1 + 2\epsilon. \tag{7.29}$$

De nuevo, bajo E, tenemos

$$\begin{aligned} |\mu_2(Y) \quad &+ \quad 2\delta\mu(Y) + \delta^2 - (b + \frac{b+1}{2}\epsilon)| \\ &\leq \quad |\mu_2(Y) - (b - 2\delta + \frac{b+1}{2}\epsilon)| + 2\delta|\mu(Y) - 1| + \delta^2 \\ &\leq \quad C\epsilon^2, \end{aligned}$$

y por lo tanto, si n es suficientemente grande, tenemos

$$\begin{aligned} b + \epsilon \quad &< \quad b + \frac{b+1}{2}\epsilon - C\epsilon^2 \\ &< \quad \mu_2(Y + \delta\mathbf{1}) < b + \frac{b+1}{2}\epsilon + C\epsilon^2 < b + b\epsilon, \end{aligned} \tag{7.30}$$

y así de (7.29) y (7.30) obtenemos la inclusión deseada (7.28), provisto que n es lo suficientemente grande. Ahora definamos μ^- y μ_2^- de la siguiente manera:

$$\mu^-(Y) := \frac{1}{n}\sum_{i=2}^{n} Y_i, \quad \mu_2^-(Y) := \frac{1}{n}\sum_{i=2}^{n} Y_i^2.$$

También definamos

$$\begin{aligned} E' := \quad & \left\{\left|\mu^-(Y) - (1 - \delta + \frac{3}{2}\epsilon - qn^{-1/2})\right| < \frac{1}{2}\epsilon^2\right\} \\ & \cap \left\{\left|\mu_2^-(Y) - (2 - 2\delta + \frac{b+1}{2}\epsilon)\right| < \frac{1}{2}\epsilon^2\right\} \\ & \cap \left\{\left|Y_1^2 - q^2 n\right| < \frac{1}{2}\epsilon^2\right\}. \end{aligned}$$

Luego probaremos que

$$E' \subseteq E \tag{7.31}$$

para todo n suficientemente grande. Veamos, bajo E' se tiene que

$$\begin{aligned}
\left|\mu_2(Y) - (b - 2\delta + \frac{b+1}{2}\epsilon)\right| &= \left|\mu_2^-(Y) + \frac{Y_1^2}{n} - (b - 2\delta + \frac{b+1}{2}\epsilon)\right| \\
&= \left|\mu_2^-(Y) - (2 - 2\delta + \frac{b+1}{2}\epsilon) + \frac{Y_1^2}{n} - (b-2)\right| \\
&\leq \left|\mu_2^-(Y) - (2 - 2\delta + \frac{b+1}{2}\epsilon)\right| + \frac{1}{n}\left|Y_1^2 - n(b-2)\right| \\
&< \frac{1}{2}\epsilon^2 + \frac{1}{n}\cdot\frac{\epsilon^2}{2} \\
&\leq \epsilon^2. \qquad (7.32)
\end{aligned}$$

De nuevo, bajo E'

$$\begin{aligned}
|Y_1 - qn^{1/2}| &= \frac{|Y_1^2 - q^2 n|}{Y_1 + qn^{1/2}} \\
&\leq \left(\frac{1}{2}\epsilon^2\right)\left(\frac{1}{qn^{1/2}}\right) \\
&= Cn^{-1/2}\epsilon^2,
\end{aligned}$$

donde la primera desigualdad es porque $Y_1 > 0$ y $Y_1 \sim exp(1)$. Además

$$\begin{aligned}
\left|\mu(Y) - (1 - \delta + \frac{3}{2}\epsilon)\right| &= \left|\mu^-(Y) + \frac{Y_1}{n} - (1 - \delta + \frac{3}{2}\epsilon)\right| \\
&= \left|\mu^-(Y) - (1 - \delta + \frac{3}{2}\epsilon - qn^{-1/2}) + (\frac{Y_1}{n} - qn^{-1/2})\right| \\
&\leq \left|\mu^-(Y) - (1 - \delta + \frac{3}{2}\epsilon - qn^{-1/2})\right| + \frac{1}{n}\left|Y_1 - qn^{1/2}\right| \\
&< \frac{1}{2}\epsilon^2 + \frac{1}{n}Cn^{-1/2}\epsilon^2 \\
&= \epsilon^2\left(\frac{1}{2} + \frac{C}{n^{3/2}}\right) \\
&\leq \epsilon^2\left(\frac{1}{2} + \frac{1}{2}\right) = \epsilon^2 \qquad (7.33)
\end{aligned}$$

donde la última desigualdad se cumple para todo n suficientemente grande tal que $Cn^{-3/2} \leq 1/2$. Finalmente concluimos de (7.32) y (7.33) que $E' \subseteq E$, lo cual prueba la inclusión (7.31).

Ahora procedemos a encontrar una cota inferior para $\mathbb{P}(E')$. Para lo cual usamos el hecho que Y_1 es independiente de $(Y_2, ..., Y_n)$, $\mathbb{E}(Y_i) = 1$ y $\mathbb{E}(Y_i^2) = 2$. Y luego aplicando el Lema 5.4 al par $(\mu^-(Y), \mu_2^-(Y))$ concluimos que para todo n suficientemente grande

$$\mathbb{P}(E') \geq C^{-1} n^{1/2} \epsilon^6 e^{-qn^{1/2}}, \tag{7.34}$$

donde C es una constante adecuada que depende de b. En efecto, por comodidad sea

$$\begin{aligned} E^* := \quad & \left\{ \left| \mu^-(Y) - (1 - \delta + \frac{3}{2}\epsilon - qn^{-1/2}) \right| < \frac{1}{2}\epsilon^2 \right\} \\ & \cap \left\{ \left| \mu_2^-(Y) - (2 - 2\delta + \frac{b+1}{2}\epsilon) \right| < \frac{1}{2}\epsilon^2 \right\}. \end{aligned}$$

Note que E' implica E^*, y además este último conjunto es equivalente a :

$$\begin{aligned} -\frac{\epsilon^2}{2} + 1 - \delta + \frac{3\epsilon}{2} - qn^{-1/2} \quad & < \mu^-(Y) = \tfrac{1}{n}\textstyle\sum_{i=2}^n Y_i < \quad \frac{\epsilon^2}{2} + 1 - \delta + \frac{3\epsilon}{2} - qn^{-1/2}, \\ -\frac{\epsilon^2}{2} + 2 - 2\delta + \frac{b+1}{2}\epsilon \quad & < \mu_2^-(Y) = \tfrac{1}{n}\textstyle\sum_{i=2}^n Y_i^2 < \quad -\frac{\epsilon^2}{2} + 2 - 2\delta + \frac{b+1}{2}\epsilon, \end{aligned}$$

lo cual equivale a

$$a_n < \frac{1}{n-1} \sum_{i=2}^n Y_i < b_n,$$

$$a'_n < \frac{1}{n-1} \sum_{i=2}^n Y_i^2 < b'_n,$$

donde $a_n = \frac{n}{n-1}\left(\frac{-\epsilon^2}{2} + 1 - \delta + \frac{3\epsilon}{2} - qn^{-1/2}\right)$, $b_n = \frac{n}{n-1}\left(\frac{\epsilon^2}{2} + 1 - \delta + \frac{3\epsilon}{2} - qn^{-1/2}\right)$, $a'_n = \frac{n}{n-1}\left(\frac{-\epsilon^2}{2} + 2 - 2\delta + \frac{b+1}{2}\epsilon\right)$, $b'_n = \frac{n}{n-1}\left(\frac{-\epsilon^2}{2} + 2 - 2\delta + \frac{b+1}{2}\epsilon\right)$. Es fácil verificar que

- $\lim_{n\to+\infty}\sqrt{n}(a_n - 1) = \lim_{n\to+\infty}\sqrt{n}(b_n - 1) = -q$,
- $\lim_{n\to+\infty}\sqrt{n}(a'_n - 2) = \lim_{n\to+\infty}\sqrt{n}(b'_n - 2) = 0$.

Aplicando el Lema 5.4 se tiene que

$$\begin{aligned} \lim_{n\to+\infty} \frac{\mathbb{P}(E^*)}{n\epsilon^2\epsilon^2} \quad & = \quad \lim_{n\to+\infty} \frac{\mathbb{P}(a_n \leq \frac{1}{n-1}\sum_{i=2}^n Y_i \leq b_n,\ a'_n \leq \frac{1}{n-1}\sum_{i=2}^n Y_i^2 \leq b'_n)}{n\epsilon^2\epsilon^2} \\ & = \quad \rho(-q, 0) > 0, \end{aligned}$$

donde ρ es una densidad Gaussiana no degenerada, de lo cual se sigue que $\exists\, n_0 \in \mathbb{N}$ tal que si $n \geq n_0$, entonces

$$0 < \frac{\mathbb{P}(E^*)}{n\epsilon^4} < 2\rho(-q,0).$$

En la primera desigualdad anterior es posible escoger una constante $C > 0$ adecuada que depende de $\rho(-q,0)$ tal que

$$\frac{1}{C} \leq \frac{\mathbb{P}(E^*)}{n\epsilon^4}.$$

Y así,

$$\mathbb{P}(E^*) \geq C^{-1} n\epsilon^4, \tag{7.35}$$

por consiguiente tenemos que

$$\begin{aligned} \mathbb{P}(E') &= \mathbb{P}\Big(E^* \cap (|Y_1^2 - q^2 n| < \epsilon^2/2)\Big) \\ &= \mathbb{P}(E^*)\mathbb{P}\big(|Y_1^2 - q^2 n| < \epsilon^2/2\big) \\ &\geq C^{-1} n\epsilon^4 \mathbb{P}\big(|Y_1^2 - q^2 n| < \epsilon^2/2\big). \end{aligned} \tag{7.36}$$

En lo que sigue mostraremos que el evento $\big(|Y_1 - qn^{1/2}| < C^{-1}n^{-1/2}\epsilon^2\big)$ implica el evento $\big(|Y_1^2 - q^2 n| < \epsilon^2/2\big)$ y que la probabilidad del primer evento tiene como cota inferior a $Cn^{-1/2}\epsilon^2 e^{-qn^{1/2}}$. En efecto, debido a la desigualdad

$$|Y_1^2 - q^2 n| \leq (Y_1 - qn^{1/2})^2 + 2qn^{1/2}|Y_1 - qn^{1/2}|,$$

se sigue que: si $|Y_1 - qn^{1/2}| < C^{-1}n^{-1/2}\epsilon^2$, entonces

$$\begin{aligned} |Y_1^2 - q^2 n| &\leq C^{-2}n^{-1}\epsilon^4 + 2qn^{1/2}C^{-1}n^{-1/2}\epsilon^2 \\ &= C^{-2}n^{-1}\epsilon^4 + 2C^{-1}q\epsilon^2 \\ &= \frac{\epsilon^2}{2}(2C^{-2}n^{-1}\epsilon^2 + 4C^{-1}q) \\ &< \frac{\epsilon^2}{2}, \quad \text{(para } C \text{ adecuada).} \end{aligned}$$

Ahora procedemos a encontrar la cota inferior del primer evento. Para $\alpha, \beta \in \mathbb{R}$, con $0 < \beta < 1$, se tiene que

$$\begin{aligned} \mathbb{P}\left(|Y_1 - \alpha| < \beta\right) &= \int_{\{|x-\alpha|<\beta\}} e^{|x-\alpha|}e^{-x}dx \\ &\geq e^{-(\alpha+\beta)}2\beta \\ &= 2e^{-\beta}\beta e^{-\alpha} \geq \frac{2}{e}\beta e^{-\alpha}. \end{aligned}$$

Luego tomando $\alpha = qn^{1/2}$ y $\beta = C^{-1}n^{-1/2}\epsilon^2$ con n suficientemente grande tal que $\beta < 1$, se obtiene que

$$\begin{aligned} \mathbb{P}\left(|Y_1 - qn^{1/2}| < C^{-1}n^{-1/2}\epsilon^2\right) &\geq \frac{2}{e}C^{-1}n^{-1/2}\epsilon^2 e^{-qn^{1/2}} \\ &= Cn^{-1/2}\epsilon^2 e^{-qn^{1/2}}, \end{aligned} \tag{7.37}$$

y así

$$\mathbb{P}\big(|Y_1^2 - q^2 n| < \epsilon^2/2\big) \geq Cn^{-1/2}\epsilon^2 e^{-qn^{1/2}} \tag{7.38}$$

Luego de (7.38) en (7.36) se obtiene la cota inferior deseada para $\mathbb{P}(E')$ dada en (7.34). Por lo tanto utilizando convenientemente (7.27),(7.28), (7.31) y (7.34) tenemos que

$$\begin{aligned} \mathbb{P}(m(Y) > \delta, Y \in \mathcal{K}^\epsilon) &= \mathbb{P}(Y + \delta\mathbf{1} \in \mathcal{K}^\epsilon)e^{-\delta n} \\ &\geq \mathbb{P}(E)e^{-\delta n} \geq \mathbb{P}(E')e^{-\delta n} \\ &\geq C^{-1}n^{1/2}\epsilon^6 e^{-qn^{1/2}}e^{-\delta n} \\ &= (C^{-1}n^{-60}e^{-qn^{1/2}})(n^{1/2}e^{-C(b)n^{-9}}) \\ &\geq C^{-1}n^{-60}e^{-qn^{1/2}}e^{-C(b)} \\ &= (Ce^{C(b)})^{-1}n^{-60}e^{-qn^{1/2}} \end{aligned} \tag{7.39}$$

Combinando (7.39) con (7.21), se tiene que

$$\mathbb{P}(|M - qn^{1/2}| > Cn^{(2-a)/4}) \leq Ce^{-C^{-1}n^{(2-a)/4}},$$

pues

$$\begin{aligned}
\mathbb{P}(|M - qn^{1/2}| &> Cn^{(2-a)/4}) \\
&\leq \frac{C\exp(-qn^{1/2} - C^{-1}n^{(2-a)/4})}{\mathbb{P}(m(Y) > C(b)\epsilon, Y \in \mathcal{K}^{\epsilon})} \\
&\leq \frac{C\exp(-qn^{1/2} - C^{-1}n^{(2-a)/4})}{C^{-1}n^{-60}e^{-qn^{1/2}}} \\
&= \frac{C\exp(-C^{-1}n^{(2-a)/4})}{C^{-1}n^{-60}} \\
&= \left(\frac{Cn^{60}}{C^{-1}\exp\left(\frac{C^{-1}}{2}n^{(2-a)/4}\right)}\right)\exp\left(-\frac{C^{-1}}{2}n^{(2-a)/4}\right) \\
&\leq 2C\exp\left(-\frac{C^{-1}}{2}n^{(2-a)/4}\right)
\end{aligned}$$

donde la última desigualdad es porque $\frac{Cn^{60}}{C^{-1}\exp\left(\frac{C^{-1}}{2}n^{(2-a)/4}\right)}$ tiende a 0. Hemos probado así que

$$\mathbb{P}\left(\left|\frac{M}{qn^{1/2}} - 1\right| > Cn^{-a/4}\right) \to 0. \tag{7.40}$$

Y dado que $Cn^{-a/4}$ converge a 0, entonces (7.40) implica la convergencia en probabilidad

$$\frac{M}{qn^{1/2}} \to 1 \quad \textit{en probabilidad.}$$

Similarmente de (7.22) obtenemos

$$\mathbb{P}(M_2 > Cn^{(2-a)/4}) \leq Ce^{-C^{-1}n^{(2-a)/4}},$$

pues

$$
\begin{aligned}
\mathbb{P}(M_2 &> Cn^{(2-a)/4}) \\
&\leq \frac{C\exp(-qn^{1/2} - C^{-1}n^{(2-a)/4})}{\mathbb{P}(m(Y) > C(b)\epsilon, Y \in \mathcal{K}^\epsilon)} \\
&\leq \frac{C\exp(-qn^{1/2} - C^{-1}n^{(2-a)/4})}{C^{-1}n^{-60}e^{-qn^{1/2}}} \\
&= \frac{C\exp(-C^{-1}n^{(2-a)/4})}{C^{-1}n^{-60}} \\
&= \left(\frac{Cn^{60}}{C^{-1}\exp\left(\frac{C^{-1}}{2}n^{(2-a)/4}\right)}\right)\exp\left(-\frac{C^{-1}}{2}n^{(2-a)/4}\right) \\
&\leq 2C\exp\left(-\frac{C^{-1}}{2}n^{(2-a)/4}\right)
\end{aligned}
$$

donde la última desigualdad es porque $\frac{Cn^{60}}{C^{-1}\exp\left(\frac{C^{-1}}{2}n^{(2-a)/4}\right)}$ tiende a 0.

Hemos probado así que

$$\mathbb{P}\left(\frac{M_2^2}{n} > C^2n^{-a/2}\right) \to 0. \tag{7.41}$$

Ya que $C^2n^{-a/2}$ converge a 0, entonces (7.41) implica la convergencia en probabilidad

$$\frac{M_2^2}{n} \longrightarrow 0 \quad \textit{en probabilidad.}$$

Esto completa la prueba de las partes (b) y (c). □

7.2. Prueba del Teorema 2.4, parte (a)

Procedemos exactamente como en la prueba de la parte (a) del Teorema 2.3. Sea $h : \mathbb{R} \to \mathbb{R}$ una función satisfaciendo $|h(x)| \leq 1$ y $|h(x) - h(y)| \leq L|x - y|$ para todo $x, y \in \mathbb{R}$, donde L es alguna constante positiva, definamos también una función $g : \mathbb{R}^n \to \mathbb{R}$ como

$$g(x) := \frac{1}{n}\sum_{i=1}^{n} h(x_i).$$

Sea $a := \mathbb{E}\, h(Y_1)$. Luego usando la desigualdad de Hoeffding, obtenemos que para cualquier $t > 0$,

$$\mathbb{P}(|g(Y) - a| > t) \leq 2e^{-nt^2/2}.$$

Usando (7.39) y la Proposición 5.3, obtenemos la siguiente analogía de (6.26):

$$\mathbb{P}(|g(Y)-a|>t) \leq Ce^{C\sqrt{n}}e^{-n(t-C_3(b)L\epsilon n)^2/2},$$

pues,

$$\begin{aligned}
\mathbb{P}(|g(X)-a|>t) &\leq \frac{e^{2B\epsilon n}\mathbb{P}(|g(Y)-a|>t-C_3(b)L\epsilon n)}{\mathbb{P}(m(Y)>C(b)\epsilon, Y\in\mathcal{K}^\epsilon)} \\
&\leq \frac{e^{2B\epsilon n}e^{-n(t-C_3(b)L\epsilon n)^2/2}}{C^{-1}n^{-60}e^{-qn^{1/2}}} \\
&\leq 2Ce^{2B}n^{60}e^{qn^{1/2}}e^{-n(t-C_3(b)L\epsilon n)^2/2} \\
&\leq Ce^{C\sqrt{n}}e^{-n(t-C_3(b)L\epsilon n)^2/2} \qquad (7.42)
\end{aligned}$$

donde la última desigualdad se cumple para todo n suficientemente grande. Afirmamos ahora que la cota cola en (7.42) decae rápidamente en el régimen $t>An^{-1/4}$, donde A es alguna constante que depende de b, y por lo cual

$$\mathbb{E}\,|g(X)-a| \leq Cn^{-1/4}. \qquad (7.43)$$

Pues,

$$\begin{aligned}
\mathbb{E}\,|g(X)-a| &= \int_0^\infty \mathbb{P}\left(|g(X)-a|>t\right)dt \\
&= \int_0^{An^{-1/4}} \mathbb{P}\left(|g(X)-a|>t\right)dt + \int_{An^{-1/4}}^\infty \mathbb{P}\left(|g(X)-a|>t\right)dt \\
&\leq An^{-1/4} + An^{-1/4}\int_1^\infty \mathbb{P}\left(|g(X)-a|>An^{-1/4}s\right)ds. \qquad (7.44)
\end{aligned}$$

Pero por (7.42)

$$\begin{aligned}
\mathbb{P}\left(|g(X)-a|>An^{-1/4}s\right)ds &\leq Ce^{C\sqrt{n}}e^{-n(An^{-1/4}s-C_3(b)L\epsilon n)^2/2} \\
&= Ce^{C\sqrt{n}}e^{-\sqrt{n}(As-C_3(b)L\epsilon n^{1+1/4})^2/2} \\
&\leq Ce^{C\sqrt{n}}e^{-\sqrt{n}(As-\frac{As}{2})^2/2} \\
&= Ce^{C\sqrt{n}}e^{-\sqrt{n}A^2s^2/8} \\
&= Ce^{-\sqrt{n}(-C+A^2s^2/8)},
\end{aligned}$$

donde la última desigualdad anterior se cumple para todo n suficientemente grande. Entonces tomando A tal que $\frac{A^2}{16} > C$ se sigue que

$$\begin{aligned}\int_1^\infty \mathbb{P}\left(|g(X)-a| > An^{-1/4}s\right) ds &\leq \int_1^\infty Ce^{-\sqrt{n}(-\frac{A^2s^2}{16}+\frac{A^2s^2}{8})}ds \\ &\leq \int_1^\infty Ce^{-\frac{A^2s^2}{16}}ds < \infty. \end{aligned} \tag{7.45}$$

Finalmente de (7.44) y (7.45) se sigue (7.43). De manera similar como se obtuvo (6.32) obtenemos,

$$|\mathbb{E}\,h(X_1) - a| \leq Cn^{-1/4}.$$

Haciendo $n \to \infty$, se sigue que $\mathbb{E}\,h(X_1) \to \mathbb{E}\,h(Y_1)$, luego aplicando la Proposición 9.9, para $k = 1$, concluimos que $X_1 \xrightarrow{ley} Y_1$.

La distribución conjunta de $(X_1, ..., X_k)$ es realizada también similarmente. Considerando k funciones $h_1, ..., h_k$ tal que $|h_i(x)| \leq 1$ y $|h_i(x) - h_i(y)| \leq L|x-y|$. También definamos $g_1, ..., g_k$ y $a_1, ..., a_k$ como

$$g_i(x) = \frac{1}{n}\sum_{j=1}^n h_i(x_j),$$

y $a_i = \mathbb{E}\,h_i(Y_i)$. Luego por un argumento telescópico, de manera similar a (6.33), obtenemos

$$\mathbb{E}\left|\prod_{i=1}^k g_i(X) - \prod_{i=1}^k a_i\right| \leq k \max_{1\leq i\leq n} \mathbb{E}\,|g_i(X) - a_i| \leq Ckn^{-1/4}, \tag{7.46}$$

y de forma similar a (6.34) obtenemos también

$$\begin{aligned}\mathbb{E}\left|\prod_{i=1}^k g_i(X) - A\right| &\geq \left|\frac{1}{n^k}\sum_{1\leq i_1,...,i_k\leq n} [\mathbb{E}\,(h_1(X_{i_1})\cdots h_k(X_{i_k})) - A]\right| \\ &\geq |\mathbb{E}\,(h_1(X_1)\cdots h_k(X_k)) - A| + \mathcal{O}(1/n), \end{aligned} \tag{7.47}$$

donde $A := \prod_{i=1}^k a_i$. Y finalmente de (7.46), (7.47) y la Proposición 9.9 se concluye la prueba de la parte (a) del teorema.

Capítulo 8

Prueba del Teorema 2.5

La prueba del Teorema 2.5 está básicamente contenida en la pruebas anteriores. Comencemos considerando el caso

$$1 < b \leq 2.$$

De forma similar a la prueba de la parte (a) del Teorema 2.3, es vez de tomar una h fija tomemos

$$h_n(x) := \begin{cases} 1 & \text{si } x < x_0 \ , \\ 1 - (x - x_0)n^5 & \text{si } x_0 \leq x < x_0 + n^{-5} \ , \\ 0 & \text{si } x \geq x_0 + n^{-5} \ , \end{cases}$$

donde $x_0 > 0$ es un número real fijo pero arbitrario. Es fácil verificar que $|h_n(x)| \leq 1$ y $|h_n(x) - h_n(y)| \leq n^5|x - y|$ para todo x, y. De aquí en adelante podemos proceder exactamente como en la prueba de (6.26) (tomando $L = n^5$) y concluir que

$$|\mathbb{E}\, h_n(X_1) - \mathbb{E}\, h_n(Y_1)| \leq C\sqrt{\frac{\log n}{n}}. \tag{8.1}$$

Ahora hallaremos una cota para $|\mathbb{P}(X_1 \leq x_0) - \mathbb{P}(Y_1 \leq x_0)|$. Dado que $1_{]-\infty, x_0]} \leq h_n$ entonces $1_{\{X_1 \leq x_0\}} \leq h_n(X_1)$ y así

$$\mathbb{P}(X_1 \leq x_0) = \mathbb{E}\, 1_{\{X_1 \leq x_0\}} \leq \mathbb{E}\, h_n(X_1). \tag{8.2}$$

Por otro lado, $h_n \leq 1_{]-\infty,x_0+n^{-5}]}$ entonces $h_n(Y_1) \leq 1_{]-\infty,x_0+n^{-5}]}(Y_1)$. Luego

$$\begin{aligned}
\mathbb{E}\, h_n(Y_1) &\leq \mathbb{E}\, 1_{]-\infty,x_0+n^{-5}]}(Y_1) \\
&= \mathbb{E}\, 1_{]-\infty,x_0]}(Y_1) + \mathbb{E}\, 1_{]x_0,x_0+n^{-5}]}(Y_1) \\
&= \mathbb{P}(Y_1 \leq x_0) + \int_{-\infty}^{\infty} 1_{]x_0,x_0+n^{-5}]}(t)\rho_{Y_1}(t)dt \\
&= \mathbb{P}(Y_1 \leq x_0) + \int_{x_0}^{x_0+n^{-5}} \rho_{Y_1}(t)dt \\
&\leq \mathbb{P}(Y_1 \leq x_0) + \int_{x_0}^{x_0+n^{-5}} M dt \quad \text{(pues } Y_1 \text{ tiene densidad acotada)} \\
&= \mathbb{P}(Y_1 \leq x_0) + Mn^{-5}. \qquad (8.3)
\end{aligned}$$

Luego de (8.2), (8.1) y (8.3)

$$\begin{aligned}
\mathbb{P}(X_1 \leq x_0) &\leq \mathbb{E}\, h_n(X_1) \\
&\leq \mathbb{E}\, h_n(Y_1) + C\sqrt{\frac{\log n}{n}} \\
&\leq \mathbb{P}(Y_1 \leq x_0) + Mn^{-5} + C\sqrt{\frac{\log n}{n}} \\
&\leq \mathbb{P}(Y_1 \leq x_0) + (M+C)\sqrt{\frac{\log n}{n}}. \qquad (8.4)
\end{aligned}$$

Ahora modifiquemos ligeramente la definición de h_n reemplazando x_0 con $x_0 - n^{-5}$. Llamemos esta nueva función $\widetilde{h}_n$, donde

$$\widetilde{h}_n(x) := \begin{cases} 1 & \text{si } x < x_0 - n^{-5}\ , \\ -(x-x_0)n^5 & \text{si } x_0 - n^{-5} \leq x < x_0\ , \\ 0 & \text{si } x \geq x_0\ . \end{cases}$$

Luego (8.1) se cumple también para esta función. Además $1_{]-\infty,x_0]} \geq \widetilde{h}_n$ entonces $1_{\{X_1 \leq x_0\}} \geq \widetilde{h}_n(X_1)$ y así

$$\mathbb{P}(X_1 \leq x_0) = \mathbb{E}\, 1_{\{X_1 \leq x_0\}} \geq \mathbb{E}\, \widetilde{h}_n(X_1). \qquad (8.5)$$

Por otro lado $\widetilde{h}_n \geq 1_{]-\infty, x_0 - n^{-5}]}$ entonces $\widetilde{h}_n(Y_1) \geq 1_{]-\infty, x_0 - n^{-5}]}(Y_1)$. Luego

$$\begin{aligned}
\mathbb{E}\,\widetilde{h}_n(Y_1) &\geq \mathbb{E}\,1_{]-\infty, x_0 - n^{-5}]}(Y_1) \\
&= \mathbb{E}\,\left(1_{]-\infty, x_0]}(Y_1) - 1_{]x_0 - n^{-5}, x_0]}(Y_1)\right) \\
&= \mathbb{E}\,1_{]-\infty, x_0]}(Y_1) - \mathbb{E}\,1_{]x_0 - n^{-5}, x_0]}(Y_1) \\
&= \mathbb{P}(Y_1 \leq x_0) - \int_{-\infty}^{\infty} 1_{]x_0 - n^{-5}, x_0]}(t)\rho_{Y_1}(t)dt \\
&= \mathbb{P}(Y_1 \leq x_0) - \int_{x_0 - n^{-5}}^{x_0} \rho_{Y_1}(t)dt \\
&\geq \mathbb{P}(Y_1 \leq x_0) - \int_{x_0 - n^{-5}}^{x_0} N dt \quad \text{(pues } Y_1 \text{ tiene densidad acotada)} \\
&= \mathbb{P}(Y_1 \leq x_0) - N n^{-5}.
\end{aligned} \tag{8.6}$$

Luego de (8.5), (8.1) y (8.6)

$$\begin{aligned}
\mathbb{P}(X_1 \leq x_0) &\geq \mathbb{E}\,\widetilde{h}_n(X_1) \\
&\geq \mathbb{E}\,\widetilde{h}_n(Y_1) - C\sqrt{\frac{\log n}{n}} \\
&\geq \mathbb{P}(Y_1 \leq x_0) - N n^{-5} - C\sqrt{\frac{\log n}{n}} \\
&\geq \mathbb{P}(Y_1 \leq x_0) - (N + C)\sqrt{\frac{\log n}{n}}.
\end{aligned} \tag{8.7}$$

De (8.4) y (8.7) se concluye que

$$\left|\mathbb{P}(X_1 \leq x_0) - \mathbb{P}(Y_1 \leq x_0)\right| \leq \widehat{C}\sqrt{\frac{\log n}{n}}, \tag{8.8}$$

donde $\widehat{C} = \text{máx}\{M + C, N + C\}$. Dado que $x_0 > 0$ fue escogido arbitrariamente y no depende de n, se sigue que

$$\sup_{x_0} \left|\mathbb{P}(X_1 \leq x_0) - \mathbb{P}(Y_1 \leq x_0)\right| \leq \widehat{C}\sqrt{\frac{\log n}{n}}, \tag{8.9}$$

lo cual prueba el teorema para $k = 1$. Para el caso general se procede de manera similar a la prueba del Teorema 2.3, donde obtuvimos

$$\left|\mathbb{E}\,(h_1(X_1)\cdots h_k(X_k)) - \mathbb{E}\,h_1(Y_1)\cdots\mathbb{E}\,h_k(Y_k)\right| \le Ck\sqrt{\frac{\log n}{n}} - \mathcal{O}(1/n). \tag{8.10}$$

Fijemos ahora $t_1, ..., t_k$ reales positivos y sean $h_n^{(1)}, ..., h_n^{(k)}$ sus funciones asociadas como en el caso $k = 1$. Luego de la desigualdad

$$1_{]-\infty,t_i]} \le h_n^{(i)}, \quad i = 1, ..., k,$$

se sigue que

$$1_{\cap_{i=1}^k\{X_i\le t_i\}} = \prod_{i=1}^k 1_{\{X_i\le t_i\}} = \prod_{i=1}^k 1_{]-\infty,t_i]}(X_i) \le \prod_{i=1}^k h_n^{(i)}(X_i),$$

y así

$$\begin{aligned}\mathbb{P}(X_1 \le t_1, ..., X_k \le t_k) &= \mathbb{E}\,1_{\cap_{i=1}^k\{X_i\le t_i\}} \\ &\le \mathbb{E}\prod_{i=1}^k h_n^{(i)}(X_i) \\ &= \mathbb{E}\left(h_n^{(1)}(X_1)\cdots h_n^{(k)}(X_k)\right). \end{aligned} \tag{8.11}$$

Usando (8.11) y (8.10) se tiene

$$\begin{aligned}\mathbb{P}(X_1 \le t_1, ..., X_k \le t_k) &\le \mathbb{E}\left(h_n^{(1)}(X_1)\cdots h_n^{(k)}(X_k)\right) \\ &\le \mathbb{E}\left(h_n^{(1)}(Y_1)\cdots h_n^{(k)}(Y_k)\right) + Ck\sqrt{\frac{\log n}{n}} - \mathcal{O}(1/n)\end{aligned} \tag{8.12}$$

Pero

$$\begin{aligned}\mathbb{E}\left(h_n^{(1)}(Y_1)\cdots h_n^{(k)}(Y_k)\right) &\le [\mathbb{P}(Y_1 \le t_1) + C_1 n^{-5}]\cdots[\mathbb{P}(Y_k \le t_k) + C_k n^{-5}] \\ &\le \mathbb{P}(Y_1 \le t_1)\times\cdots\times\mathbb{P}(Y_k \le t_k) + C\mathcal{O}(1/n),\end{aligned} \tag{8.13}$$

entonces de (8.12) y (8.13) se sigue que

$$\mathbb{P}(X_1 \le t_1, ..., X_k \le t_k) \le \mathbb{P}(Y_1 \le t_1)\times\cdots\times\mathbb{P}(Y_k \le t_k) + Ck\sqrt{\frac{\log n}{n}} + C\mathcal{O}(1/n). \tag{8.14}$$

Ahora consideremos k funciones $\tilde{h}_n^{(1)}, ..., \tilde{h}_n^{(k)}$ como en el caso $k = 1$. Luego tenemos que $1_{]-\infty,t_i]} \geq \tilde{h}_n^{(i)}$ para $i = 1, 2, ..., k$. De lo cual se sigue que

$$1_{\cap_{i=1}^k \{X_i \leq t_i\}} = \prod_{i=1}^k 1_{\{X_i \leq t_i\}} = \prod_{i=1}^k 1_{]-\infty,t_i]}(X_i) \geq \prod_{i=1}^k \tilde{h}_n^{(i)}(X_i),$$

y así

$$\begin{aligned} \mathbb{P}(X_1 \leq t_1, ..., X_k \leq t_k) &= \mathbb{E}\, 1_{\cap_{i=1}^k \{X_i \leq t_i\}} \\ &\geq \mathbb{E} \prod_{i=1}^k \tilde{h}_n^{(i)}(X_i) \\ &= \mathbb{E}\left(\tilde{h}_n^{(1)}(X_1) \cdots \tilde{h}_n^{(k)}(X_k)\right). \end{aligned} \tag{8.15}$$

Usando (8.10) y (8.15) se tiene

$$\begin{aligned} \mathbb{P}(X_1 \leq t_1, ..., X_k \leq t_k) &\geq \mathbb{E}\left(\tilde{h}_n^{(1)}(X_1) \cdots \tilde{h}_n^{(k)}(X_k)\right) \\ &\geq \mathbb{E}\left(\tilde{h}_n^{(1)}(Y_1) \cdots \tilde{h}_n^{(k)}(Y_k)\right) - Ck\sqrt{\frac{\log n}{n}} + \mathcal{O}(1/n) \end{aligned} \tag{8.16}$$

Pero

$$\begin{aligned} \mathbb{E}\left(\tilde{h}_n^{(1)}(Y_1) \cdots \tilde{h}_n^{(k)}(Y_k)\right) &= \mathbb{E}\left(\tilde{h}_n^{(1)}(Y_1)\right) \times \cdots \times \mathbb{E}\left(\tilde{h}_n^{(k)}(Y_k)\right) \\ &\geq [\mathbb{P}(Y_1 \leq t_1) + C_1 n^{-5}] \times \cdots \times [\mathbb{P}(Y_k \leq t_k) + C_k n^{-5}] \\ &\geq \mathbb{P}(Y_1 \leq t_1) \times \cdots \times \mathbb{P}(Y_k \leq t_k) + C\mathcal{O}(1/n), \end{aligned} \tag{8.17}$$

entonces de (8.16) y (8.17) se sigue que

$$\mathbb{P}(X_1 \leq t_1, ..., X_k \leq t_k) \geq \mathbb{P}(Y_1 \leq t_1) \times \cdots \times \mathbb{P}(Y_k \leq t_k) - Ck\sqrt{\frac{\log n}{n}} + C\mathcal{O}(1/n). \tag{8.18}$$

Finalmente de las ecuaciones (8.14),(8.18) y de la arbitrariedad de $t_1, ..., t_k$ se completa la prueba.

Cuando $b > 2$, la prueba es exactamente la misma, excepto que la cota en (8.1) se convierte en $Cn^{-1/4}$.

Capítulo 9

Anexos

En este capítulo trataremos de algunos resultados usados durante el texto. 2

Fijemos un espacio vectorial de dimensión finita con producto interno

$$(V, \langle \cdot\,,\cdot\rangle)\ .$$

Denotaremos por $\|v\|$ a la norma de cada vector $v \in V$ correspondiendo al producto interno fijado. Por cada $r > 0$ definimos la esfera $\mathbb{S}_r$ de radio $\sqrt{r}$ como

$$\mathbb{S}_r \ := \ \{x \in V \mid \langle x, x\rangle = r\}\ .$$

Definición 9.1. *La aplicación $A : V \to V$ es llamada una isometría lineal si A es lineal y preserva la norma, i.e.*

$$\|Av\| = \|v\|\ , \quad \textit{para todo } v \in V\,.$$

Claramente cada esfera $\mathbb{S}_r$ es invariante por isometrías lineales. El siguiente resultado fue usado para definir la distribución uniforme sobre $\mathcal{K}$. Considere el σ-álgebra de borelianos $\mathcal{B}_V$, que es el σ-álgebra generado por los abiertos según la norma $\|\cdot\|$.

Teorema 9.2. *Por cada $r > 0$ existe una única medida λ sobre $(V, \mathcal{B}_V)$ tal que :*

1. $\lambda(\mathbb{S}_r) = 1$.
2. *λ es invariante por todas las isometrías lineales, i.e. $\lambda \circ A^{-1} = \lambda$ para toda isometría lineal $A : V \to V$.*

Demostración. Comenzaremos probando la unicidad. Sean μ, ν dos medidas sobre V tales que cumplen las condiciones 1 y 2 del teorema. Considere la función característica φ_μ de μ:

$$\varphi_\mu(x) \ := \ \int_V e^{i\langle x,y\rangle}\mu(dy) \ , \quad x \in V \ .$$

Afirmamos que φ_μ es invariante por la acción de cualquier isometría lineal, es decir

$$\varphi_\mu(Ax) \ = \ \varphi_\mu(x) \ .$$

para toda $A : V \to V$ isometría lineal. En efecto:

$$\begin{aligned} \varphi_\mu(Ax) &= \int_V e^{i\langle Ax,y\rangle}\mu(dy) \\ &= \int_V e^{i\langle x,A^T y\rangle}\mu(dy) \\ &= \int_V e^{i\langle x,A^{-1} y\rangle}\mu(dy) \\ &= \int_V e^{i\langle x,z\rangle}\mu(dz) \\ &= \varphi_\mu(x). \end{aligned}$$

Dado que para todo par $x, y \in V$ tal que $\|x\| = \|y\|$ siempre existe una isometría lineal $A : V \to V$ tal que $Ax = y$ entonces φ_μ debe ser constante en cada esfera $\mathbb{S}_r$. En otras palabras, existe una función $g : [0, \infty[\to \mathbb{C}$ tal que

$$\varphi_\mu(x) \ = \ g(\|x\|) \ , \quad \forall x \in V \ .$$

Para la otra medida ν podemos concluir del mismo modo que

$$\varphi_\nu(x) \ = \ h(\|x\|) \ , \quad \forall x \in V \ .$$

para alguna función $h : [0, \infty[\to \mathbb{C}$. Para cualquier $s > 0$ arbitrario tenemos

$$g(sr) \ = \ \int_{\mathbb{S}_r} \varphi_\mu(sx)\nu(dx) \ = \ \int_V \varphi_\mu(sy)\nu(dy)$$

debido a que φ_μ es constante sobre $\mathbb{S}_{sr}$ y que la ν-medida del complemento de $\mathbb{S}_r$ es cero. Usando la definición de φ_μ y luego el Teorema de Fubini tenemos que la última expresión coincide con

$$\int_V \int_V e^{i\langle x,sy\rangle}\mu(dx)\nu(dy) \ = \ \int_V \int_V e^{i\langle sx,y\rangle}\nu(dy)\mu(dx) \ = \ \int_V \varphi_\nu(sx)\mu(dx) \ .$$

Ya que la μ-medida del complemento de la esfera $\mathbb{S}_r$ es cero y φ_ν es constante sobre $\mathbb{S}_r$ la última expresión coincide con

$$\int_{\mathbb{S}_r} \varphi_\nu(sx)\mu(dx) \;=\; h(sr)\;.$$

Juntando las última igualdades obtenemos

$$g(sr) \;=\; h(sr)\;, \quad \forall s > 0\;.$$

Se sigue que $g = h$ y así $\varphi_\mu = \varphi_\nu$ lo cual completa la prueba de unicidad.

Para probar la existencia considere la medida de Lebesgue $\boldsymbol{m}_V$ sobre $(V, \mathcal{B}_V)$ obtenida como la medida inducida por la medida de Lebesgue sobre el espacio euclideano $\mathbb{R}^n$, de la misma dimensión de V, y cualquier isomorfismo lineal entre V y $\mathbb{R}^n$. Ahora, para cada $D \subseteq \mathbb{S}_r$ definimos

$$\Gamma(D) \;=\; \big\{rx : x \in D\,,\; 0 \le r \le 1\big\}\;.$$

Es fácil verificar que si $D \in \mathcal{B}_V$ entonces $\Gamma(D) \in \mathcal{B}_V$ y que $\Gamma(\mathbb{S}_r)$ es la bola unitaria cerrada y por tanto tiene $\boldsymbol{m}_V$-medida estrictamente positiva y finita. Estas observaciones nos permiten definir

$$\lambda(E) \;:=\; \frac{\boldsymbol{m}_V(\Gamma(E \cap \mathbb{S}_r))}{\boldsymbol{m}_V(\Gamma(\mathbb{S}_r))}\;, \quad E \in \mathcal{B}_V\;.$$

Es simple verificar que λ es en efecto una medida de probabilidad. Además, ya que $\boldsymbol{m}$ es invariante por isometrías lineales y

$$\Gamma\big((A^{-1}E) \cap \mathbb{S}_r\big) \;=\; A^{-1}\Gamma(E \cap \mathbb{S}_r)\;, \quad E \in \mathcal{B}_V$$

para cualquier isometría lineal $A : V \to V$, entonces λ resulta también invariante por isometrías lineales como queríamos. $\square$

Definición 9.3. *Una medida $\lambda : \mathcal{B}_V \to [0, \infty]$, donde $\mathcal{B}_V$ es la colección de los borelianos de V, se llama **distribución uniforme** sobre la esfera $\mathbb{S}_r$ si satisface las condiciones 1 y 2 del teorema anterior.*

El siguiente resultado fue también usado en la construcción de la distribución uniforme sobre $\mathcal{K}$. Nos dice que si una medida ν es invariante por isometrías lineales y la restringimos a un boreliano W que es invariante por las isometrías lineales entonces la restricción de ν a W es también invariante por isometrías lineales.

Lema 9.4. Sean ν una medida sobre $(V, \mathcal{B}_V)$, $W \in \mathcal{B}_V$ y $A : V \to V$ una isometría lineal tal que $A(W) = W$. Si $\nu \circ A^{-1} = \nu$ entonces

$$\tilde{\nu} \circ A^{-1} = \tilde{\nu},$$

donde

$$\tilde{\nu}(E) \;=\; \nu(E \cap W)\,, \quad E \in \mathcal{B}_V\,.$$

Demostración. Sea $B \in \mathcal{B}_V$, entonces

$$\begin{aligned}
\tilde{\nu} \circ A^{-1}(B) &= \nu\left(A^{-1}(B) \cap W\right) \\
&= \nu\left(A^{-1}(B) \cap A^{-1}(W)\right) \\
&= \nu \circ A^{-1}(B \cap W) \\
&= \nu(B \cap W) \\
&= \tilde{\nu}(B).
\end{aligned}$$

□

El siguiente resultado fue usado al final de la prueba del Lema 5.7. Como no encontramos una referencia preferimos escribir una demostración en este trabajo.

Lema 9.5. Sean μ una probabilidad sobre $(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n})$ tal que $\mu(S^c) = 0$ para algún $S \in \mathcal{B}_{\mathbb{R}^n}$. Supongamos que $\forall\, x \in S$, $\exists\, \delta > 0$ tal que

$$\mu(B(x,r)) \le C\, \boldsymbol{m}^n(B(x,r)) \quad \forall r \le \delta,$$

donde $C > 0$ es una constante que no depende de x y $\boldsymbol{m}^n$ es la medida de Lebesgue sobre $\mathbb{R}^n$. Entonces $\mu \ll \boldsymbol{m}^n$ y $\dfrac{d\mu}{d\boldsymbol{m}^n} \le C \quad \boldsymbol{m}^n - c.t.p.$

Demostración. Sea A un conjunto medible tal que $\boldsymbol{m}^n(A) = 0$. Denotemos $A' = A \cap S$. Entonces basta probar que $\mu(A') = 0$. Fijemos $\epsilon > 0$ arbitrario. Ya que $\boldsymbol{m}^n(A') = 0$ entonces $\exists\, G \subseteq \mathbb{R}^n$ abierto tal que

$$A' \subseteq G \quad \text{y} \quad \boldsymbol{m}^n(G) < \epsilon\ .$$

Por hipótesis $\forall\, x \in A'$, $\exists$ una bola B_x de centro x tal que

$$B_x \subseteq G \quad \text{y} \quad \mu(B_x) \leq C\, \boldsymbol{m}^n(B_x)\ .$$

Llamemos $a := \mu\left(\bigcup_{x \in A'} B_x\right)$. Si $a = 0$ entonces no hay nada que probar. Supongamos entonces que $a > 0$. Por el Lema 3.15 de [5] se sigue que $\exists\, x_1, x_2, ..., x_m \in A'$ tal que

$$\mu\left(B_{x_1} \cup \cdots \cup B_{x_m}\right) > \left(\frac{a}{2}\right) \cdot \frac{1}{3^n},$$

donde $B_{x_1}, ..., B_{x_m}$ son bolas disjuntas. Ahora bien,

$$\begin{aligned}
\left(\frac{a}{2}\right) \cdot \frac{1}{3^n} < \sum_{j=1}^{m} \mu(B_{x_j}) \ &\leq\ \sum_{j=1}^{m} C\boldsymbol{m}^n(B_{x_j}) \\
&=\ C\boldsymbol{m}^n\left(B_{x_1} \cup \cdots \cup B_{x_m}\right) \\
&\leq\ C\boldsymbol{m}^n(G) < C\epsilon.
\end{aligned}$$

Eso prueba que

$$\mu(A')\ \leq\ a\ \leq\ 2 \cdot 3^n C\epsilon\ .$$

Como ϵ es arbitrario, concluimos que $\mu(A') = 0$ probando que $\mu \ll \boldsymbol{m}^n$.

Luego, por el Teorema de Radon-Nikodym, existe

$$f\ =\ \frac{d\mu}{d\boldsymbol{m}^n} \in L^1(\mathbb{R}^2, \boldsymbol{m}^n)\ .$$

Finalmente, gracias al Teorema de diferenciación de Lebesgue (Ver Teorema 3.21 de[5]) se sigue que para $\boldsymbol{m}^n$-casi todo x,

$$f(x)\ =\ \lim_{r \to 0} \frac{1}{\boldsymbol{m}^n\left(B_r(x)\right)} \int_{B_r(x)} f\, d\boldsymbol{m}^n\ =\ \lim_{r \to 0} \frac{\mu\left(B_r(x)\right)}{\boldsymbol{m}^n\left(B_r(x)\right)}\ \leq\ C\ ,$$

donde la última desigualdad se sigue gracias a la hipótesis. □

Ahora probaremos la desigualdad de Hoeffding. Para ello, comenzaremos probando el siguiente lema.

Lema 9.6. Sea X una variable aleatoria en el espacio de probabilidad $([a,b], \mathcal{B}_{[a,b]}, \mathbb{P})$. Si X tiene media 0, entonces $\mathbb{E}\, e^{tX} \leq e^{t^2(b-a)^2/8}$.

Demostración. Debido a la convexidad de e^{tx}, tenemos que

$$e^{tx} \leq \lambda e^{ta} + (1-\lambda)e^{tb},$$

donde $x = \lambda a + (1-\lambda)b$, $\lambda \in [0,1]$. Sea $\lambda = (b-x)/(b-a)$ entonces se sigue que

$$e^{tx} \leq \frac{b-x}{b-a}e^{ta} + \frac{x-a}{b-a}e^{tb}.$$

Tomando la esperanza en ambos lados de la desigualdad, obtenemos

$$\begin{aligned} \mathbb{E}\, e^{tX} &\leq \frac{b-\mathbb{E}\, X}{b-a}e^{ta} + \frac{\mathbb{E}\, X - a}{b-a}e^{tb} \\ &= \frac{be^{ta} - ae^{tb}}{b-a} \\ &\leq e^{t^2(b-a)^2/8}. \end{aligned}$$

Esta última desigualdad se demostrará a continuación. Sea $p = -a/(b-a)$, entonces

$$\begin{aligned} \frac{be^{ta} - ae^{tb}}{b-a} &= (1-p+pe^{t(b-a)})e^{-pt(b-a)} \\ &= e^{\phi(u)}, \end{aligned}$$

donde $u = t(b-a)$ y $\phi(u) = -pu + \log(1-p+pe^u)$. Dado que

$$\phi'(u) = -p + \frac{p}{p+(1-p)e^{-u}},$$

por lo tanto $\phi(0) = \phi'(0) = 0$. Más aún $\phi''(u) \leq \frac{1}{4}$. Ahora por el Teorema de Taylor se tiene que para algún $\theta \in [0,u]$

$$\begin{aligned} \phi(u) &= \phi(0) + u\phi'(0) + \frac{u^2}{2}\phi''(\theta) \\ &\leq \frac{u^2}{8} \\ &= \frac{t^2(b-a)^2}{8}. \end{aligned}$$

□

Proposición 9.7. *(Desigualdad de Hoeffding) Sean $X_1, ..., X_n$ variables aleatorias independientes acotadas tal que $a_i \leq X_i \leq b_i$ y sea $S_n = X_1 + \cdots X_n$. Entonces para todo $\epsilon > 0$,*

$$\mathbb{P}\left(S_n - \mathbb{E}\, S_n \geq \epsilon\right) \leq e^{-2\epsilon^2/D} \quad y \quad \mathbb{P}\left(S_n - \mathbb{E}\, S_n \leq -\epsilon\right) \leq e^{-2\epsilon^2/D},$$

donde $D = \sum_{i=1}^n (b_i - a_i)^2$.

Demostración. Solo probaremos la cota inferior (la prueba para la cota superior es similar). Sea $Y_i = X_i - \mathbb{E}\, X_i$, entonces las Y_i son variables aleatorias independientes, con media cero y rango $[a_i - \mathbb{E}\, X_i, b_i - \mathbb{E}\, X_i]$. Para cualquier $t > 0$, se sigue

$$\begin{aligned} \mathbb{P}(S_n - \mathbb{E}\, S_n \geq \epsilon) &= \mathbb{P}(Y_1 + \cdots Y_n \geq \epsilon) \\ &= \mathbb{P}(e^{t(Y_1 + \cdots Y_n)} \geq e^{t\epsilon}) \leq \frac{\mathbb{E}\, e^{t(Y_1 + \cdots + Y_n)}}{e^{t\epsilon}}, \end{aligned}$$

donde en la última desigualdad se ha utilizado la desigualdad de Markov. Luego por la independencia de las Y_i y usando el Lema 9.6 para cada Y_i, obtenemos que

$$\begin{aligned} \mathbb{P}(S_n - \mathbb{E}\, S_n \geq \epsilon) &\leq \frac{\mathbb{E}\, e^{tY_1} \cdots \mathbb{E}\, e^{tY_1}}{e^{t\epsilon}} \\ &\leq \frac{e^{t^2(b_1 - a_1)^2/8} \cdots e^{t^2(b_n - a_n)^2/8}}{e^{t\epsilon}}. \end{aligned}$$

Finalmente, tomando $t = 4\epsilon/D$ concluimos que

$$\mathbb{P}(S_n - \mathbb{E}\, S_n \geq \epsilon) \leq e^{-2\epsilon^2/D}.$$

□

Para terminar este capítulo, daremos la definición de convergencia en distribución de vectores aleatorios y probaremos en la Proposición 9.9 que podemos usar funciones Lipschitz para probar la convergencia.

Definición 9.8. *(Convergencia distribucional) Para $k \in \mathbb{N}$ fijo, decimos que*

$$(X_1, X_2, ..., X_k) \xrightarrow{ley} (Y_1, Y_2, ..., Y_k)$$

si $\forall a_1, ..., a_n \in \mathbb{R}$ tales que $\mathbb{P}(Y_1 = a_1) = \mathbb{P}(Y_2 = a_2) = ... = \mathbb{P}(Y_k = a_k) = 0$ se cumple

$$\mathbb{P}(X_1 \leq a_1, ..., X_k \leq a_k) \to \mathbb{P}(Y_1 \leq a_1, ..., Y_k \leq a_k)$$

La siguiente proposición fue usada para probar la convergencia en distribución enunciada en nuestros teoremas principales.

Proposición 9.9. *Sean $h_1, ..., h_k : \mathbb{R} \to \mathbb{R}$ k funciones lipschitzianas tales que $|h_i(x)| \leq 1$, y $(Y_1, Y_2, ..., Y_k)$ un vector aleatorio, cuyas coordenadas tienen densidad proporcional a $\exp(-rx^2 - sx)$ sobre $[0, \infty[$. Si*

$$\mathbb{E}\left[\prod_{j=1}^{k} h_j(X_j)\right] \to \mathbb{E}\left[\prod_{j=1}^{k} h_j(Y_j)\right],$$

entonces

$$(X_1, X_2, ..., X_k) \xrightarrow{ley} (Y_1, Y_2, ..., Y_k).$$

Demostración. Considerando k funciones $h_1, ..., h_k : \mathbb{R} \to \mathbb{R}$ definidas como:

$$h_j(x) := \begin{cases} 1 & \text{si } x < a_j \\ 1 - \frac{1}{\delta}(x - a_j) & \text{si } a_j \leq x < a_j + \delta \\ 0 & \text{si } x \geq a_j + \delta, \end{cases}$$

para $j = 1, ..., k$ y $\delta > 0$ fijo pero arbitrario. Dado que $1_{\{X_j \leq a_j\}} \leq h_j(X_j)$ para $j = 1, ..., k$ entonces,

$$\mathbb{P}\left(X_1 \leq a_1, ..., X_k \leq a_k\right) \leq \mathbb{E}\left[h_1(X_1)h_2(X_2) \cdots h_k(X_k)\right]. \tag{9.1}$$

Por otro lado,

$$\mathbb{E}\left[h_1(Y_1)h_2(Y_2) \cdots h_k(Y_k)\right] \leq \mathbb{P}\left(Y_1 \leq a_1 + \delta, Y_2 \leq a_2 + \delta, ..., Y_k \leq a_k + \delta\right). \tag{9.2}$$

Como

$$\mathbb{E}\left[h_1(X_1)h_2(X_2) \cdots h_k(X_k)\right] \to \mathbb{E}\left[h_1(Y_1)h_2(Y_2) \cdots h_k(Y_k)\right],$$

se sigue de (9.1) y (9.2) que para $\delta > 0$

$$\limsup_{n\to\infty} \mathbb{P}\left(X_1 \leq a_1, ..., X_k \leq a_k\right) \leq \mathbb{P}\left(Y_1 \leq a_1 + \delta, Y_2 \leq a_2 + \delta, ..., Y_k \leq a_k + \delta\right).$$

Luego haciendo $\delta \downarrow 0$, obtenemos

$$\limsup_{n\to\infty} \mathbb{P}\left(X_1 \leq a_1, ..., X_k \leq a_k\right) \leq \mathbb{P}\left(Y_1 \leq a_1, Y_2 \leq a_2, ..., Y_k \leq a_k\right). \tag{9.3}$$

Por otro lado, considerando k funciones $\tilde{h}_1, ..., \tilde{h}_k : \mathbb{R} \to \mathbb{R}$ definidas como:

$$\tilde{h}_j(x) := \begin{cases} 1 & \text{si } x < a_j - \delta \\ -\frac{1}{\delta}(x - a_j) & \text{si } a_j - \delta \leq x < a_j \\ 0 & \text{si } x \geq a_j, \end{cases}$$

para $j = 1, ..., k$ y $\delta > 0$ fijo pero arbitrario. Dado que $1_{\{X_j \leq a_j\}} \geq \tilde{h}_j(X_j)$ para $j = 1, ..., k$ entonces,

$$\mathbb{P}\left(X_1 \leq a_1, ..., X_k \leq a_k\right) \geq \mathbb{E}\left[\tilde{h}_1(X_1)\tilde{h}_2(X_2)\cdots\tilde{h}_k(X_k)\right]. \tag{9.4}$$

Por otro lado,

$$\mathbb{E}\left[\tilde{h}_1(Y_1)\tilde{h}_2(Y_2)\cdots\tilde{h}_k(Y_k)\right] \geq \mathbb{P}\left(Y_1 \leq a_1 - \delta, Y_2 \leq a_2 - \delta, ..., Y_k \leq a_k - \delta\right). \tag{9.5}$$

Como

$$\mathbb{E}\left[\tilde{h}_1(X_1)\tilde{h}_2(X_2)\cdots\tilde{h}_k(X_k)\right] \to \mathbb{E}\left[\tilde{h}_1(Y_1)\tilde{h}_2(Y_2)\cdots\tilde{h}_k(Y_k)\right],$$

se sigue de (9.4) y (9.5) que para $\delta > 0$

$$\liminf_{n\to\infty} \mathbb{P}\left(X_1 \leq a_1, ..., X_k \leq a_k\right) \geq \mathbb{P}\left(Y_1 \leq a_1 - \delta, Y_2 \leq a_2 - \delta, ..., Y_k \leq a_k - \delta\right).$$

Luego haciendo $\delta \downarrow 0$, obtenemos

$$\liminf_{n\to\infty} \mathbb{P}\left(X_1 \leq a_1, ..., X_k \leq a_k\right) \geq \mathbb{P}\left(Y_1 < a_1, Y_2 < a_2, ..., Y_k < a_k\right). \tag{9.6}$$

Finalmente de (9.3) y (9.6) se concluye que

$$\mathbb{P}(X_1 \leq a_1, ..., X_k \leq a_k) \to \mathbb{P}(Y_1 \leq a_1, ..., Y_k \leq a_k).$$

Lo cual completa la prueba. □

Bibliografía

[1] Athreya, K. B., & Lahiri, S. N. (2006). Measure theory and probability theory. Springer Science & Business Media.

[2] Bhattacharya, R. N., & Rao, R. R (1986). Normal Approximation and Asymptotic Expansion.

[3] Chatterjee, S. (2016). A note about the uniform distribution on the intersection of a simplex and a sphere. arXiv preprint arXiv:1011.4043.

[4] Diaconis, P. and Friedman, D. (1987). A dozen de Finetti-style results in search of a theory. *Ann. Inst. H. Poincaré Probab. Statist.* **23** no. 2, suppl., 397–423.

[5] Folland, G. B. (2013). Real analysis: modern techniques and their applications. John Wiley & Sons.

[6] Grimmett, G., & Stirzaker, D. (2001). Probability and random processes. Oxford university press.

[7] Rudin, W. (1987). Real and complex analysis. Tata McGraw-Hill Education.

[8] Rumpf, B. (2004). Simple statistical explanation for the localization of energy in nonlinear lattices with two conserve quantities. *Phys. Rev.* E **69**, 016618.

[9] Shakarchi, R., & Stein, E. M. (2009). Real analysis: measure theory, integration, and Hilbert spaces. Princeton University Press.