
**UNA k -DEFORMACIÓN PARA LA VARIEDAD DE INFORMACIÓN
ESTADÍSTICA**

JUAN CARLOS ARANGO PARRA

ESCUELA DE CIENCIAS Y HUMANIDADES
DEPARTAMENTO DE CIENCIAS BÁSICAS
MAESTRÍA EN MATEMÁTICAS APLICADAS
MEDELLÍN
2012

Una k -deformación para la variedad de información estadística



Trabajo de investigación presentado para optar al título de Magister en
Matemáticas Aplicadas

JUAN CARLOS ARANGO PARRA

Asesor
Gabriel Ignacio Loaiza Ossa
Doctor en Ciencias Matemáticas

Escuela de Ciencias y Humanidades
Departamento de Ciencias Básicas
Maestría en Matemáticas Aplicadas
Medellín
2012

Nota de aceptación

Coordinador de la Maestría

Director del proyecto

Ciudad y fecha (día, mes, año):

Agradecimientos

A mi familia que siempre me apoyada en mis estudios y con sabiduría soportan mis errores y displicencia; y de manera silenciosa han permitido mi avance académico y personal.

A cada uno de esos grandes profesores y maestros que me han conducido por este camino de la disciplina, la responsabilidad, el respeto y la superación constante. Especialmente a los profesores Hernán Monsalve que para mí ha sido ese “Papá matemático” que me facilitado el acceso a este hermoso e intricado mundo del formalismo matemático. A Gabriel Ignacio Loaiza por sus asesorías, buen humor, apoyo y su vasto conocimiento del mundo y en estos últimos días de la gramática. Un sentido respeto a ese gran matemático y persona que es **Gerardo Arango**, paz en su tumba; estarás vivo en mi corazón y mis pensamientos.

También un agradecimiento a Hector Román Quiceno por sus asesorías, sugerencias, sus ideas, proyectos, expectativas y vivencias. De nuevo a esos amigos bohemios que tras charlas, copas, obstáculos y proyectos de vida me han enriquecido desde lo académico, profesional y sobre todo lo personal. A esos tres amores que han compartido conmigo este abstracto mundo de las matemáticas, que han sabido comprender el poco tiempo de dedicación por resolver algunos “problemitas matemáticos”.

Resumen

A partir de una medida de probabilidad dada μ , sobre el espacio \mathfrak{M}_μ de densidades estrictamente positivas, se construye una variedad topológica a través de modelos exponenciales k -deformados que conectan a densidades $p, q \in \mathfrak{M}_\mu$ mediante $q = e_k^{u \ominus K_{p,k}(u)} p$, donde: $0 < |k| < 1$, \ominus es la k -diferencia según G. Kaniadakis y $K_{p,k}$ es una k -deformación del mapeo acumulante definido para una variable aleatoria u por $K_p(u) = \ln(E_p[e^u])$ (siendo E_p el valor esperado respecto a la medida con densidad p). El sentido de las k -deformaciones implica que cuando $k \rightarrow 0$, se recuperan tanto los modelos exponenciales como la variedad de información de Pistone y Sempi presentada en 1995. Los espacios modeladores son espacios de funciones de Orlicz $L^{\phi_k}(p \cdot \mu)$, donde $\phi_k(\cdot) = \cosh_k(\cdot) - 1$, los cuales son espacios de Banach infinito-dimensionales; a pesar que la variedad construida sea topológica, se estudia la analiticidad del mapeo $K_{p,k}$ usando teoría de convergencia de series de potencias entre espacios de Banach. La variedad construida es diferente a la presentada por Pistone (*K-exponential models from the geometrical viewpoint* en 2009), que se basa en k -divergencias $\left(D_k(p \parallel q) = E_p \left[\ln_k \left(\frac{p}{q} \right) \right]\right)$ y modelos k -exponenciales de la forma $q = e_k^{u - D_k(p \parallel q)} p$, donde $u \in L_0^{1/k}(p \cdot \mu)$. En el capítulo 3 se amplían las demostraciones de la variedad estándar de Pistone y Sempi.

Índice general

1. Propuesta	1
1.1. Introducción	1
1.2. Objetivos	3
1.2.1. Objetivo General	3
1.2.2. Objetivos Específicos	3
2. Preliminares	5
2.1. Aspectos sobre teoría de la medida y probabilidad	5
2.2. Espacios de Funciones de Orlicz	9
2.3. Analiticidad entre espacios de Banach	20
2.4. k-Deformaciones	23
3. Variedad de Información	27
3.1. Introducción	27
3.2. Generalidades sobre el espacio $L^{\phi_1}(p \cdot \mu)$	27
3.3. Funcional Acumulante	30
3.4. Variedad de Información	41
3.5. Entropía Relativa	52
3.6. B_p como un espacio de Hilbert	54
3.7. Gradiente del funcional Acumulante	64
3.8. Espacio Tangente	68
4. Modelos k-exponenciales según G. Pistone	71
4.1. Introducción	71

4.2. Cartas	71
4.3. Espacio Tangente	81
5. Variedad de información k-deformada	85
5.1. Generalidades sobre el espacio $L^{\phi_k}(p \cdot \mu)$	85
5.2. Analiticidad	87
5.3. Funcional acumulante k -deformado	98
5.4. Variedad de información k -deformada	104
6. Problemas Abiertos	115
7. Conclusiones	117
Bibliografía	119

Capítulo 1

Propuesta

1.1. Introducción

En análisis funcional, la teoría de espacios de funciones de Orlicz [27, 31] es una generalización natural de los espacios de Lebesgue, que ha tenido diversas e importantes aplicaciones [26]. Particularmente, los espacios de funciones de Orlicz junto a la geometría diferencial, han fundamentado aplicaciones relacionadas con la estadística y la teoría de la información. El desarrollo comenzó en 1945 con C.R. Rao [25] y H. Jeffeys [11], quienes interpretaron modelos de teoría de la información como modelos estadísticos parametrizados, usaron la geometría Riemanianna. Para las familias de funciones de densidad de probabilidad parametrizadas se establece una estructura de variedad y desde 1985 algunas importantes aplicaciones han sido estudiadas, por ejemplo en los trabajos de Amari [1, 2].

En el artículo publicado en 1995 y titulado *An infinite-dimensional geometric structure on the space of all the probability measures equivalent to a given one* [22], G. Pistone y C. Sempi introducen una variedad de Banach (modelada sobre espacios de funciones de Orlicz) sobre las funciones de densidad de probabilidad no parametrizadas, la cual permite interpretar la entropía clásica relativa a dos funciones de densidad de probabilidad que estén conectadas por un modelo exponencial. La variedad construida es conocida como variedad de información estadística de Pistone y Sempi, y trabajos posteriores [29, 12] han permitido avances en el estudio de los estados físicos de sistemas clásicos y cuánticos. En teoría estadística clásica, el trabajo de Pistone y Sempi desarrolló una teoría de los mejores estimadores (de mínima varianza) entre todos los estimadores localmente insesgados para la estimación no paramétrica.

El estudio y aplicación de la variedad de información estadística se han realizado en diversos trabajos (por ejemplo [21, 5]) y además, el mismo G. Pistone, en el artículo titulado *k-exponential models from the geometrical viewpoint* [23], presentó en el año 2009 una nueva variedad de Banach que contiene, como límite, a la variedad construida por él

en 1995. La variedad introducida en dicho trabajo depende de un parámetro real k (con $0 < |k| < 1$) correspondiente al parámetro k de la mecánica estadística según el formalismo de G. Kaniadakis [14] y se basa en las k -divergencias dadas para densidades positivas p y q por $D_k(p \parallel q) = E_p \left[\ln_k \left(\frac{p}{q} \right) \right]$ ver [23, 14] y en modelos k -exponenciales de la forma $q = e_k^{u - D_k(p \parallel q)} p$ (donde $u \in L_0^{1/k}(p \cdot \mu)$). Esta variedad recupera a la variedad de Pistone y Sempi cuando k tiende a cero y se construye con metodología diferente a la empleada en el artículo de Pistone y Sempi [22], donde las divergencias no son la base de la construcción.

En el trabajo del año 2009 [23], G. Pistone propone, además, una discusión sobre posibles estructuras que permitan interpretaciones geométricas de aspectos relacionados con la información en el contexto del formalismo de la mecánica estadística según G. Kaniadakis; que se trata de una generalización de la mecánica estadística de Boltzmann Gibbs [14, 15]. Una estructura geométrica allí propuesta, que aún no aparece desarrollada en la literatura, es una k -deformación de la variedad de información estadística, mediante otra k -deformación de los modelos exponenciales. Es ese caso, se proponen modelos que conectan a densidades positivas p y q mediante $q = e_k^{u \ominus \Psi(u)} p$, donde: \ominus es la k -diferencia según G. Kaniadakis y $\Psi(u)$ es alguna función de la variable aleatoria u .

En este trabajo se desarrolla dicha propuesta de G. Pistone y se establece que $\Psi(u)$ es una k -deformación del mapeo acumulante definido para una variable aleatoria u por $K_p(u) = \ln(E_p[e^u])$, siendo E_p el valor esperado respecto a la medida con densidad p ; precisamente, en el último capítulo se verá que $\Psi(u) := K_{p,k}(u)$.

Además de ampliar la discusión sobre k -deformaciones para la variedad de Pistone y Sempi, esta memoria de tesis también presenta la divulgación de aspectos importantes sobre variedades de información; ampliando los detalles de pruebas que aparecen en artículos y presentando las pruebas que son omitidas. Por ello, el capítulo 2 presenta preliminares necesarios sobre elementos de teoría de medida, espacios de funciones de Orlicz, analiticidad entre espacios de Banach y k -Deformaciones. En el capítulo 3 se presenta la construcción de la variedad de información de Pistone y Sempi [22, 21, 6, 5] y el capítulo 4 presenta la variedad k -deformada construida por Pistone [23]. En el capítulo 5 se presenta el desarrollo de una nueva forma de construir una k -deformación para la variedad de Pistone y Sempi, acorde con la conjetura hecha en [23] y probando en detalle los nuevos resultados. Finalmente, se presentan secciones de conclusiones y problemas abiertos.

1.2. Objetivos

1.2.1. Objetivo General

A partir del parámetro k según la mecánica estadística de G. Kaniadakis y un tipo de k -deformación para los modelos exponenciales, construir una variedad topológica (k -deformación de la variedad de información de Pistone y Sempi) sobre las densidades de probabilidad estrictamente positivas y, además, divulgar algunos aspectos relevantes de las variedades de información estadística.

1.2.2. Objetivos Específicos

1. Mostrar detalles relevantes de la variedad de la información estadística de Pistone y Sempi.
2. Presentar el estudio como espacio de Hilbert del conjunto de variables aleatorias centradas respecto de la norma inducida por el producto interno.
3. Dar una interpretación geométrica en el espacio de variables aleatorias centradas, de conceptos estadísticos como son: la varianza, la covarianza, la desviación típica y el coeficiente de correlación.
4. Presentar la construcción de la variedad k -deformada según Pistone, a partir de las k -divergencias y un tipo de k -deformación de los modelos exponenciales.
5. Introducir y presentar el estudio de otra k -deformación de modelos exponenciales, que permita construir una k -deformación de la variedad de Pistone y Sempi diferente a la ya construida por Pistone y acorde con su propuesta.

Capítulo 2

Preliminares

En este capítulo se presentan conceptos necesarios para el desarrollo del trabajo. Se presentan en primera instancia conceptos y propiedades relativos a la teoría de la medida y aspectos de la teoría de la probabilidad, cuyo sustento teórico se encuentra en Royden y Liliana Blanco [28, 3]. Se continúa con aspectos de la teoría de espacios de funciones Orlicz, según los libros de Rao y Ren [26, 27] y de la tesis doctoral de Pablo Gregori [10]; en este caso se presentan algunas pruebas para contextualizar mejor la teoría. Posteriormente, se continúa con aspectos de analiticidad entre espacios de Banach, desde la teoría de convergencia de series en espacios de Banach, tomados del libro de Kadets and Kadets [13]. Por último, se presenta la teoría de k -deformaciones según Kaniadakis en [14] y las tesis de maestría de Ruben Borja y de Dora Deossa, respectivamente [4, 7].

2.1. Aspectos sobre teoría de la medida y probabilidad

Con la idea de establecer notación y recordar algunos resultados, que se usarán en pruebas o resaltan la validez de algunas definiciones, en esta sección se presentan algunos aspectos sobre teoría de la medida y probabilidad, tomando como referencia los libros *Real Analysis* y *Teoría avanzada de probabilidad* [28, 3]. Para ello se tendrá presente que en este trabajo se hace referencia a variables aleatorias realvaluadas, esto es, funciones \mathcal{A} -Borel medibles de Ω en \mathbb{R} , donde $(\Omega, \mathcal{A}, \mu)$ es un espacio de probabilidad. De ahora en adelante se dirá función medible en lugar de \mathcal{A} -Borel medible.

En primer lugar, se recuerda que dado un espacio de medida $(\Omega, \mathcal{A}, \mu)$ y una función \mathcal{A} -Borel medible $X : \Omega \rightarrow \mathbb{R}$, se obtiene otra medida ν en (Ω, \mathcal{A}) llamada medida con densidad X , que es dada para cada $E \in \mathcal{A}$ por $\nu(E) := \int_E X d\mu$; en tal caso se denota $\nu := X \cdot \mu$, para indicar que la nueva medida depende de la anterior y de X . Bajo una versión particular de lo anterior, se puede escribir la siguiente definición; que dá sentido a la definición de la medida y notación empleadas para la definición de los espacios modeladores de las variedades en los próximos capítulos.

Definición 2.1. Sean $(\Omega, \mathcal{A}, \mu)$ un espacio de probabilidad y p una función de densidad de probabilidad. Se denota por $p \cdot \mu$, a la medida de probabilidad con densidad p , que es dada para cada $A \in \mathcal{A}$ por

$$(p \cdot \mu)(A) = \int_A p d\mu .$$

Un resultado que permite caracterizar las funciones medibles bajo la medida $p \cdot \mu$ en términos de la medida μ es el siguiente: una función f es $p \cdot \mu$ medible, si y sólo si, fp es μ medible, en cuyo caso se presenta la igualdad

$$\int_A f d(p \cdot \mu) = \int_A fp d\mu . \quad (2.1)$$

Una función f es convexa si, para cada $x, y \in \mathbb{R}$ y λ tal que $0 \leq \lambda \leq 1$, se tiene

$$f((1 - \lambda)x + \lambda y) \leq (1 - \lambda)f(x) + \lambda f(y) .$$

La convexidad de una función medible permite definir propiedades importantes. Sean, $(\Omega, \mathcal{A}, \mu)$ un espacio de probabilidad, una función f convexa sobre \mathbb{R} y ϕ una función μ -integrable entonces, se presenta la conocida desigualdad de Jensen:

$$f \left[\int_{\Omega} \phi(t) dt \right] \leq \int_{\Omega} f(\phi(t)) dt . \quad (2.2)$$

A manera de ejemplo, se puede resaltar que en este trabajo se usará en varias ocasiones la desigualdad $e^{\int_{\mathbb{R}} \phi(u) d\mu} \leq \int_{\mathbb{R}} e^{\phi(u)} d\mu$, donde ϕ es una función μ -integrable y claramente la función exponencial es convexa en todos los reales.

La antesala del teorema de la convergencia monótona es el lema de Fatou, que reviste una gran importancia en el estudio de la integral de Lebesgue porque permite establecer condiciones para intercambiar el orden del límite de una sucesión de funciones medibles y la integral de Lebesgue. El lema de Fatou se usa en algunas pruebas más adelante y se enuncia como sigue. Sea $\{f_n\}$ una sucesión de funciones medibles no-negativas tal que $f_n(x) \rightarrow f(x)$ casi en todas partes de un conjunto A entonces,

$$\int_A f d\mu \leq \liminf_{n \rightarrow \infty} \int_A f_n d\mu = \underline{\lim}_{n \rightarrow \infty} \int_A f_n d\mu .$$

El teorema de la convergencia monótona establece que si $\{f_n\}$ es una sucesión creciente de funciones medibles no-negativas tal que $f = \lim_{n \rightarrow \infty} f_n$ entonces

$$\int_A f d\mu = \int_A \left(\lim_{n \rightarrow \infty} f_n \right) d\mu = \lim_{n \rightarrow \infty} \int_A f_n d\mu .$$

De tal forma, si la sucesión creciente f_n converge a f casi en todas partes, entonces se permite el intercambio entre el límite y la integral; situación similar sucede entre la serie y la integral.

Considérese ahora $\{f_n\}$, una sucesión de funciones medibles. Se dice que $\{f_n\}$ converge a f en medida si: dado $\epsilon > 0$, existe $N \in \mathbb{N}$, tal que para todo $n \geq N$ se tiene que

$$\mu(\{x : |f(x) - f_n(x)| \geq \epsilon\}) < \epsilon.$$

Tanto el lema de Fatou como el teorema de la convergencia monótona son válidos si la convergencia es en medida.

Dado un espacio de probabilidad $(\Omega, \mathcal{A}, \mu)$, la *esperanza matemática* o *valor esperado* de una variable aleatoria $u : \Omega \rightarrow \mathbb{R}$, es el número real dado por $E[u] = \int_{\Omega} u d\mu$, siempre que la integral sea finita. Si p es una función de densidad de probabilidad entonces, en términos de la medida de probabilidad $p \cdot \mu$, la esperanza se denotará por

$$E_p[u] := \int_{\Omega} u d(p \cdot \mu) = \int_{\Omega} u p d\mu.$$

A continuación se ilustran algunas de las propiedades del valor esperado.

1. Si u, v son variables aleatorias cuya esperanza existe y $\alpha, \beta \in \mathbb{R}$ entonces, $E_p[\alpha u + \beta v] = \alpha E_p[u] + \beta E_p[v]$.
2. Si las variables aleatorias u y v son tales que $u \leq v$ entonces, $E_p[u] \leq E_p[v]$. En particular se tiene para cualquier variable aleatoria que $|E_p[u]| \leq E_p[|u|]$.

Asociado a una variable aleatoria u , se encuentra el n -ésimo momento de u alrededor del cero, el cual se denota como $\mu'_n = E_p[u^n]$ siempre y cuando dicho valor esperado existe. En este sentido aparece el momento central alrededor de la media $E_p[u]$, dado por $\mu_n = E_p[(u - E_p[u])^n]$. El segundo momento central alrededor del cero recibe el nombre de *varianza* de la variable aleatoria u y está expresada como $var_p[u] = E_p[(u - E_p[u])^2]$. La desviación típica denotada $\sigma_p(u)$ es la raíz cuadrada positiva de la varianza, esto es, $\sigma_p(u) = \sqrt{var_p[u]}$. Algunas propiedades de la varianza son

1. Para una variable aleatoria u se satisface $var_p[u] = E_p[u^2] - (E_p[u])^2$.
2. Si u es una variable aleatoria cuya esperanza existe entonces, $var_p[u] \geq 0$; la igualdad ocurre en el caso en que u sea constante.
3. Para una constante α , se verifica $var_p[\alpha u] = \alpha^2 var_p[u]$.
4. Si β es una constante entonces, las variables aleatorias u y $u + \beta$ tienen la misma varianza, $var_p[u + \beta] = var_p[u]$.

En general, la varianza es una medida de la dispersión respecto de la media o esperanza. Sea u variable aleatoria; en caso que $E_p[e^{ut}] < \infty$ para t en algún intervalo $I = (-\alpha, \alpha)$ entonces, la función $\hat{u}_p(t) = E_p[e^{ut}]$ recibe el nombre de *función generadora de momentos*. Si para una variable aleatoria u , la función generadora de momentos existe entonces, se garantiza la existencia de todos los momentos centrales alrededor del cero, es decir, $E_p[u^n]$ existe para todo $n \in \mathbb{N}$.

Sean u y v dos variables aleatorias para las cuales los segundos momentos centrales alrededor del cero existan $E_p[u^2] < \infty$ y $E_p[v^2] < \infty$. La *covarianza* entre ambas variables aleatorias está dada por $cov_p[u, v] = E_p[(u - E_p[u])(v - E_p[v])]$. Lo anterior se puede presentar como $cov_p[u, v] = E_p[uv] - E_p[u]E_p[v]$. Con base en ambas igualdades se tienen las propiedades:

1. La covarianza es simétrica, $cov_p[u, v] = cov_p[v, u]$.
2. La covarianza de una variable con sí misma es la varianza de dicha variable $cov_p[u, u] = var_p[u]$.
3. Para a, b reales se verifica que $cov_p[au + b, v] = a cov_p[u, v]$.
4. Si la covarianza existe entre dos variables u y v entonces, se cumple la desigualdad de Cauchy-Schwarz: $|E_p[uv]|^2 \leq E_p[u^2]E_p[v^2]$.
5. Para variables aleatorias u y v , resulta la igualdad $var_p[u + v] = var_p[u] + var_p[v] + 2cov_p[u, v]$.

Ahora, el coeficiente de correlación entre las variables aleatorias u, v es el número

$$\rho_p(u, v) = \frac{cov_p[u, v]}{\sqrt{var_p[u]}\sqrt{var_p[v]}}$$

siempre que la varianza de una y otra variable aleatoria sean números reales positivos. Finalmente, el coeficiente de correlación satisface lo siguiente.

1. El coeficiente de correlación es simétrico $\rho_p(u, v) = \rho_p(v, u)$.
2. $|\rho_p(u, v)| \leq 1$.
3. $\rho_p(u, u) = 1$ y $\rho_p(u, -u) = -1$.
4. Para a, b reales, con $a > 0$, se satisface $\rho_p(au + b, v) = \rho_p(u, v)$.

Finalmente, es pertinente aclarar la notación respecto a espacios de Lebesgue. Considerando un espacio de probabilidad $(\Omega, \mathcal{A}, \mu)$, una función de densidad de probabilidad p y un número real $a > 1$, los espacios definidos a continuación (espacios de Lebesgue) se presentan respecto a la medida de probabilidad $p \cdot \mu$.

Como es usual, se denotará por $L^a(p \cdot \mu)$ al espacio de Lebesgue constituido por todas aquellas funciones medibles f sobre Ω (identificando las que son iguales casi en todas partes de Ω) para las cuales

$$E_p[|f|^a] = \int_{\Omega} |f(t)|^a p \, d\mu < \infty ,$$

dotado de la norma

$$\| \cdot \|_{L^a} := \left(\int_{\Omega} | \cdot |^a p \, d\mu \right)^{\frac{1}{a}} = (E_p[| \cdot |^a])^{\frac{1}{a}} .$$

Para $a = 1$, el espacio de Lebesgue $L^1(p \cdot \mu)$ está conformado por aquellas funciones para las cuales el valor absoluto es integrable, dotado la norma dada por

$$\| \cdot \|_{L^1} := \left(\int_{\Omega} | \cdot | p \, d\mu \right) = E_p[| \cdot |] ,$$

que también es un espacio de Banach.

Varias propiedades importantes de los espacios de Lebesgue se pueden escribir en términos del valor esperado, por ejemplo, la desigualdad de Hölder se puede enunciar como sigue. Sean r y s exponentes conjugados, es decir, $\frac{1}{r} + \frac{1}{s} = 1$. Si $f \in L^r(p \cdot \mu)$ y $g \in L^s(p \cdot \mu)$ entonces, $fg \in L^1(p \cdot \mu)$ y se presenta la desigualdad

$$E_p[|fg|] \leq (E_p[|f|^r])^{\frac{1}{r}} (E_p[|g|^s])^{\frac{1}{s}} .$$

2.2. Espacios de Funciones de Orlicz

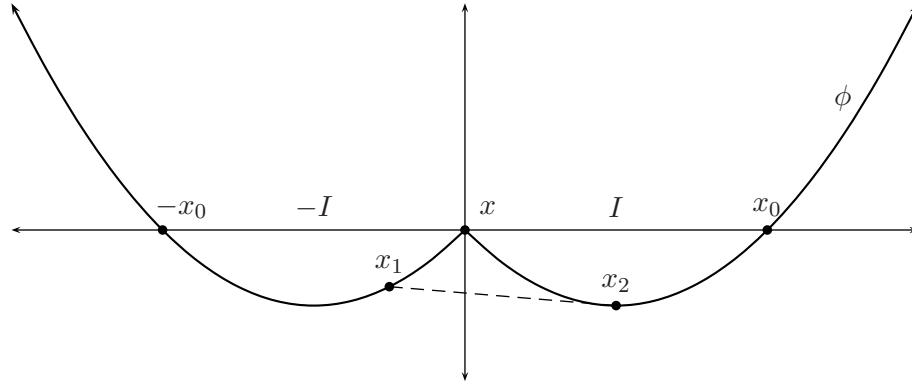
Los espacios de funciones de Orlicz son generalizaciones naturales de los espacios de Lebesgue. Esta sección presenta algunos aspectos básicos sobre la teoría de espacios de funciones de Orlicz, tomados del libro *Theory of Orlicz Spaces* [27], que serán útiles para la presentación de este trabajo. Se expondrá una versión para una medida finita μ , una densidad de probabilidad p y correspondiente a la medida $p \cdot \mu$; tomando $(\Omega, \mathcal{A}, p \cdot \mu)$ como espacio de probabilidad. También es oportuno citar el libro *Application of Orlicz spaces* [26] y de la tesis doctoral de Pablo Gregori [10].

Definición 2.2. Una función de Young $\phi : \mathbb{R} \rightarrow \overline{\mathbb{R}}$ es una función convexa tal que $\phi(0) = 0$, $\phi(x) = \phi(-x)$ y $\lim_{x \rightarrow \infty} \phi(x) = \infty$.

Como consecuencia de la definición de una función de Young se tiene que

$$\phi(x) \geq 0 \quad \forall x \in \mathbb{R} . \tag{2.3}$$

Para ver esto, razonando por método indirecto, supóngase que existe un intervalo $I = [0, x_0]$ donde $\phi(x) < 0$, para todo $x \in I$. Por ser la función par, se tiene que $\phi(x) < 0$ para $x \in -I = [-x_0, 0]$. De lo anterior se sigue un absurdo, precisamente que ϕ no puede ser función convexa; ya que tomando $x_1 \in -I$ y $x_2 \in I$ entonces, el segmento de línea que contiene a x_1 y x_2 es tal que existe un punto $x \in [x_1, x_2]$ para el cual $\phi(x) \geq (1 - \lambda)\phi(x_1) + \lambda\phi(x_2)$, con $0 \leq \lambda \leq 1$. A continuación se presenta un gráfico que ilustra la situación



Por un razonamiento análogo al expuesto anteriormente, se tiene que una función de Young ϕ es creciente en el intervalo $[0, \infty)$ y por ser par se tiene que ϕ es decreciente para $(-\infty, 0]$; lo cual implica el siguiente lema.

Lema 2.1. Sean α y β reales positivos con $\alpha \leq \beta$. Si ϕ es una función de Young entonces $\phi(\alpha x) \leq \phi(\beta x)$ para todo $x \in \mathbb{R}$.

Demostración

Debido a que x es un número real entonces, por tricotomía se presenta una y solo una de las posibilidades: $x = 0$, $x < 0$ o $x > 0$. Analicemos cada uno de los casos.

- i. Si $x = 0$ entonces, $\phi(\alpha x) = \phi(\beta x) = 0$ por ser $\phi(0) = 0$.
- ii. Si $x > 0$ entonces, $\alpha x \leq \beta x$ debido a que $\alpha \leq \beta$. Como la función de Young es creciente en el intervalo $[0, \infty)$ entonces, $\phi(\alpha x) \leq \phi(\beta x)$ y se concluye este caso.
- iii. En el caso $x < 0$, se sigue que $\alpha x \geq \beta x$ donde ambos números αx y βx son negativos. Al ser ϕ decreciente en el intervalo $(-\infty, 0]$ entonces, $\phi(\alpha x) \leq \phi(\beta x)$, lo cual concluye la demostración. \square

Las funciones de Young se pueden escribir en términos de integrales como

$$\phi(x) = \int_0^{|x|} \varphi(t) dt, \quad (2.4)$$

donde φ es una función no-negativa, con dominio $\mathbb{R}^+ \cup \{0\}$, no-decreciente, continua, la cual satisface las condiciones $\varphi(0) = 0$ y $\lim_{x \rightarrow \infty} \varphi(x) = \infty$. Se define ahora la función complementaria de una función de Young.

Definición 2.3. *Sea ϕ una función de Young. La función ϕ^* que es llamada complementaria de ϕ , es la función definida para cada, número real y , por*

$$\phi^*(y) = \sup\{xy - \phi(x) : x \in \mathbb{R}\}.$$

En tal caso se cumple

$$xy \leq \phi(x) + \phi^*(y). \quad (2.5)$$

En efecto, de acuerdo con la definición de supremo se tiene que $\phi^*(y) \geq xy - \phi(x)$ para todo $x \in \mathbb{R}$, si fijamos x_0 en los reales entonces, $\phi^*(y) \geq x_0y - \phi(x_0)$; de donde $\phi^*(y) + \phi(x_0) \geq x_0y$. Como x_0 es arbitrario se sigue $\phi^*(y) + \phi(x) \geq xy$ para todo x e y en los reales. La desigualdad (2.5) es conocida como la desigualdad de Young.

En términos de integración, si ϕ es una función de Young que se puede escribir como $\phi(x) = \int_0^{|x|} \varphi(t)dt$ entonces, la función complementaria es dada por

$$\phi^*(y) = \int_0^{|y|} \theta(t)dt, \quad (2.6)$$

donde $\theta(t)$ es la función inversa de $\varphi(t)$.

Teorema 2.1. *Sea ϕ una función de Young entonces, la función complementaria ϕ^* es una función de Young.*

Demostración

Para la función ϕ^* se probarán las tres condiciones definidas en (2.2), esto es, $\phi^*(0) = 0$, $\phi^*(-y) = \phi^*(y)$ y $\lim_{y \rightarrow \infty} \phi^*(y) = \infty$. Por la definición de la función complementaria (2.3) se tiene

$$\phi^*(0) = \sup\{x \cdot 0 - \phi(x) : x \in \mathbb{R}\} = \sup\{-\phi(x) : x \in \mathbb{R}\}.$$

Puesto que ϕ es una función de Young entonces $\phi(x) \geq 0$ para todo $x \in \mathbb{R}$, así $-\phi(x) \leq 0$ y por tanto $\sup\{-\phi(x) : x \in \mathbb{R}\} = 0$. En consecuencia $\phi^*(0) = 0$.

Se demuestra ahora que la función complementaria es par. Con base en la definición para ϕ^* se escribe

$$\phi^*(-y) = \sup\{x(-y) - \phi(x) : x \in \mathbb{R}\} = \sup\{(-x)y - \phi(x) : x \in \mathbb{R}\}.$$

Como ϕ es par entonces $\phi^*(-y) = \sup\{(-x)y - \phi(-x) : x \in \mathbb{R}\}$, al hacer el cambio de variable $z = -x$ resulta $\phi^*(-y) = \sup\{zy - \phi(z) : z \in \mathbb{R}\}$, y por tanto $\phi^*(-y) = \phi^*(y)$.

Con base en la desigualdad de Young resulta la desigualdad

$$\lim_{y \rightarrow \infty} xy \leq \lim_{y \rightarrow \infty} (\phi(x) + \phi^*(y)) = \phi(x) + \lim_{y \rightarrow \infty} \phi^*(y).$$

Para x fijo, $\lim_{y \rightarrow \infty} xy = \infty$, entonces por ser las funciones de Young definidas en los reales extendidos se tiene $\infty \leq \lim_{y \rightarrow \infty} \phi^*(y) \leq \infty$, es decir, $\lim_{y \rightarrow \infty} \phi^*(y) = \infty$. Se concluye con esto que si ϕ es una función de Young entonces su función complementaria también lo es. \square

Teorema 2.2. *Sea ϕ una función de Young, se verifica que $\phi^{**} = \phi$.*

La función de Young ϕ induce un funcional denotado por i_ϕ (ver [10]) y definido para cada $u \in L^1(p \cdot \mu)$ como

$$i_\phi(u) = \int_{\Omega} \phi(u) p d\mu. \quad (2.7)$$

Dicho funcional no es lineal y conlleva a las definiciones de Clase de Orlicz y espacio de funciones de Orlicz que se precisan a continuación.

Definición 2.4. *Sea ϕ una función de Young. La clase de Orlicz asociada a la función ϕ es el conjunto definido por*

$$\mathcal{L}^\phi(p \cdot \mu) := \{u \in L^1(p \cdot \mu) : i_\phi(u) < \infty\} = \left\{ u \in L^1(p \cdot \mu) : \int_{\Omega} \phi(u) p d\mu < \infty \right\}. \quad (2.8)$$

Claramente $\mathcal{L}^\phi(p \cdot \mu) \subset L^1(p \cdot \mu)$. La clase de Orlicz no es en sí mismo un espacio vectorial real (ver la tesis doctoral de Gregori [10] página 98), sin embargo es un conjunto convexo como se indica en el siguiente teorema.

Teorema 2.3. *Sea ϕ una función de Young entonces $\mathcal{L}^\phi(p \cdot \mu)$ es un conjunto convexo.*

Demostración

Sean u y v en $\mathcal{L}^\phi(p \cdot \mu)$, se sigue de acuerdo con la definición (2.4) que $i_\phi(u) < \infty$ y $i_\phi(v) < \infty$. Sea $0 \leq \lambda \leq 1$; puesto que ϕ es una función convexa entonces $\phi((1-\lambda)u + \lambda v) \leq (1-\lambda)\phi(u) + \lambda\phi(v)$. Aplicando el operador integral a ambos lados y algunas propiedades de este operador resulta,

$$\int_{\Omega} \phi((1-\lambda)u + \lambda v) p d\mu \leq \int_{\Omega} [(1-\lambda)\phi(u) + \lambda\phi(v)] p d\mu$$

$$\leq (1 - \lambda) \int_{\Omega} \phi(u) p d\mu + \lambda \int_{\Omega} \phi(v) p d\mu .$$

En términos del operador i_{ϕ} se escribe $i_{\phi}((1 - \lambda)u + \lambda v) \leq (1 - \lambda)i_{\phi}(u) + \lambda i_{\phi}(v)$, como $i_{\phi}(u)$ y $i_{\phi}(v)$ son finitas, entonces $i_{\phi}((1 - \lambda)u + \lambda v) < \infty$. Así $(1 - \lambda)u + \lambda v$ es un elemento de $L^{\phi}(p \cdot \mu)$, lo cual permite concluir la convexidad de dicho conjunto. \square

Definición 2.5. *Sea ϕ una función de Young. El espacio de funciones de Orlicz asociado a ϕ es el conjunto*

$$\begin{aligned} L^{\phi}(p \cdot \mu) &:= \{u \in L^1(p \cdot \mu) : i_{\phi}(\alpha u) < \infty \text{ para alguna } \alpha > 0\} \\ &= \left\{ u \in L^1(p \cdot \mu) : \int_{\Omega} \phi(\alpha u) p d\mu < \infty \text{ para alguna } \alpha > 0 \right\} . \end{aligned}$$

En el siguiente teorema se demuestra que el espacio de funciones de Orlicz es convexo, al igual que la clase de Orlicz y además posee la estructura de espacio lineal.

Teorema 2.4. *Sea ϕ una función de Young, entonces $L^{\phi}(p \cdot \mu)$ es un espacio lineal convexo.*

Demostración

Sean u y v en $L^{\phi}(p \cdot \mu)$, existen escalares α_1 y α_2 positivos tales que $i_{\phi}(\alpha_1 u) < \infty$ y $i_{\phi}(\alpha_2 v) < \infty$. Sea λ un escalar tal que $0 \leq \lambda \leq 1$ y tomemos $\alpha = \min\{\alpha_1, \alpha_2\}$ en cuyo caso $\alpha \leq \alpha_1$ y $\alpha \leq \alpha_2$, aplicando el lema (2.1) se sigue $\phi(\alpha u) \leq \phi(\alpha_1 u)$ y $\phi(\alpha v) \leq \phi(\alpha_2 v)$. Debido a que ϕ es una función convexa, se tiene

$$\begin{aligned} \phi(\alpha((1 - \lambda)u + \lambda v)) &= \phi((1 - \lambda)(\alpha u) + \lambda(\alpha v)) \\ &\leq (1 - \lambda)\phi(\alpha u) + \lambda\phi(\alpha v) \\ &\leq (1 - \lambda)\phi(\alpha_1 u) + \lambda\phi(\alpha_2 v) . \end{aligned}$$

Al aplicar el operador integral a ambos lados y hacer uso del funcional i_{ϕ} , resulta

$$i_{\phi}(\alpha((1 - \lambda)u + \lambda v)) \leq (1 - \lambda)i_{\phi}(\alpha_1 u) + \lambda i_{\phi}(\alpha_2 v) .$$

Como $i_{\phi}(\alpha_1 u)$ y $i_{\phi}(\alpha_2 v)$ son finitas, conduce a $i_{\phi}(\alpha((1 - \lambda)u + \lambda v)) < \infty$; así $(1 - \lambda)u + \lambda v$ es un elemento de $L^{\phi}(p \cdot \mu)$; por tanto $L^{\phi}(p \cdot \mu)$ es un conjunto convexo.

Sea $\beta > 0$ se demuestra que βu es un elemento en el espacio del Orlicz, si $\alpha' = \frac{\alpha_1}{\beta}$ entonces $\alpha' \beta = \alpha_1$ y por tanto $\phi(\alpha'(\beta u)) = \phi(\alpha_1 u)$, al integrar se tiene $i_{\phi}(\alpha'(\beta u)) = i_{\phi}(\alpha_1 u) < \infty$, por lo que existe $\alpha' > 0$ tal que $\int \phi(\alpha'(\beta u)) p d\mu < \infty$ para concluir que βu es un elemento

en $L^\phi(p \cdot \mu)$. Consideremos ahora $v \in L^\phi(p \cdot \mu)$, se demuestra que $u + v$ también es un elemento en el espacio de Orlicz. Por propiedades algebraicas y de la convexidad de ϕ , se tiene

$$\begin{aligned} \phi(\alpha(u + v)) &= \phi(\alpha u + \alpha v) = \phi\left(\frac{\alpha}{2}(2u) + \frac{\alpha}{2}(2v)\right) \\ &\leq \frac{\alpha}{2}\phi(2u) + \frac{\alpha}{2}\phi(2v) . \end{aligned}$$

Como $2u$ y $2v$ son elementos en el espacio de Orlicz se concluye con esto que existe $\alpha = \min\{\alpha_1, \alpha_2\} > 0$ tal que $i_\phi(\alpha(u + v)) < \infty$, y por tanto $u + v$ es un elemento del espacio de Orlicz, así $L^\phi(p \cdot \mu)$ es un espacio lineal convexo. \square

Sea ϕ una función de Young, de acuerdo con la definición de clase de Orlicz y de espacio de funciones de Orlicz se sigue que $\mathcal{L}^\phi(p \cdot \mu) \subset L^\phi(p \cdot \mu)$. En efecto, si $u \in \mathcal{L}^\phi(p \cdot \mu)$ entonces $i_\phi(u) < \infty$, si se hace $\alpha = 1$ entonces $i_\phi(\alpha u) < \infty$ y por lo tanto $u \in L^\phi(p \cdot \mu)$. Sin embargo, la inclusión contraria no siempre se presenta, es decir, no es cierto para toda ϕ que $L^\phi(p \cdot \mu) \subset \mathcal{L}^\phi(p \cdot \mu)$. Esta inclusión es cierta cuando la función de Young ϕ satisface la siguiente condición.

Definición 2.6. Una función de Young ϕ satisface la condición Δ_2 (se escribe $\phi \in \Delta_2$) sii

$$\phi(2x) \leq K\phi(x) ,$$

para $x \geq x_0 \geq 0$ y una constante $K > 0$.

Si hacemos que ϕ cumpla la condición Δ_2 entonces, existe un $K > 0$ tal que $\phi(2u) \leq K\phi(u)$ para $x \geq x_0 \geq 0$. Al escribir $\phi(u) = \phi\left(2\frac{u}{2}\right)$ resulta $\phi(u) \leq K\phi\left(\frac{1}{2}u\right)$, esto por la condición Δ_2 . En términos del operador i_ϕ se tiene la desigualdad

$$i_\phi(u) \leq K i_\phi\left(\frac{1}{2}u\right) .$$

Si $\alpha = \frac{1}{2}$ entonces $i_\phi(u) \leq K i_\phi(\alpha u)$ y si u es una variable aleatoria en $L^\phi(p \cdot \mu)$ entonces, $i_\phi(\alpha u) < \infty$ y así $i_\phi(u) < \infty$, es decir, $u \in \mathcal{L}^\phi(p \cdot \mu)$. Se concluye con esto que $L^\phi(p \cdot \mu) \subset \mathcal{L}^\phi(p \cdot \mu)$ para las funciones que verifican la condición Δ_2 y así $L^\phi(p \cdot \mu) = \mathcal{L}^\phi(p \cdot \mu)$.

Al espacio de funciones de Orlicz se le asocia una norma, la cual se conoce como la norma de Luxemburgo, definida para $u \in L^\phi(p \cdot \mu)$ por

$$\|u\|_{\phi,p} = \inf \left\{ k > 0 : i_\phi\left(\frac{u}{k}\right) \leq 1 \right\} = \inf \left\{ k > 0 : \int_{\Omega} \phi\left(\frac{u}{k}\right) p d\mu \leq 1 \right\} . \quad (2.9)$$

En este caso se hace uso del término p para indicar la dependencia de la norma de la medida $p \cdot \mu$; en términos de la medida μ se escribe $\|u\|_{\phi}$. La norma de Luxemburgo caracteriza el

espacio de funciones de Orlicz como

$$L^\phi(p \cdot \mu) = \{u \in L^1(p \cdot \mu) : \|u\|_{\phi,p} < \infty\} .$$

Por el teorema de la convergencia monótona se sigue que si u es una variable aleatoria no nula ($u \neq 0$)

$$i_\phi \left(\frac{u}{\|u\|_{\phi,p}} \right) \leq 1 \quad \text{es equivalente a} \quad \int_{\Omega} \phi \left(\frac{u}{\|u\|_{\phi,p}} \right) p d\mu \leq 1 . \quad (2.10)$$

En el siguiente teorema se demuestra que $\|\cdot\|_{\phi,p}$ es precisamente una norma sobre el espacio de funciones $L^\phi(p \cdot \mu)$.

Teorema 2.5. *Sea ϕ una función de Young, entonces el espacio de funciones de Orlicz $L^\phi(p \cdot \mu)$ es normado.*

Demostración

Si $u = 0$ casi en todas partes de Ω entonces, $i_\phi \left(\frac{u}{k} \right) = 0$ y así el ínfimo respecto de k debe ser cero, lo cual escribimos $\|u\|_{\phi,p} = 0$. Consideremos ahora que $\|u\|_{\phi,p} > 0$, razonando por el absurdo, supongamos que u no es cero casi en todas partes de Ω entonces, para algún $\epsilon_0 > 0$, el conjunto $A = \{x \in \Omega : |u(x)| > \epsilon_0\}$ tiene medida positiva y se escribe $\mu_0 = \mu(A)$. Como $i_\phi \left(\frac{u}{k} \right) \leq 1$, y por las propiedades de la integral de Lebesgue (se hace $\frac{1}{k} = n$) resulta

$$1 \geq i_\phi(nu) = \int_{\Omega} \phi(nu) p d\mu > \int_A \phi(nu) p d\mu = \mu_0 \phi(nu) ,$$

si $n \rightarrow \infty$ entonces $\phi(nu) \rightarrow \infty$ lo cual contradice la desigualdad anterior, por lo que $u = 0$ casi en todas partes de Ω .

Sean u, v variables aleatorias no nulas en $L^\phi(p \cdot \mu)$, por la convexidad de la función ϕ se tiene

$$\begin{aligned} i_\phi \left(\frac{u+v}{\|u\|_{\phi,p} + \|v\|_{\phi,p}} \right) &= \int_{\Omega} \phi \left(\frac{u+v}{\|u\|_{\phi,p} + \|v\|_{\phi,p}} \right) p d\mu \\ &= \int_{\Omega} \phi \left(\frac{u}{\|u\|_{\phi,p} + \|v\|_{\phi,p}} + \frac{v}{\|u\|_{\phi,p} + \|v\|_{\phi,p}} \right) p d\mu \\ &= \int_{\Omega} \phi \left(\frac{\|u\|_{\phi,p}}{\|u\|_{\phi,p} + \|v\|_{\phi,p}} \cdot \frac{u}{\|u\|_{\phi,p}} + \frac{\|v\|_{\phi,p}}{\|u\|_{\phi,p} + \|v\|_{\phi,p}} \cdot \frac{v}{\|v\|_{\phi,p}} \right) p d\mu \\ &\leq \frac{\|u\|_{\phi,p}}{\|u\|_{\phi,p} + \|v\|_{\phi,p}} \int_{\Omega} \phi \left(\frac{u}{\|u\|_{\phi,p}} \right) p d\mu + \frac{\|v\|_{\phi,p}}{\|u\|_{\phi,p} + \|v\|_{\phi,p}} \int_{\Omega} \phi \left(\frac{v}{\|v\|_{\phi,p}} \right) p d\mu \end{aligned}$$

$$\leq \frac{\|u\|_{\phi,p}}{\|u\|_{\phi,p} + \|v\|_{\phi,p}} i_\phi \left(\frac{u}{\|u\|_{\phi,p}} \right) + \frac{\|u\|_{\phi,p}}{\|u\|_{\phi,p} + \|v\|_{\phi,p}} i_\phi \left(\frac{v}{\|v\|_{\phi,p}} \right).$$

Para u y v se satisface la expresión (2.10), donde $i_\phi \left(\frac{u}{\|u\|_{\phi,p}} \right) \leq 1$ y $i_\phi \left(\frac{v}{\|v\|_{\phi,p}} \right) \leq 1$, por tanto

$$i_\phi \left(\frac{u+v}{\|u\|_{\phi,p} + \|v\|_{\phi,p}} \right) \leq \frac{\|u\|_{\phi,p}}{\|u\|_{\phi,p} + \|v\|_{\phi,p}} + \frac{\|v\|_{\phi,p}}{\|u\|_{\phi,p} + \|v\|_{\phi,p}} = 1$$

Se logra con esto que $\|u+v\|_{\phi,p}$ es finita. Si en (2.9) se hace $k = \|u\|_{\phi,p} + \|v\|_{\phi,p}$ y se toma ínfimo resulta la desigualdad

$$\begin{aligned} \|u+v\|_{\phi,p} &\leq \frac{\|u\|_{\phi,p}}{\|u\|_{\phi,p} + \|v\|_{\phi,p}} \|u\|_{\phi,p} + \frac{\|v\|_{\phi,p}}{\|u\|_{\phi,p} + \|v\|_{\phi,p}} \|v\|_{\phi,p} \\ &\leq \frac{\|u\|_{\phi,p}^2}{\|u\|_{\phi,p} + \|v\|_{\phi,p}} + \frac{\|v\|_{\phi,p}^2}{\|u\|_{\phi,p} + \|v\|_{\phi,p}} = \frac{\|u\|_{\phi,p}^2 + \|v\|_{\phi,p}^2}{\|u\|_{\phi,p} + \|v\|_{\phi,p}} \\ &\leq \frac{(\|u\|_{\phi,p} + \|v\|_{\phi,p})^2}{\|u\|_{\phi,p} + \|v\|_{\phi,p}} \leq \|u\|_{\phi,p} + \|v\|_{\phi,p} \end{aligned}$$

lo cual prueba la desigualdad triangular y la aplicación $\|\cdot\|_{\phi,p}$ es una norma sobre el espacio $L^\phi(p \cdot \mu)$. \square

La completitud del espacio de funciones de Orlicz se puede ver en [10] página 98, por lo que $L^\phi(p \cdot \mu)$ es un espacio lineal normado y completo, por tanto un espacio de Banach.

Definición 2.7. *Dos funciones de Young ϕ y ψ se dice que son equivalentes y se escribe $\phi \equiv \psi$ si existen constantes C_1 y C_2 positivas, tales que $C_1 \leq C_2$ y*

$$\phi(C_1 x) \leq \psi(x) \leq \phi(C_2 x),$$

para $x \geq x_0 \geq 0$.

De la definición 2.7 se sigue que: si ϕ y ψ son funciones equivalentes entonces los espacios de funciones Orlicz asociados a cada función son iguales $L^\phi(p \cdot \mu) = L^\psi(p \cdot \mu)$. Además se presenta la equivalencia entre las normas y de las funciones complementarias $\phi^* \equiv \psi^*$.

Teorema 2.6. *Sean ϕ y ψ funciones de Young equivalentes entonces, $\phi^* \equiv \psi^*$ y $L^\phi(p \cdot \mu) = L^\psi(p \cdot \mu)$.*

Demostración

Supongamos que ϕ y ψ son funciones de Young equivalentes, existen constantes positivas C_1 y C_2 tales que $\phi(C_1x) \leq \psi(x) \leq \phi(C_2x)$ con $x \geq x_0 \geq 0$ y $C_1 \leq C_2$. Multiplicando por -1 y sumando el término C_1C_2xy resulta la desigualdad

$$(C_2x)(C_1y) - \phi(C_2x) \leq x(C_1C_2y) - \psi(x) \leq (C_1x)(C_2y) - \phi(C_1x) .$$

Por las propiedades del supremo y la definición de la función de Young complementaria (2.3) resulta la desigualdad $\phi^*(C_1y) \leq \psi^*(C_1C_2y) \leq \phi^*(C_2y)$, si se hace el cambio de variable $C_1C_2y = y'$ entonces

$$\phi^* \left(\frac{y'}{C_2} \right) \leq \psi^*(y') \leq \phi^* \left(\frac{y'}{C_1} \right) ,$$

expresión equivalente a $\phi^*(C'_1y') \leq \psi^*(y') \leq \phi^*(C'_2y')$ donde $C'_1 = \frac{1}{C_2}$ y $C'_2 = \frac{1}{C_1}$. Como $C_1 \leq C_2$ entonces $C'_1 \leq C'_2$ y así las funciones de Young ϕ^* y ψ^* son equivalentes y se escribe $\phi^* \equiv \psi^*$.

Veamos ahora que $L^\phi(p \cdot \mu) = L^\psi(p \cdot \mu)$. Si $u \in L^\phi(p \cdot \mu)$, existe una constante positiva α tal que $i_\phi(\alpha u) < \infty$. Como ϕ y ψ son equivalentes entonces, existen constantes positivas C_1 y C_2 para las que $\phi(C_1u) \leq \psi(u) \leq \phi(C_2u)$. Si se sustituye u por αu resulta $\phi(C_1(\alpha u)) \leq \psi(\alpha u) \leq \phi(C_2(\alpha u))$, en términos del funcional i_ϕ se puede escribir como

$$i_\phi(C_1(\alpha u)) \leq i_\psi(\alpha u) \leq i_\phi(C_2(\alpha u)) . \quad (1)$$

Como $i_\phi(C_2(\alpha u)) < \infty$ en (1) se sigue que $i_\psi(\alpha u) < \infty$, luego $u \in L^\psi(p \cdot \mu)$, por lo que $L^\phi(p \cdot \mu) \subset L^\psi(p \cdot \mu)$. Recíprocamente, si $u \in L^\psi(p \cdot \mu)$ entonces existe una constante $\beta > 0$ tal que $i_\psi(\beta u) < \infty$, para β se tiene la desigualdad $\phi(C_1\beta u) \leq \psi(\beta u) \leq \phi(C_2\beta u)$ que es equivalente a $i_\phi(C_1\beta u) \leq i_\psi(\beta u)$, y por tanto $i_\phi((C_1\beta)u) < \infty$. Si $\theta = C_1\beta$ entonces $\theta > 0$ y $i_\phi(\theta u) < \infty$, lo que permite concluir que $u \in L^\phi(p \cdot \mu)$ y $L^\psi(p \cdot \mu) \subset L^\phi(p \cdot \mu)$. Obteniendo así la igualdad de los espacios. \square

Como consecuencia de la igualdad de los espacios de Orlicz, se tiene que las normas son equivalentes. Sea ϕ una función de Young y ϕ^* su función complementaria, se verifica la siguiente versión de la desigualdad de Hölder presentada en (2.7) para los espacios de Lebesgue.

Teorema 2.7. *Sea ϕ una función de Young. Si $u \in L^\phi(p \cdot \mu)$ y $v \in L^{\phi^*}(p \cdot \mu)$ entonces $uv \in L^1(p \cdot \mu)$ y*

$$\int_{\Omega} |uv| p d\mu \leq 2 \|u\|_{\phi, p} \|v\|_{\phi^*, p} .$$

Demostración

En el caso en que $u = v = 0$ entonces $\|u\|_{\phi,p} = \|v\|_{\phi,p} = 0$ y se presenta la igualdad. Si u, v son variables aleatorias no nulas entonces por la desigualdad de Young presentada en (2.5) se tiene

$$\frac{|u|}{\|u\|_{\phi,p}} \cdot \frac{|v|}{\|v\|_{\phi^*,p}} \leq \phi\left(\frac{|u|}{\|u\|_{\phi,p}}\right) + \phi^*\left(\frac{|v|}{\|v\|_{\phi^*,p}}\right).$$

Integrando en ambos lados y haciendo uso de las propiedades de la integral de Lebesgue se tiene

$$\int_{\Omega} \frac{|uv|}{\|u\|_{\phi,p}\|v\|_{\phi^*,p}} p d\mu \leq i_{\phi}\left(\frac{|u|}{\|u\|_{\phi,p}}\right) + i_{\phi^*}\left(\frac{|v|}{\|v\|_{\phi^*,p}}\right) \leq 1 + 1 \leq 2.$$

Ya que $\|u\|_{\phi,p}\|v\|_{\phi^*,p}$ es una constante positiva (respecto de medida $p \cdot \mu$) se obtiene que

$$\int_{\Omega} |uv| p d\mu \leq 2\|u\|_{\phi,p}\|v\|_{\phi^*,p},$$

Por lo que uv es un elemento en $L^1(p \cdot \mu)$. \square

La desigualdad de Hölder presentada en (2.7) implica la existencia del número

$$\begin{aligned} N_{\phi,p}(u) &= \sup \left\{ \int_{\Omega} |uv| p d\mu : v \in L^{\phi^*}(p \cdot \mu), \int_{\Omega} \phi^*(v) p d\mu \leq 1 \right\} \\ N_{\phi,p}(u) &= \sup \left\{ \int_{\Omega} |uv| p d\mu : v \in L^{\phi^*}(p \cdot \mu), i_{\phi^*}(v) \leq 1 \right\}. \end{aligned} \quad (2.11)$$

La condición que $i_{\phi^*}(v) \leq 1$ implica que $\|v\|_{\phi^*,p} \leq 1$, es decir, v es un elemento en la bola unitaria cerrada del espacio de funciones de Orlicz asociado a la función de Young complementaria. La función $N_{\phi,p}$ recibe el nombre de Norma de Orlicz y se relaciona con la norma de Luxemburgo mediante la desigualdad $\|u\|_{\phi,p} \leq N_{\phi,p}(u) \leq 2\|u\|_{\phi,p}$, lo cual garantiza la equivalencia de las normas; y por tanto inducen la misma topología sobre el espacio de funciones de Orlicz $L^{\phi}(p \cdot \mu)$.

A manera de ilustración consideremos las funciones de Young ϕ_1 y ϕ_2 definidas como

$$\phi_1(x) = \cosh(x) - 1 \quad \text{y} \quad \phi_2(x) = e^{|x|} - |x| - 1. \quad (2.12)$$

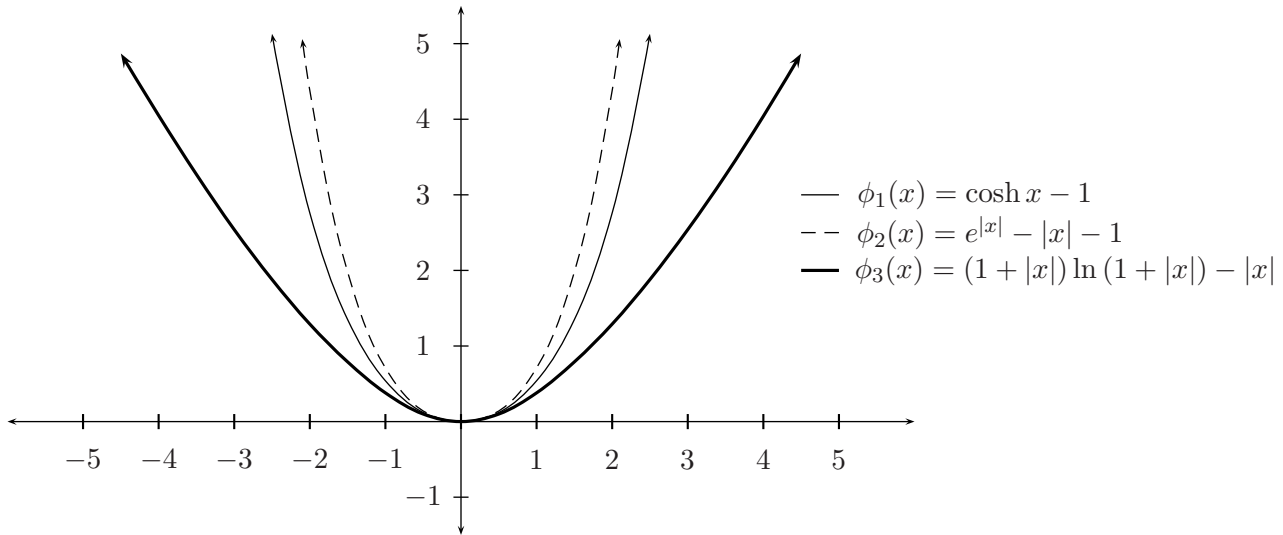
Se afirma que la función ϕ_2 es equivalente a ϕ_1 , por lo que se debe encontrar constantes positivas $C_1 < C_2$ para las cuales $\phi_2(C_1x) \leq \phi_1(x) \leq \phi_2(C_2x)$. Como la función $\phi_2(x) = e^{|x|} - |x| - 1$ es de Young, por (2.4) se puede escribir de la forma

$$\phi_2(x) = \int_0^{|x|} \underbrace{(e^t - 1)}_{\varphi(t)} dt = e^{|x|} - |x| - 1.$$

Para determinar la función complementaria de ϕ_2 , se requiere hallar la inversa de $\varphi(t)$. Haciendo $y = e^t - 1$ y despejando t en función de y se logra $t = \ln(1 + y)$, por lo que $\theta(t) = \ln(1 + t)$; por (2.6) se tiene que la función complementaria es

$$\phi_2^*(y) = \int_0^{|y|} \ln(1 + t) dt = (1 + t) \ln(1 + t) - (1 + t)|_0^{|y|} = (1 + |y|) \ln(1 + |y|) - |y|,$$

Si se denota a ϕ_2^* por ϕ_3 es claro que $\phi_2^*(x) = \phi_3(x)$ y $\phi_3^*(x) = \phi_2(x)$ (ver teorema 2.2).



Por su parte, la función ϕ_3 satisface la condición Δ_2 expuesta en la definición 2.6, al demostrar la desigualdad $\phi_3(2x) \leq 4\phi_3(x)$ se sigue que $K = 4$. Para $x \geq 0$, la derivada de $\phi_3(x)$ es $\phi_3'(x) = \ln(1 + x)$, mientras que la derivada de $\phi_3(2x)$ es $\phi_3'(2x) = 2 \ln(1 + 2x)$. Consideremos la función $f(x) = \frac{\phi_3(2x)}{\phi_3(x)}$, la derivada de $f(x)$ es

$$f'(x) = \frac{\ln(1 + 2x) \ln(1 + x) - 2x \ln\left(\frac{1+2x}{1+x}\right)}{\phi_3^2(x)}.$$

Como $\ln(1 + 2x) \ln(1 + x) \leq 2x \ln\left(\frac{1+2x}{1+x}\right)$ para todo $x > -\frac{1}{2}$ entonces, $f'(x) < 0$ y por tanto f es una función decreciente. La función f no está definida en cero, sin embargo es posible hallar $\lim_{x \rightarrow 0} f(x)$, por la regla de L'Hopital se tiene,

$$\lim_{x \rightarrow 0} f(x) = \lim_{x \rightarrow 0} \frac{(1 + 2x) \ln(1 + 2x) - 2x}{(1 + x) \ln(1 + x) - x} = \lim_{x \rightarrow 0} \frac{2 \ln(1 + 2x)}{\ln(1 + x)} = \lim_{x \rightarrow 0} \frac{4(x + 1)}{1 + 2x} = 4.$$

Se redefine f de tal forma que $f(0) = 4$. Como f es decreciente entonces, para $x \geq 0$ se sigue que $f(0) \geq f(x)$, esto es $\frac{\phi_3(2x)}{\phi_3(x)} \leq 4$ para concluir que $\phi_3(2x) \leq 4\phi_3(x)$.

2.3. Analiticidad entre espacios de Banach

Sean E y F espacios de Banach. Se denota por $\mathcal{L}(E; F)$ el espacio de todos los mapeos lineales y continuos (acotados) de E en F ; que con la norma uniforme (del supremo) tiene estructura de espacio de Banach. Por $\mathcal{L}^n(E; F)$ al espacio de todos los mapeos multilineales y continuos de E^n a F , donde $n \in \mathbb{N}$ y $E^n := \underbrace{E \times E \times \dots \times E}_{n\text{-veces}}$ y la multilinealidad es

la linealidad por cada una de las componentes. Y $\mathcal{L}_s^n(E; F) \subset \mathcal{L}^n(E; F)$ es el operador de todos los operadores lineales, continuos y simétricos, donde se entiende por simetría el intercambio del orden de las componentes sin alterar el resultado de la aplicación. En el caso en que $F = \mathbb{R}$, se denotan los conjuntos anteriores como $\mathcal{L}(E)$, $\mathcal{L}^n(E)$ y $\mathcal{L}_s^n(E)$.

Un operador multilineal λ es acotado si y solo si existe $k > 0$ tal que $\|\lambda(v_1, v_2, \dots, v_n)\|_F \leq k\|v_1\|_E \dots \|v_n\|_E$, para todo $v_1, v_2, \dots, v_n \in E^n$.

Definición 2.8. Una función $\hat{\lambda} : E \rightarrow F$ entre espacios de Banach es llamado un polinomio n -homogéneo continuo, si existe un $\lambda \in \mathcal{L}_s^n(E; F)$ tal que $\hat{\lambda}(x) = \lambda x^n$ para todo $x \in E$.

Se denota por $\mathcal{P}^n(E; F)$ al espacio de todos los polinomios continuos n -homogéneos de E en F . En caso que $n = 0$ se asume $\mathcal{P}^0(E; F) := F$. Nótese que la linealidad del operador λ no la hereda el operador $\hat{\lambda}$.

El operador simétrico multilineal λ está determinado de forma única por el operador $\hat{\lambda}$, en efecto, si se aplica el operador diferencial n -veces a $\hat{\lambda}$ (en la definición 2.8) resulta la denominada forma polar de $\hat{\lambda}$, esto es, $D^n \hat{\lambda}(x) = D^n \lambda x^n = n! \lambda$.

Sobre $\mathcal{P}^n(E; F)$ se define la norma

$$\|\hat{\lambda}\|_{\mathcal{P}^n} = \sup_{x \neq 0} \frac{\|\hat{\lambda}(x)\|_F}{\|x\|_E^n} = \sup_{\|x\|_E=1} \|\hat{\lambda}(x)\|_F,$$

de donde resulta $\|\hat{\lambda}\|_{\mathcal{P}^n} \leq \|\lambda\|_{\mathcal{L}^n}$ y procediendo de forma inductiva se obtiene la desigualdad

$$\|\hat{\lambda}\|_{\mathcal{P}^n} \leq \|\lambda\|_{\mathcal{L}^n} \leq \frac{n^n}{n!} \|\hat{\lambda}\|_{\mathcal{P}^n}, \quad (2.13)$$

que establece la equivalencia entre las normas de \mathcal{P}^n y \mathcal{L}^n .

Definición 2.9. Sea $\{\hat{\lambda}_n\}_{n \in \mathbb{N}}$ una sucesión con $\hat{\lambda}_n \in \mathcal{P}^n(E; F)$ entonces,

$$\sum_{n=0}^{\infty} \hat{\lambda}_n$$

es llamada serie de potencias de E en F . El radio de convergencia $\hat{\rho}$ de la serie es definido como

$$\hat{\rho} = \sup \left\{ r \in [0, \infty) : \sum_{n=0}^{\infty} \|\hat{\lambda}_n\|_{\mathcal{P}^n} r^n < \infty \right\} .$$

El radio de convergencia restringido ρ de la serie de potencias está definido como

$$\rho = \sup \left\{ r \in [0, \infty) : \sum_{n=0}^{\infty} \|\lambda_n\|_{\mathcal{L}^n} r^n < \infty \right\} .$$

Un resultado importante es la ecuación de Cauchy-Hadamard dada por

$$\frac{1}{\hat{\rho}} = \limsup_{n \rightarrow \infty} \|\hat{\lambda}_n\|_{\mathcal{P}^n}^{\frac{1}{n}} \quad \text{y} \quad \frac{1}{\rho} = \limsup_{n \rightarrow \infty} \|\lambda_n\|_{\mathcal{L}^n}^{\frac{1}{n}} . \quad (2.14)$$

Si cada miembro de la desigualdad presentada en (2.13) se eleva a la potencia $\frac{1}{n}$ para n fijo, resulta

$$\|\hat{\lambda}\|_{\mathcal{P}^n}^{\frac{1}{n}} \leq \|\lambda\|_{\mathcal{L}^n}^{\frac{1}{n}} \leq \left(\frac{n^n}{n!}\right)^{\frac{1}{n}} \|\hat{\lambda}\|_{\mathcal{P}^n}^{\frac{1}{n}} ,$$

se sigue que sus límites superiores conservan la desigualdad, por lo que

$$\limsup_{n \rightarrow \infty} \|\hat{\lambda}\|_{\mathcal{P}^n}^{\frac{1}{n}} \leq \limsup_{n \rightarrow \infty} \|\lambda\|_{\mathcal{L}^n}^{\frac{1}{n}} \leq \limsup_{n \rightarrow \infty} \left(\frac{n^n}{n!}\right)^{\frac{1}{n}} \|\hat{\lambda}\|_{\mathcal{P}^n}^{\frac{1}{n}} .$$

Haciendo uso de la igualdad presentada en (2.14) y el hecho que $\lim_{n \rightarrow \infty} \left(\frac{n^n}{n!}\right)^{\frac{1}{n}} = e$ resulta la desigualdad $\frac{1}{\hat{\rho}} \leq \frac{1}{\rho} \leq \frac{e}{\hat{\rho}}$ que es equivalente a

$$\frac{\hat{\rho}}{e} \leq \rho \leq \hat{\rho} . \quad (2.15)$$

Esta desigualdad muestra la equivalencia entre los radios de convergencia de la serie de potencias $\sum_{n=0}^{\infty} \hat{\lambda}_n$. La desigualdad $\|\hat{\lambda}_n(x)\|_F \leq \|\hat{\lambda}_n\|_{\mathcal{P}^n} \|x\|_E^n$ implica que $\sum_{n=0}^{\infty} \hat{\lambda}_n(x)$ converge absolutamente y uniformemente en la bola cerrada $\overline{B}(0, r)$, para cada $r < \hat{\rho}$. Además, el radio de convergencia de dicha serie se puede escribir como

$$\hat{\rho} = \sup \left\{ r : \sum_{n=0}^{\infty} \hat{\lambda}_n(x) \text{ converge uniformemente para } \|x\|_E \leq r \right\} .$$

Definición 2.10. Sean E y F espacios de Banach y $U \subset E$ un conjunto abierto. Un mapeo $f : U \rightarrow F$ es llamado analítico si para cada $x_0 \in U$ existe una serie de potencias convergente $\sum_{n=0}^{\infty} \hat{\lambda}_n$ con radio de convergencia positivo tal que, para cada x en una vecindad de x_0

$$f(x) = \sum_{n=0}^{\infty} \hat{\lambda}_n (x - x_0) .$$

La serie de potencias $\sum_{n=0}^{\infty} \hat{\lambda}_n(x)$ es una función analítica en la bola $B(0, \rho)$, es decir, es analítica una vecindad abierta del cero con radio de convergencia ρ . De acuerdo con la definición de la función polinomial n -homogénea continua dada en (2.8), la función f de la definición 2.10 se puede escribir como

$$f(x) = \sum_{n=0}^{\infty} \lambda_n (x - x_0)^n .$$

Una función que es analítica es infinitamente diferenciable sobre la vecindad de x_0 dada en la definición 2.10, donde f es tal que $D^k f(x_0) = k! \lambda_k$ y $D^k f$ tiene la siguiente expansión como serie de potencia

$$D^k f(x) = \sum_{n=0}^{\infty} \frac{(k+n)!}{n!} \lambda_{k+n} (x - x_0)^n .$$

Teorema 2.8. Sean E y F espacios de Banach, $U \subset E$ un conjunto abierto y $f : U \rightarrow F$ una función suave. Si existe una constante M tal que para todo $x \in B(x_0, r) \subset U$

$$\|D^n f(x)\|_{\mathcal{L}^n} \leq \frac{Mn!}{r^n}$$

entonces, f tiene una representación en serie de Taylor centrada en x_0 como

$$f(x) = \sum_{n=0}^{\infty} \frac{1}{n!} D^n f(x_0) (x - x_0)^n$$

y además f es analítica en $B(x_0, r)$.

El siguiente teorema será útil más adelante.

Teorema 2.9. Sean E, F y G espacios de Banach, $U \subset E$ y $V \subset F$ conjuntos abiertos. Si $f : U \rightarrow F$ y $g : V \rightarrow G$ son funciones analíticas tales que $f(U) \subset V$ entonces $g \circ f$ es analítica.

2.4. k -Deformaciones

Sin hacer una exposición detallada, esta sección presenta algunos aspectos de las matemáticas k -deformadas según Kaniadakis, como aparecen en los artículos [14, 15]. Para ampliar detalles, se puede consultar la tesis de Dora Deossa [7].

Considérese un parámetro real k , con $-1 < k < 1$. Para cada número real x , se define $x_{\{k\}}$ y $x^{\{k\}}$ por: $x_{\{k\}} := \frac{1}{k} \operatorname{arcsenh}(kx)$ si $k \neq 0$, $x_{\{0\}} = x$, $x^{\{k\}} = \frac{1}{k} \operatorname{senh}(kx)$ si $k \neq 0$ y $x^{\{0\}} = x$. Lo anterior, define funciones inversas y por tanto $(x^{\{k\}})_{\{k\}} = (x_{\{k\}})^{\{k\}} = x$. La k -suma se define por

$$x \oplus^k y = (x_{\{k\}} + y_{\{k\}})^{\{k\}} = \frac{1}{k} \operatorname{senh}[\operatorname{arcsenh}(kx) + \operatorname{arcsenh}(ky)]$$

y dota al conjunto de números reales de estructura de grupo abeliano, donde el elemento neutro es cero y el inverso de un real w es $-w$. La k -suma se reduce a la suma usual cuando $k \rightarrow 0$, esto es, $x \oplus^0 y = x + y$. Resulta muy útil expresar la k -suma por $(x \oplus^k y)_{\{k\}} = x_{\{k\}} + y_{\{k\}}$ de la siguiente forma

$$x \oplus^k y = x \sqrt{1 + k^2 y^2} + y \sqrt{1 + k^2 x^2}. \quad (2.16)$$

Como es de esperar, la k -diferencia (denotada con \ominus^k) es definida por $x \ominus^k y = x \oplus^k (-y)$. También se introduce el k -producto mediante la ley de composición \otimes^k , definida por $x \otimes^k y := (x_{\{k\}} \cdot y_{\{k\}})^{\{k\}}$ y, como antes, $x \otimes^0 y = xy$. La estructura algebraica $(\mathbb{R}, \oplus^k, \otimes^k)$ es un campo, isomorfo al usual en los reales, en el cual: el elemento neutro multiplicativo (k -multiplicativo) es $1^{\{k\}} = \frac{1}{k} \operatorname{senh}(k)$, el elemento inverso multiplicativo (k -multiplicativo) para cada real $y \neq 0$ es $\bar{y} = (1/y_{\{k\}})^{\{k\}}$ y, desde luego, la k -división \oslash^k está definida por $x \oslash^k y = x \otimes^k \left(\frac{1}{y_{\{k\}}} \right)^{\{k\}}$.

De otro lado, la función exponencial k -deformada se puede introducir mediante una modificación del concepto de derivada (a derivada k -deformada), de manera que la función exponencial k -deformada quede invariante bajo dicha derivada y se puedan extender con facilidad algunas propiedades de la exponencial natural. La modificación de la derivada propuesta por Kaniadakis se puede expresar como sigue: la k -derivada una de una función real de variable real f , que sea derivable, es dada por

$$\frac{df(x)}{d_{\{k\}}x} := \sqrt{1 + k^2 x^2} \frac{df(x)}{dx}. \quad (2.17)$$

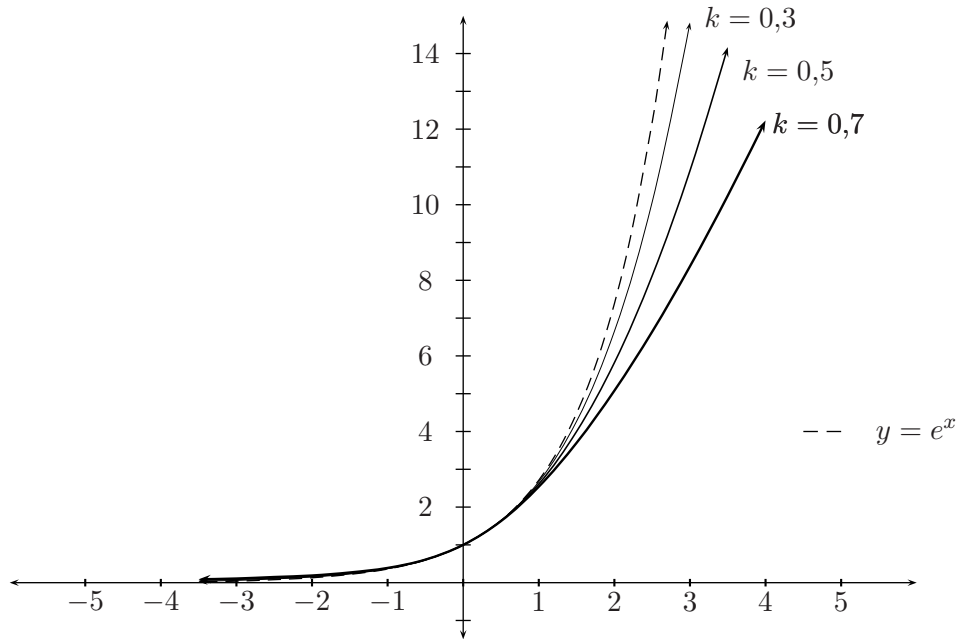
El k -diferencial se define por $d_{\{k\}}x = \sqrt{1+k^2x^2}dx$. Es claro que la k -derivada es regida por las mismas reglas de la derivada usual y que nuevamente, se obtiene la derivada usual cuando $k \rightarrow 0$. Ahora, se define la función **exponencial k -deformada o k -exponencial**, cuando $k = 0$ por $\exp_0(x) = \exp(x)$ y cuando $k \neq 0$ por

$$\exp_k(x) := \exp(x_{\{k\}}) = (\sqrt{1+k^2x^2} + kx)^{\frac{1}{k}}. \quad (2.18)$$

Esta función tiene gran similitud con la exponencial natural; ambas se definen de los reales en los reales positivos, son de clase $C^\infty(\mathbb{R})$, estrictamente crecientes, convexas,

$$\lim_{x \rightarrow -\infty} \exp_k(x) = 0 \text{ y } \lim_{x \rightarrow \infty} \exp_k(x) = \infty.$$

En el siguiente gráfico se hace un esbozo de algunas funciones exponenciales deformadas obtenidas. La función punteada hace alusión a la exponencial usual, nótese que a medida que $k \rightarrow 0$ entonces la exponencial deformada tiene a la exponencial usual.



Según la k -derivada y la derivada usual se cumple

$$\frac{d}{d_{\{k\}}x} \exp_k(x) = \exp_k(x) \text{ y } \frac{d}{dx} \exp_k(x) = \sqrt{1+k^2x^2} \exp_k(x). \quad (2.19)$$

La función k -exponencial cumple las siguiente propiedades.

1. $\exp_k(0) = 1$ y $\exp_{\{-k\}}(x) = \exp_k(x)$.
2. $\exp_k(x) \exp_k(-x) = 1$ y $(\exp_k(x))^r = \exp_{\{k/r\}}(rx)$, ($r \neq 0$).

$$3. \exp_k(x) \exp_k(y) = \exp_k(x \overset{k}{\oplus} y).$$

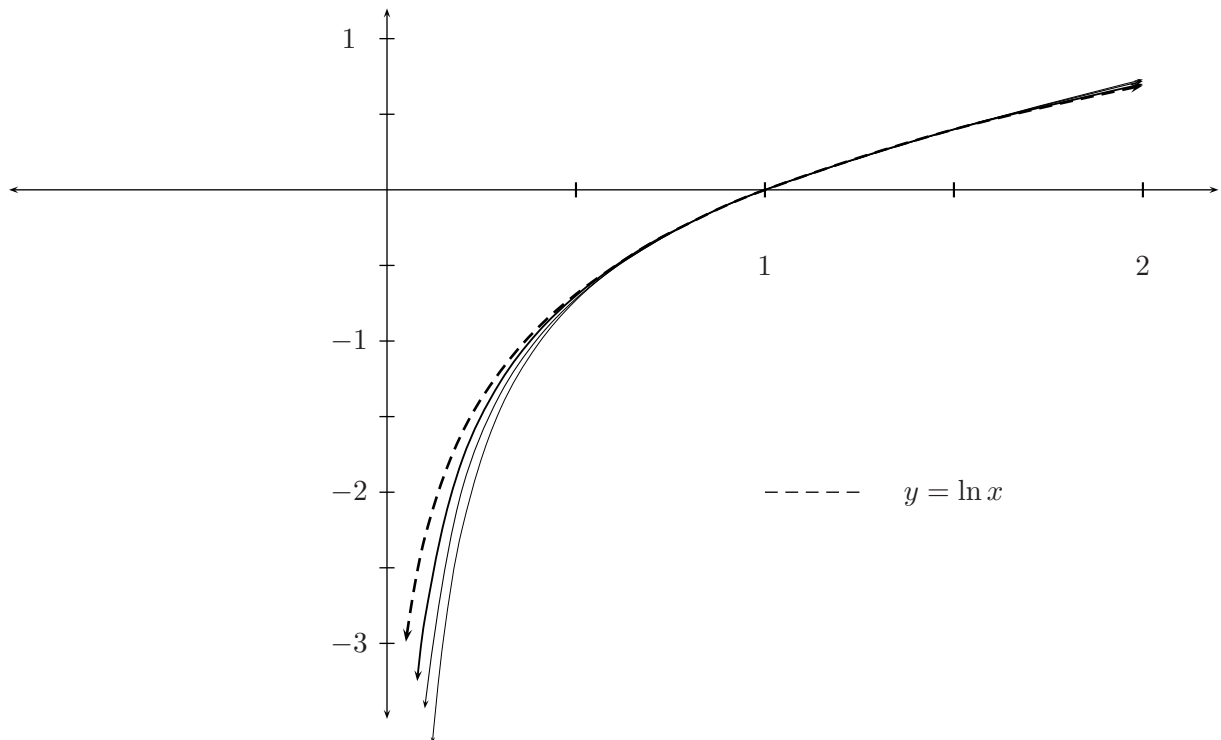
$$4. \exp_k(x) / \exp_k(y) = \exp_k(x \overset{k}{\ominus} y).$$

La función k -**logaritmo**, denotada por $\ln_k(\cdot)$, es definida como la función inversa del k -exponencial. Precisamente, es función de los reales positivos en los reales y está dada: cuando $k = 0$ por $\ln_{\{0\}}(x) := \ln x$ y cuando $k \neq 0$ por

$$\ln_k(x) := (\ln x)^{\{k\}} = \frac{x^k - x^{-k}}{2k}. \quad (2.20)$$

Esta función es: de clase $C^\infty(\mathbb{R}^+)$, estrictamente creciente, cóncava, $\lim_{x \rightarrow 0^+} \ln_k(x) = -\infty$ y $\lim_{x \rightarrow \infty} \ln_k(x) = \infty$.

En el siguiente gráfico, se ilustran algunas de las funciones logarítmicas deformadas, de abajo hacia arriba los valores de la deformación son $k = 0,9$, $k = 0,7$ y $k = 0,5$ donde la línea punteada representa la función logarítmica usual.



Además, se cumplen las siguientes propiedades.

$$1. \ln_k(1) = 0 \text{ y } \ln_{\{-k\}}(x) = \ln_k(x).$$

2. $\ln_k(\frac{1}{x}) = -\ln_k(x)$ y $\ln_k(x^r) = r \ln_{\{rk\}}(x)$, ($r \in \mathbb{R}$).
3. $\ln_k(xy) = \ln_k(x) \oplus^k \ln_k(y)$.
4. $\ln_k(x/y) = \ln_k(x) \ominus^k \ln_k(y)$.

Para extender la k -exponencial a los números complejos, se puede usar la expansión en serie de Taylor para $\exp_k(\cdot)$ en torno a $x_0 = 0$, la cual está dada por:

$$\exp_k(x) = \sum_{n=0}^{\infty} a_n(k) \frac{x^n}{n!}, \quad k^2 x^2 < 1,$$

donde los coeficientes a_n son definidos por $a_0(k) = 1$, $a_1(k) = 1$, $a_{2m}(k) = \prod_{j=0}^{m-1} [1 - (2j)^2 k^2]$ y $a_{2m+1}(k) = \prod_{j=1}^m [1 - (2j-1)^2 k^2]$. La k -exponencial de un número imaginario ix se define por: $\exp_k(ix) = \sum_{n=0}^{\infty} a_n(k) \frac{(ix)^n}{n!}$ y además se prueba que

$$\exp_k(ix) = \sum_{n=0}^{\infty} (-1)^n a_{2n}(k) \frac{x^{2n}}{(2n)!} + i \sum_{n=0}^{\infty} (-1)^n a_{2n+1}(k) \frac{x^{2n+1}}{(2n+1)!}.$$

Las series de la expresión anterior, convergen absolutamente cuando $|x| < |k|^{-1}$. Si se denota $\cos_k(x) = \sum_{n=0}^{\infty} (-1)^n a_{2n}(k) \frac{x^{2n}}{(2n)!}$ y $\sin_k(x) = \sum_{n=0}^{\infty} (-1)^n a_{2n+1}(k) \frac{x^{2n+1}}{(2n+1)!}$, se tiene una versión de la fórmula de Euler, es decir, $\exp_k(\pm ix) = \cos_k(x) \pm i \sin_k(x)$.

Muchos desarrollos pueden ser considerados si se cambian las funciones exponencial y logarítmica naturales por las correspondientes k -deformadas. Por ejemplo, la trigonometría circular k -deformada se puede definir a partir de la trigonometría hiperbólica k -deformada, la cual se inicia con las definiciones de k -seno y k -coseno hiperbólicos como $\sinh_{\{k\}}(x) = \frac{1}{2}(\exp_k(x) - \exp_k(-x))$ y $\cosh_{\{k\}}(x) = \frac{1}{2}(\exp_k(x) + \exp_k(-x))$. Se obtienen versiones de las fórmulas de Euler y Moivre, y en general, muchas identidades que son k -deformaciones de las conocidas.

Capítulo 3

Variedad de Información

3.1. Introducción

En este capítulo se presenta la construcción de la variedad de información de G. Pistone y C. Sempì [22]. Sea $(\Omega, \mathcal{A}, \mu)$ un espacio de probabilidad. Considérese el conjunto \mathfrak{M}_μ constituido por todas las densidades de probabilidad positivas casi en todas partes de Ω , es decir,

$$\mathfrak{M}_\mu = \left\{ p : \Omega \rightarrow \mathbb{R} : p > 0 \text{ c.t.p. y } \int_{\Omega} p d\mu = 1 \right\}.$$

Precisamente, se construye una variedad en \mathfrak{M}_μ , modelada sobre espacios de Banach tales que para cada $p \in \mathfrak{M}_\mu$, son subespacios cerrados del correspondiente espacio de Orlicz $L^{\phi_1}(p \cdot \mu)$. Es por ello que se requiere obtener las propiedades de los espacios de Orlicz $L^{\phi_1}(p \cdot \mu)$, para luego demostrar la analiticidad de los mapeos M_p y K_p (denominados función generadora de momentos y función acumulante respectivamente); mediante los cuales definen posteriormente los mapeos transición para la variedad. La analiticidad de dichos mapeos, implicará que la variedad construida sea de clase C^∞ . Lo anterior establece el orden a seguir en la presentación de este capítulo, cuyo contenido corresponde a los trabajos de: Pistone-Sempì, Pistone-Rogantin y la tesis doctoral de Alberto Cena, respectivamente [22, 21, 5].

3.2. Generalidades sobre el espacio $L^{\phi_1}(p \cdot \mu)$

El conjunto \mathfrak{M}_μ no tiene la estructura de espacio lineal, ya que si p y q están en \mathfrak{M}_μ , no se cumple $(p + q) \in \mathfrak{M}_\mu$ ya que $\int_{\Omega} (p + q) d\mu = 2$. Sin embargo, siempre se cumple que $\frac{1}{2}(p + q) \in \mathfrak{M}_\mu$. A pesar de no ser un espacio lineal, se cuenta con el siguiente resultado.

Teorema 3.1. *El conjunto \mathfrak{M}_μ es convexo.*

Demostración

Sean p_1 y p_2 densidades en \mathfrak{M}_μ , así $\int_\Omega p_1 d\mu = 1$ y $\int_\Omega p_2 d\mu = 1$. Sea λ un escalar tal que $0 \leq \lambda \leq 1$ entonces, se tiene

$$\int_\Omega (\lambda p_1 + (1 - \lambda)p_2) d\mu = \lambda \int_\Omega p_1 d\mu + (1 - \lambda) \int_\Omega p_2 d\mu = \lambda + (1 - \lambda) = 1.$$

De lo anterior, $\lambda p_1 + (1 - \lambda)p_2 \in \mathfrak{M}_\mu$ y así se concluye que \mathfrak{M}_μ es un conjunto convexo. \square

Si $p \in \mathfrak{M}_\mu$, p es una densidad de probabilidad y por tanto determina una medida dada por $p \cdot \mu$ como se presentó en los preliminares (ver 2.1).

Consideremos ahora la función $\phi_1(\cdot) : \mathbb{R} \rightarrow \overline{\mathbb{R}}$ definida por $\phi_1(x) = \cosh x - 1$, la cual es una función de Young. Por tanto ϕ_1 determina a la clase de Orlicz

$$\mathcal{L}^{\phi_1}(p \cdot \mu) = \{u \in L^1(p \cdot \mu) : E_p[\phi_1(u)] < \infty\}$$

y al espacio de funciones de Orlicz

$$L^{\phi_1}(p \cdot \mu) = \{u \in L^1(p \cdot \mu) : E_p[\phi_1(\alpha u)] < \infty \text{ para alguna } \alpha > 0\},$$

dotado de la norma de Luxemburgo:

$$\|u\|_{\phi_1, p} = \inf \left\{ r > 0 : E_p \left[\phi_1 \left(\frac{u}{r} \right) \right] \leq 1 \right\}.$$

En este trabajo, dicha norma se escribirá como $\|u\|_{\phi_1, p} := \|u\|_{\phi_1}$. En consecuencia, $L^{\phi_1}(p \cdot \mu)$ es un espacio de Banach (subespacio cerrado de $L^1(p \cdot \mu)$), que puede ser descrito por $L^{\phi_1}(p \cdot \mu) = \{u \in L^1(p \cdot \mu) : \|u\|_{\phi_1} < \infty\}$ o también por

$$L^{\phi_1}(p \cdot \mu) = \{u \in L^1(p \cdot \mu) : E_p[\cosh(\alpha u) - 1] < \infty \text{ para alguna } \alpha > 0\}$$

y la norma de Luxemburgo será

$$\|u\|_{\phi_1, p} = \inf \left\{ r > 0 : E_p \left[\cosh \left(\frac{u}{r} \right) - 1 \right] \leq 1 \right\}.$$

Ahora bien, se denotará por $B_{\phi_1}(0, 1)$ (o simplemente B_{ϕ_1}) a la bola unitaria abierta (centrada en cero) en el espacio $L^{\phi_1}(p \cdot \mu)$. Así, si $u \in B_{\phi_1}(0, 1)$ entonces, $\|u\|_{\phi_1, p} < 1$ y por lo tanto existe un $\alpha > 1$ tal que $E_p[\cosh(\alpha u) - 1] \leq 1$, o equivalentemente $E_p[\cosh(\alpha u)] \leq 2$.

Por $S_{\phi_1}(0, 1)$ (o simplemente S_{ϕ_1}) se denotará la esfera unitaria sobre el espacio de funciones de Orlicz $L^{\phi_1}(p \cdot \mu)$, es decir, $u \in S_{\phi_1}$ si y solo si $\|u\|_{\phi_1, p} = 1$.

Si u, v en $L^{\phi_1}(p \cdot \mu)$ y $\epsilon > 0$ entonces, $\|u - v\|_{\phi_1, p} < \epsilon$ si y sólo si existe un $h > \frac{1}{\epsilon}$ tal que $E_p[\phi_1(h(u - v))] \leq 1$; que en términos del operador integral se escribe

$$\int_{\Omega} \phi_1(h(u - v)) p d\mu \leq 1$$

Si consideramos ahora una sucesión $\{u_n\}$ de variables aleatorias en $L^{\phi_1}(p \cdot \mu)$ entonces, $u_n \rightarrow u$ si y sólo si para todo $h > 1$ se cumple que $E_p[\phi_1(h(u_n - u))] \leq 1$, es decir,

$$\int_{\Omega} \cosh(h(u_n - u)) p d\mu \leq 2.$$

El siguiente lema será de gran utilidad más adelante.

Lema 3.1. *Sea $p \in \mathfrak{M}_{\mu}$. Para $u \in L^{\phi_1}(p \cdot \mu)$ se satisface que $\|u\|_{\phi_1, p} \leq 1$ si y sólo si $E_p[\cosh u] \leq 2$.*

Demostración

Si $u = 0$ entonces $\|u\|_{\phi_1, p} \leq 1$ y $E_p[\cosh u] = E_p[1] = 1 \leq 2$ y por tanto la propiedad se cumple para este caso. Tomemos ahora $u \neq 0$ y definamos la función real valuada

$$\begin{aligned} \alpha(\cdot) : (0, \infty) &\rightarrow \mathbb{R} \\ k &\mapsto \alpha(k) = E_p \left[\phi_1 \left(\frac{u}{k} \right) \right], \end{aligned}$$

esto para $u \in L^{\phi_1}(p \cdot \mu)$. Veamos que la función α es decreciente, para ello sean k_1, k_2 reales positivos tales que $k_1 < k_2$.

Si $u > 0$ entonces, $0 < \frac{u}{k_2} < \frac{u}{k_1}$. Como ϕ_1 es una función creciente en el intervalo $[0, \infty)$, se sigue que $\phi_1\left(\frac{u}{k_2}\right) < \phi_1\left(\frac{u}{k_1}\right)$ y como la integral de Lebesgue es monótona se obtiene $E_p\left[\phi_1\left(\frac{u}{k_2}\right)\right] < E_p\left[\phi_1\left(\frac{u}{k_1}\right)\right]$. Así, en términos de la función α se tiene que $\alpha(k_2) < \alpha(k_1)$, por lo que α es decreciente.

En el caso en que $u < 0$ resulta $0 > \frac{u}{k_2} > \frac{u}{k_1}$ como ϕ_1 es decreciente en el intervalo $(-\infty, 0]$ entonces $\phi_1\left(\frac{u}{k_2}\right) < \phi_1\left(\frac{u}{k_1}\right)$ que procediendo como en el caso anterior se obtiene $\alpha(k_2) < \alpha(k_1)$. En ambos casos se concluye que α es una función decreciente.

Como $\|u\|_{\phi_1, p} \leq 1$ y la función α es decreciente resulta

$$\alpha(\|u\|_{\phi_1, p}) \geq \alpha(1) \tag{3.1}$$

Ya que $\alpha(1) = E_p[\phi_1(u)] = E_p[\cosh u - 1] = E_p[\cosh u] - 1$. De acuerdo a (2.10) aplicada a la función de Young ϕ_1 , se tiene

$$\alpha(\|u\|_{\phi_1, p}) = E_p \left[\phi_1 \left(\frac{u}{\|u\|_{\phi_1, p}} \right) \right] = \int_{\Omega} \phi_1 \left(\frac{u}{\|u\|_{\phi_1, p}} \right) \leq 1.$$

Por transitividad con la desigualdad (3.1) se obtiene $\alpha(1) \leq 1$ y así $E_p[\cosh u] - 1 \leq 1$, esto es, $E_p[\cosh u] \leq 2$. \square

3.3. Funcional Acumulante

Se presenta ahora el estudio del funcional acumulante K_p , introducido en trabajos de Pistone [21, 5], a través del funcional generador de momentos M_p . Las pruebas de algunos resultados se consiguen en forma análoga a las de los resultados que se probarán en el capítulo 5; por tal motivo, se omitirán dichas pruebas en esta sección y se referencia la respectiva fuente bibliográfica.

Definición 3.1. Sea $p \in \mathfrak{M}_\mu$. Para cada $a \geq 1$, $n \in \mathbb{N}$ y $u \in B_{\phi_1}$, se define $\lambda_{a,n}(u)$ por

$$\begin{aligned} \lambda_{a,n}(u)(\cdot) : \quad L^{\phi_1}(p \cdot \mu) \times \dots \times L^{\phi_1}(p \cdot \mu) &\rightarrow L^a(p \cdot \mu) \\ (w_1, \dots, w_n) &\mapsto \frac{w_1}{a} \dots \frac{w_n}{a} e^{\frac{u}{a}}. \end{aligned}$$

En la definición 3.1 el operador $\lambda_{a,n}(u)$ depende de tres parámetros fijos como son: a que genera el espacio de Lebesgue $L^a(p \cdot \mu)$, u que es una variable aleatoria en la bola unidad del espacio de funciones de Orlicz y n que indica la cantidad de componentes que se asumen en el dominio.

Teorema 3.2. La función $\lambda_{a,n}(u)$ es continua, simétrica y n -multilineal, es decir,

$$\lambda_{a,n}(u) \in \mathcal{L}_s^n \left(L^{\phi_1}(p \cdot \mu), L^a(p \cdot \mu) \right).$$

La demostración de este teorema se encuentra en la tesis de Cena, [5], donde se presenta la desigualdad

$$\int_{\Omega} |w_1 w_2 \dots w_n e^{\frac{u}{a}}|^a p \, d\mu \leq \|w_1\|_{\phi_1}^a \dots \|w_n\|_{\phi_1}^a \frac{4}{r^{na}}, \quad (3.2)$$

donde cada w_i es una variable aleatoria en $L^{\phi_1}(p \cdot \mu)$, $u \in B_p$ y $r = \frac{1 - \|u\|_{\phi_1}}{n} > 0$.

Consideremos ahora $u \in B_{\phi_1}$, por las propiedades de la función exponencial como $\left|e^{\frac{u}{a}}\right| = e^{\frac{u}{a}}$ y $e^u \leq e^{|u|} \leq 2 \cosh u$ se tiene

$$\begin{aligned} \int_{\Omega} \left|e^{\frac{u}{a}}\right|^a p d\mu &= \int_{\Omega} \left(e^{\frac{u}{a}}\right)^a p d\mu = \int_{\Omega} e^u p d\mu \\ &\leq \int_{\Omega} e^{|u|} p d\mu \leq 2 \int_{\Omega} \cosh u p d\mu . \end{aligned}$$

Como $u \in B_{\phi_1}(0, 1)$ entonces $\|u\|_{\phi_1, p} < 1$, lo cual garantiza que se puede utilizar el lema 3.1 para tener que $E_p[\cosh u] \leq 2$. Por tanto, de la desigualdad anterior se sigue que

$$\int_{\Omega} \left|e^{\frac{u}{a}}\right|^a p d\mu \leq 4 .$$

en términos de la norma en el espacio de Lebesgue $L^a(p \cdot \mu)$, se tiene

$$\left\|e^{\frac{u}{a}}\right\|_{L^a} = \left(\int_{\Omega} \left|e^{\frac{u}{a}}\right|^a p d\mu\right)^{\frac{1}{a}} \leq 4^{\frac{1}{a}} < \infty .$$

En consecuencia, $\left\|e^{\frac{u}{a}}\right\|_{L^a} < \infty$ y por tanto $e^{\frac{u}{a}} \in L^a(p \cdot \mu)$ siempre que $u \in B_{\phi_1}$.

Definición 3.2. Sea $p \in \mathfrak{M}_{\mu}$, $a \geq 1$ y $u \in B_{\phi_1}(0, 1)$ fijos. Se define el operador $\lambda_{a,0}(u)$ como $\lambda_{a,0}(u) = e^{\frac{u}{a}}$.

Una consecuencia directa del teorema 3.2 es que si se hace $a = 1$ entonces el mapeo $(w_1, w_2, \dots, w_n) \mapsto E_p[w_1 w_2 \dots w_n e^u]$ es continuo. Si hacemos $a = 1$ en (3.2) resulta

$$\left|\int_{\Omega} w_1 w_2 \dots w_n e^u p d\mu\right| \leq \int_{\Omega} |w_1 w_2 \dots w_n e^u| p d\mu \leq \|w_1\|_{\phi} \dots \|w_n\|_{\phi_1} \frac{4}{r^n} .$$

Del acotamiento del operador se sigue su continuidad. En el caso en que cada w_i sea la misma variable aleatoria v entonces el mapeo $v \mapsto E_p[v^n e^u]$ es continuo.

Con base en la definición 3.1 se tiene que $\lambda_{a,n}(u)(w_1, \dots, w_n) = \frac{w_1}{a} \dots \frac{w_n}{a} e^{\frac{u}{a}}$ donde $n \in \mathbb{N}$, $a \geq 1$, $u \in B_{\phi_1}$ y $w_i \in L^{\phi_1}(p \cdot \mu)$ para $i = 1, 2, \dots, n$. Si en la definición 3.1 se toma $w_i = v$ para todo $i = 1, 2, \dots, n$ entonces,

$$\lambda_{a,n}(u)v^n = \lambda_{a,n}(u)(v, \dots, v) = \frac{v}{a} \dots \frac{v}{a} e^{\frac{u}{a}} = \left(\frac{v}{a}\right)^n e^{\frac{u}{a}} ,$$

donde $v^n = (v, v, \dots, v)$ es la notación para el lado izquierdo de la anterior ecuación. A continuación se define la función continua polinomial n -homogénea como

$$\begin{aligned} \hat{\lambda}_{a,n}(u)(\cdot) : L^{\phi_1}(p \cdot \mu) &\rightarrow L^a(p \cdot \mu) \\ v &\mapsto \hat{\lambda}_{a,n}(u)(v) = \lambda_{a,n}(u)v^n . \end{aligned} \tag{3.3}$$

En términos de la igualdad anterior y de la definición 3.2 se sigue que

$$\begin{aligned}\hat{\lambda}_{a,n}(u)(v) &= \lambda_{a,n}(u)v^n = \left(\frac{v}{a}\right)^n e^{\frac{u}{a}} \\ &= \left(\frac{v}{a}\right)^n \lambda_{a,0}(u).\end{aligned}$$

Por el teorema 3.2 se concluye que $\hat{\lambda}_{a,n}(u) \in \mathcal{P}^n(L^{\phi_1}(p \cdot \mu), L^a(p \cdot \mu))$. Al diferenciar respecto de v se determina la forma polar para $\hat{\lambda}_{a,n}(u)$ por medio de

$$\begin{aligned}D^n \hat{\lambda}_{a,n}(u)(v) &= D^n \lambda_{a,n}(u)v^n = D^n \left[\left(\frac{v}{a}\right)^n e^{\frac{u}{a}} \right] \\ &= n! \left(\frac{1}{a}\right)^n e^{\frac{u}{a}} = n! \lambda_{a,n}(u)(1).\end{aligned}$$

Además si se hacemos $u = 0$ entonces,

$$\hat{\lambda}_{a,n}(0)(v) = \lambda_{a,n}(0)v^n = \left(\frac{v}{a}\right)^n. \quad (3.4)$$

Teorema 3.3. Sean $p \in \mathfrak{M}_\mu$ y $a \geq 1$. La serie

$$A(v) = \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{v}{a}\right)^n$$

es una serie de potencias de $L^{\phi_1}(p \cdot \mu)$ en $L^a(p \cdot \mu)$, con radio de convergencia $\hat{\rho} \geq 1$.

Definición 3.3. Para cada $a \geq 1$ y $p \in \mathfrak{M}_\mu$, la función exponencial $\exp_{p,a}$ está definido por

$$\begin{aligned}\exp_{p,a} : B_{\phi_1} &\rightarrow L^a(p \cdot \mu) \\ v &\mapsto \exp_{p,a}(v) = \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{v}{a}\right)^n.\end{aligned}$$

El teorema 3.3 garantiza que la función $\exp_{p,a}$ (definida en 3.3) esté bien definida entre la bola unidad B_{ϕ_1} y el espacio de Lebesgue $L^a(p \cdot \mu)$.

Sea $v \in B_{\phi_1}$. Puntualmente se sabe que $\sum_{k=0}^{\infty} \frac{1}{k!} \left(\frac{v}{a}\right)^k$ converge a $e^{\frac{v}{a}}$, esto es, $\sum_{k=0}^{\infty} \frac{1}{k!} \left(\frac{v}{a}\right)^k = e^{\frac{v}{a}}$. Puesto que $p \cdot \mu$ es una medida finita entonces, se da la convergencia absoluta $\sum_{k=0}^{\infty} \frac{1}{k!} \left(\frac{v}{a}\right)^k \rightarrow e^{\frac{v}{a}}$. Por el teorema 3.3 se sigue que $\exp_{p,a}(\cdot)$ converge en norma $\|\cdot\|_{L^a}$, y

como esta convergencia implica convergencia en medida entonces, $\sum_{k=0}^{\infty} \frac{1}{k!} \left(\frac{v}{a}\right)^k \rightarrow \exp_{p,a}(v)$ en $p \cdot \mu$ medida. Por unicidad se concluye que

$$\exp_{p,a}(v) = e^{\frac{v}{a}}. \quad (3.5)$$

Equivalentemente $\exp_{p,a}(v) = \lambda_{a,0}(v)$, siempre que v sea una variable aleatoria en el conjunto B_{ϕ_1} .

Teorema 3.4. *La función exponencial $\exp_{p,a}(\cdot)$ satisface las siguientes propiedades*

- i. $\exp_{p,a}(0) = 1$.
- ii. $\exp_{p,a}(u + v) = \exp_{p,a}(u) \cdot \exp_{p,a}(v)$ donde $u, v \in B_{\phi_1}$ son tales que $(u + v) \in B_{\phi_1}$.
- iii. Para cada $u \in B_{\phi_1}$, $\exp_{p,a}(u)$ tiene una inversa $(\exp_{p,a}(u))^{-1}$ en $L^a(p \cdot \mu)$ dada por $(\exp_{p,a}(u))^{-1} = \exp_{p,a}(-u)$.

Demostración

- i. Por la igualdad (3.5), $\exp_{p,a}(0) = e^{\frac{0}{a}} = 1$.
- ii. Por la definición 3.3 y propiedades de las series resulta

$$\begin{aligned} \exp_{p,a}(u + v) &= \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{u+v}{a}\right)^n = \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{u}{a} + \frac{v}{a}\right)^n \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \left(\sum_{k=0}^n \binom{n}{k} \left(\frac{u}{a}\right)^{n-k} \left(\frac{v}{a}\right)^k \right) \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \left(\sum_{k=0}^n \frac{n!}{k!(n-k)!} \left(\frac{u}{a}\right)^{n-k} \left(\frac{v}{a}\right)^k \right) \\ &= \sum_{n=0}^{\infty} \sum_{k=0}^n \frac{1}{k!(n-k)!} \left(\frac{u}{a}\right)^{n-k} \left(\frac{v}{a}\right)^k. \end{aligned}$$

En la igualdad anterior se hace $i = n - k$ para tener

$$\exp_{p,a}(u + v) = \sum_{i=0}^{\infty} \sum_{k=0}^n \frac{1}{k!i!} \left(\frac{u}{a}\right)^i \left(\frac{v}{a}\right)^k = \left(\sum_{i=0}^{\infty} \frac{1}{i!} \left(\frac{u}{a}\right)^i \right) \left(\sum_{k=0}^{\infty} \frac{1}{k!} \left(\frac{v}{a}\right)^k \right).$$

Por la definición (3.3) se concluye que $\exp_{p,a}(u + v) = \exp_{p,a}(u) \cdot \exp_{p,a}(v)$.

iii. Como $u + (-u) = 0$ entonces aplicando la función $\exp_{p,a}(\cdot)$ se sigue que

$$\exp_{p,a}(u + (-u)) = \exp_{p,a}(0) = 1$$

Puesto que $u + (-u)$ es un elemento de B_{ϕ_1} entonces, se utiliza el literal *ii.* de este mismo teorema para tener $\exp_{p,a}(u) \cdot \exp_{p,a}(-u) = 1$; así $\exp_{p,a}(-u)$ es el inverso multiplicativo de $\exp_{p,a}(u)$, lo cual se escribe como $(\exp_{p,a}(u))^{-1} = \exp_{p,a}(-u)$. \square

Teorema 3.5. *Sean $p \in \mathfrak{M}_\mu$ y $a \geq 1$. El mapeo $\exp_{p,a}$ es analítico. En una vecindad de cada $u_0 \in B_{\phi_1}$ se puede expandir en una serie de Taylor como*

$$\exp_{p,a}(u_0 + v) = \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{v}{a}\right)^n e^{\frac{u_0}{a}}.$$

Si $v \in B_{\phi_1}$ y en el teorema 3.5 se hace $u_0 + v = u$, se logra demostrar que $u \in V_\epsilon(u_0)$, con esta sutitución la serie del teorema 3.5 se escribe

$$\begin{aligned} \exp_{p,a}(v) &= \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{v - u_0}{a}\right)^n e^{\frac{u_0}{a}} = \sum_{n=0}^{\infty} \frac{1}{n!} \lambda_{a,n}(u_0)(v - u_0)^n \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \hat{\lambda}_{a,n}(u_0)(v - u_0), \end{aligned}$$

que representa la serie de potencias de $L^{\phi_1}(p \cdot \mu)$ en $L^a(p \cdot \mu)$, la cual está centrada en $u_0 \in B_{\phi_1}$.

Definición 3.4. *Sea $p \in \mathfrak{M}_\mu$, el funcional generador de momentos es el mapeo*

$$\begin{aligned} M_p(\cdot) : L^{\phi_1}(p \cdot \mu) &\rightarrow \overline{\mathbb{R}^+} \\ u &\mapsto M_p(u) = E_p[e^u]. \end{aligned}$$

En el caso en que M_p se restrinja a B_{ϕ_1} entonces, $M_p(u)$ coincide con $E_p[\exp_{p,1}(u)]$. El teorema siguiente presenta las propiedades del funcional generador de momentos.

Teorema 3.6. *El funcional generador de momentos M_p satisface las siguientes propiedades*

- i.* $M_p(0) = 1$ y $M_p(u) > 1$, si $u \neq 0$ es una variable aleatoria centrada.
- ii.* M_p es convexa.
- iii.* M_p es continuo inferiormente.

iv. $B_{\phi_1} \subset \text{dom}(M_p)$, donde $\text{dom}(M_p)$ representa el dominio del funcional generador de momentos.

v. Es infinitamente Gâteaux diferenciable en el interior de su propio dominio. La n -ésima derivada en $u \in \text{dom}^\circ M_p$ en la dirección de $v \in L^{\phi_1}(p \cdot \mu)$ es

$$\frac{d^n}{dt^n} M_p(u + tv)|_{t=0} = E_p[v^n e^u].$$

vi. Es infinitamente Fréchet-diferenciable y analítica sobre B_{ϕ_1} . La n -ésima derivada en $u \in B_{\phi_1}$ evaluada en $(v_1, \dots, v_n) \in L^{\phi_1}(p \cdot \mu) \times \dots \times L^{\phi_1}(p \cdot \mu)$ es

$$D^n M_p(u)(v_1, \dots, v_n) = E_p[v_1 \dots v_n e^u].$$

En particular $DM_p(u)(\cdot) = E_p[\cdot]$.

La demostración del teorema 3.6 se puede encontrar en [5, 22] y en el capítulo 5 se generaliza el teorema anterior (ver teorema 5.5).

La función generadora de momentos de una variable aleatoria $u \in L^{\phi_1}(p \cdot \mu)$ es el mapeo $\hat{u}_p(\cdot) : \mathbb{R} \rightarrow \overline{\mathbb{R}}$ dado por

$$\hat{u}_p(t) = E_p[e^{tu}] = \int_{\Omega} e^{tu} p d\mu.$$

La clase de Cramer en p , denotada \mathcal{C}_p , es el conjunto de variables aleatorias $u \in L^{\phi_1}(p \cdot \mu)$ tales que $\hat{u}_p(t) < \infty$, para t en una vecindad de cero, es decir, la clase de Cramer la constituyen las variables aleatorias cuya función generadora de momentos es finita y se escribe

$$\mathcal{C}_p = \{u \in L^{\phi_1}(p \cdot \mu) : \hat{u}_p(t) < \infty \text{ para } t \in I \text{ con } I \text{ una vecindad del cero}\}. \quad (3.6)$$

De acuerdo con la definición de la clase de Cramer en p se siguen dos teoremas, sobre su estructura lineal y la igualdad (conjuntista) con el espacio de funciones de Orlicz $L^{\phi_1}(p \cdot \mu)$.

Teorema 3.7. Sea $p \in \mathfrak{M}_\mu$. La clase de Cramer en p es un espacio lineal convexo.

Demostración

Sea $u \in \mathcal{C}_p$ entonces, $\hat{u}_p(t) < \infty$ para todo $t \in I$, donde I es un abierto que contiene al cero. Sea α en los reales, se demuestra que $\alpha u \in \mathcal{C}_p$. Si $\alpha \neq 0$ entonces el conjunto $I' = \{\frac{t}{\alpha} : t \in I\}$ es abierto y contiene al cero. Sea $t' \in I'$, por la definición de la clase de Cramer $\hat{\alpha u}_p(t') = E_p[e^{t'(\alpha u)}]$, como $t' \in I'$ entonces existe $t \in I$ tal que $t' = \frac{t}{\alpha}$, luego

$\alpha \hat{u}_p(t') = E_p[e^{tu}] < \infty$. En particular, si $u \in \mathcal{C}_p$ entonces $2u \in \mathcal{C}_p$.

Sean $u, v \in \mathcal{C}_p$. Existen intervalos I_1 y I_2 tales que $\hat{u}_p(t) < \infty$ para $t \in I_1$ y $\hat{v}_p(t) < \infty$ para $t \in I_2$. Sea $I = I_1 \cap I_2$ entonces, I es abierto y contiene a cero. Se probará que $(u + v) \in \mathcal{C}_p$. sea $t \in I$, por la definición de la clase de Cramer

$$(u \hat{+} v)_p(t) = E_p[e^{t(u+v)}] = E_p \left[e^{\frac{1}{2}t(2u) + \frac{1}{2}t(2v)} \right] .$$

De la convexidad de la función exponencial, resulta

$$(u \hat{+} v)_p(t) \leq \frac{1}{2}E_p \left[e^{t(2u)} \right] + \frac{1}{2}E_p \left[e^{t(2v)} \right] = \frac{1}{2}(\hat{2u})_p(t) + \frac{1}{2}(\hat{2v})_p(t) .$$

Como $2u, 2v \in \mathcal{C}_p$ entonces, $(u \hat{+} v)_p(t) < \infty$ para todo $t \in I$. Se concluye así que la clase de Cramer es un espacio lineal.

Para la convexidad, sean $0 \leq \lambda \leq 1$; $u, v \in \mathcal{C}_p$ y $t \in I = I_1 \cap I_2$. Por la convexidad de la función exponencial resulta

$$E_p \left[e^{t((1-\lambda)u + \lambda v)} \right] \leq (1 - \lambda)\hat{u}_p(t) + \lambda\hat{v}_p(t) ,$$

de donde $E_p \left[e^{t((1-\lambda)u + \lambda v)} \right] < \infty$ y así $(1 - \lambda)u + \lambda v$ está en \mathcal{C}_p para cualquier $0 \leq \lambda \leq 1$. La clase de Cramer \mathcal{C}_p es un espacio lineal convexo. \square

El siguiente teorema establece que la clase de Cramer \mathcal{C}_p es conjuntistamente igual al espacio de Orlicz $L^{\phi_1}(p \cdot \mu)$, demostración que se encuentra en [21] y que en el capítulo 5 se presenta para el espacio $L^{\phi_k}(p \cdot \mu)$.

Teorema 3.8. *Para $p \in \mathfrak{M}_\mu$, se verifica que $\mathcal{C}_p = L^{\phi_1}(p \cdot \mu)$ conjuntistamente.*

Si la función generadora de momentos existe entonces, los momentos de todos los órdenes existen, por lo que si $\hat{u}_p(t) < \infty$ entonces, el primer momento o esperanza existe. Así, en la clase de Cramer de p se distingue el subconjunto de variables aleatorias cuya esperanza es cero, se denota por B_p , esto es,

$$B_p := \{u \in \mathcal{C}_p : E_p[u] = 0\} = \{u \in L^{\phi_1}(p \cdot \mu) : E_p[u] = 0\} . \quad (3.7)$$

A dicho conjunto se le llama clase de Cramer centrada en $L^{\phi_1}(p \cdot \mu)$. En el siguiente teorema se demuestra que el espacio de variables aleatorias centradas es cerrado en $L^{\phi_1}(p \cdot \mu)$.

Teorema 3.9. *El espacio B_p es un subespacio convexo y cerrado de $L^{\phi_1}(p \cdot \mu)$.*

Demostración

Sean u, v en B_p y α, β reales. Debido a que $E_p[u] = E_p[v] = 0$ entonces

$$E_p[\alpha u + \beta v] = \alpha E_p[u] + \beta E_p[v] = 0 ,$$

por lo que $\alpha u + \beta v \in B_p$ y así B_p es un subespacio de $L^{\phi_1}(p \cdot \mu)$. B_p es un subespacio convexo, ya que si $0 \leq \lambda \leq 1$ entonces, $E_p[(1 - \lambda)u + \lambda v] = (1 - \lambda)E_p[u] + \lambda E_p[v] = 0$. Veamos ahora que B_p es cerrado, para ello tomemos una sucesión $\{u_n\}_{n \in \mathbb{N}}$ en B_p , esto es, $E_p[u_n] = 0$ para todo $n \in \mathbb{N}$. Supongamos ahora que dicha sucesión converge en medida $p \cdot \mu$ a la variable aleatoria u , escribimos $u_n \rightarrow u$, y se debe concluir que $u \in B_p$. Supongamos que $u \geq 0$, por lo que $E_p[u] \geq 0$ y en consecuencia existe una subsucesión u_{n_k} de u_n tal que $u_{n_k} \rightarrow u$. Por el lema de Fatou

$$E_p[u] = E_p \left[\lim_{k \rightarrow \infty} u_{n_k} \right] \leq \liminf_{k \rightarrow \infty} E_p[u_{n_k}] ,$$

así $E_p[u] \leq 0$, y por antisimetría en los reales se tiene $E_p[u] = 0$. \square

De acuerdo con el teorema 3.9 se sigue que B_p es un subespacio cerrado del espacio de Banach $L^{\phi_1}(p \cdot \mu)$ por lo que en sí mismo B_p tiene la estructura de un espacio de Banach. En éste se denotará la bola unitaria por \mathcal{V}_p , es decir,

$$\mathcal{V}_p = \{u \in B_p : \|u\|_{\phi_1, p} < 1\} .$$

Por la definición de \mathcal{V}_p se sigue que éste es un conjunto abierto y se presentan las inclusiones conjuntistas $\mathcal{V}_p \subset B_p \subset \mathcal{C}_p = L^{\phi_1}(p \cdot \mu) \subset L^1(p \cdot \mu)$.

Definición 3.5. *Dado $p \in \mathfrak{M}_\mu$, se dice que f es un modelo exponencial unidimensional si existe $u \in \mathcal{C}_p$ y una función real valuada ϕ tal que para todo $t \in I$ con I un abierto que contiene al cero, se cumple*

$$f(t) = e^{tu - \phi(t)} p .$$

En tal caso se dice que f es un modelo exponencial paramétrico con parámetro t , y además se escribe $f(\cdot) : I \rightarrow \mathfrak{M}_\mu$.

Definición 3.6. *Se dice que dos densidades de probabilidad p y z en \mathfrak{M}_μ están relacionadas por un modelo exponencial unidimensional si existen $r \in \mathfrak{M}_\mu$, $u \in \mathcal{C}_p$, una función ϕ real valuada y $\lambda > 0$ tal que $\forall t \in (-\lambda, \lambda)$ la función*

$$f(t) = e^{ut - \phi(t)} r ,$$

cumple que existen t_0 y t_1 en $(-\lambda, \lambda)$ tales que $f(t_0) = p$ y $f(t_1) = z$.

Dados p y z en \mathfrak{M}_μ , se dice que p está relacionado con z , se escribe $p \sim z$ si y sólo si p y z están conectados por un modelo exponencial paramétrico, de acuerdo con la definición 3.6. La relación \sim es de equivalencia y por tanto particiona a \mathfrak{M}_μ en clases de equivalencia, donde dos densidades están en la misma clase si están conectados por un modelo exponencial.

Teorema 3.10. *Sean p y z dos densidades de probabilidad en \mathfrak{M}_μ relacionadas por un modelo exponencial entonces, $L^{\phi_1}(p \cdot \mu) = L^{\phi_1}(z \cdot \mu)$.*

En “*Geometric Structures on the non-parametric statistical manifold*” [5] se demuestra que además de ser iguales conjuntistamente, los espacios $L^\phi(p \cdot \mu)$ y $L^\phi(z \cdot \mu)$ (con p y z conectados exponencialmente), el mapeo identidad

$$Id : (L^{\phi_1}(p \cdot \mu), \|\cdot\|_{\phi_1,p}) \rightarrow (L^{\phi_1}(z \cdot \mu), \|\cdot\|_{\phi_1,z})$$

es un homeomorfismo.

Teorema 3.11. *Sean $p \in \mathfrak{M}_\mu$ y $q = \frac{e^u}{E_p[e^u]}p$ con $u \in \mathcal{V}_p$.*

1. *La variable aleatoria e^u es $p \cdot \mu$ integrable.*
2. *$q \in \mathfrak{M}_\mu$.*
3. *$L^{\phi_1}(p \cdot \mu) = L^{\phi_1}(q \cdot \mu)$.*

Demostración

1. Por hipótesis $u \in \mathcal{V}_p$ entonces, $\|u\|_{\phi_1,p} < 1$ y existe $\alpha > 1$ tal que $E_p[\cosh(\alpha u)] \leq 2$, equivalente a $E_p[e^{\alpha u} + e^{-\alpha u}] \leq 4$. De lo anterior se desprende que $E_p[e^{\alpha u}] \leq 4$ y $E_p[e^{-\alpha u}] \leq 4$.

Si $u > 0$ entonces, al ser $\alpha > 1$, se sigue que $\alpha u > u$ y al ser la función exponencial creciente entonces, $e^{\alpha u} > e^u$. Por ser el operador $E_p[\cdot]$ monótono entonces, $E_p[e^{\alpha u}] > E_p[e^u]$, dado que $E_p[e^{\alpha u}] < 4$ se tiene $E_p[e^u] < 4$. Por lo tanto e^u es $p \cdot \mu$ integrable.

Si por el contrario suponemos que $u < 0$ entonces, $-u > 0$ y así $-\alpha u > -u > 0 > u$. Análogamente con el caso anterior $E_p[e^{-\alpha u}] > E_p[e^u]$ y así $E_p[e^u] < 4$ ya que $E_p[e^{-\alpha u}] \leq 4$. Luego, e^u es integrable para $u < 0$. Con esto se concluye que e^u es $p \cdot \mu$ integrable para todo $u \in \mathcal{V}_p$.

2. Ahora bien, debido a que $q = \frac{e^u}{E_p[e^u]}p$ entonces por el literal anterior el término $E_p[e^u]$ está bien definido y es finito. Ya que $E_p[e^u]$ es una constante respecto de la medida $p \cdot \mu$, así

$$\begin{aligned} \int_{\Omega} q d\mu &= \int_{\Omega} \frac{e^u}{E_p[e^u]} p d\mu = \frac{1}{E_p[e^u]} \int_{\Omega} e^u p d\mu \\ &= \frac{1}{E_p[e^u]} E_p[e^u] = 1. \end{aligned}$$

Como además $q > 0$, por ser $p > 0$, entonces se concluye $q \in \mathfrak{M}_{\mu}$, es decir, q es una densidad de probabilidad.

3. Puesto que $E_p[e^u] > 0$ entonces, la expresión $E_p[e^u] = e^{\ln E_p[e^u]}$ está bien definida, así

$$q = \frac{e^u}{E_p[e^u]} p = \frac{e^u}{e^{\ln E_p[e^u]}} p = e^{u - \ln E_p[e^u]} p, \quad (3.8)$$

lo que garantiza que q y p están conectados por un modelo exponencial unidimensional no paramétrico. Y por el teorema 3.10 se sigue que $L^{\phi_1}(p \cdot \mu) = L^{\phi_1}(q \cdot \mu)$. \square

De acuerdo con el teorema 3.11 se tiene que si $q = e^{u - \ln(E_p[e^u])} p$ y $u \in \mathcal{V}_p \subset L^{\phi_1}(p \cdot \mu)$ entonces u también es una variable aleatoria en $L^{\phi_1}(q \cdot \mu)$. El argumento $u - \ln(E_p[e^u])$ de la función exponencial definida en la expresión 3.8 permite definir las cartas que sirven para modelar la variedad en \mathfrak{M}_{μ} .

Definición 3.7. Sea $p \in \mathfrak{M}_{\mu}$ se define el mapeo acumulante como

$$\begin{aligned} K_p(\cdot) : B_p &\rightarrow \overline{\mathbb{R}^+} \\ u &\mapsto K_p(u) = \ln(M_p(u)) = \ln(E_p[e^u]). \end{aligned}$$

De acuerdo con la definición 3.7 se tiene que

$$\text{dom}(K_p) = B_p \cap \text{dom}(M_p). \quad (3.9)$$

Por otro lado, $K_p(u)$ es un número real y al aplicar el funcional $E_p[\cdot]$ a dicha función, resulta $E_p[K_p(u)] = K_p(u)$. En el siguiente teorema se ilustran las propiedades del funcional acumulante (para las demostraciones ver [21]), algunas de las cuales se heredan de la función generadora de momentos.

Teorema 3.12. *El funcional acumulante K_p satisface:*

- i. $\mathcal{V}_p \subset \text{dom}(K_p)$.
- ii. $K_p(0) = 0$ y $K_p(u) > 0$ si $u \neq 0$ es centrada.
- iii. K_p es analítico en $B_p \cap B_{\phi_1}$.
- iv. K_p es convexo.
- v. es infinitamente Fréchet-diferenciable sobre \mathcal{V}_p .
- vi. para todo $u \in \mathcal{V}_p$, $q = e^{u-K_p(u)}p$ es una densidad de probabilidad en \mathfrak{M}_μ . El valor de la n -ésima derivada en u en la dirección $v \in B_p$ de K_p es

$$D^n K_p(u) \cdot v^n = \frac{d^n}{dt^n} \ln (E_q [e^{tv}]) |_{t=0} .$$

Como un caso particular del teorema tenemos el siguiente corolario el cual caracteriza el diferencial del acumulante en términos de $v \in B_p$.

Corolario 3.1. *Si $v \in B_p$ y $q = e^{u-K_p(u)}p$ entonces $DK_p(u) \cdot v = E_q[v]$.*

Demostración

Para $p \in \mathfrak{M}_\mu$ consideremos $q = e^{u-K_p(u)}p$. De acuerdo con la definición 3.3 se escribe $e^{tv} = \sum_{k=0}^{\infty} \frac{1}{k!} t^k v^k$ y al aplicar el operador $E_q[\cdot]$, con t constante respecto de la medida $p \cdot \mu$, resulta $E_q[e^{tv}] = \sum_{k=0}^{\infty} \frac{E_q[v^k]}{k!} t^k$. Al derivar respecto de t se tiene

$$\frac{d}{dt} E_q[e^{tv}] = E_q[v] + E_q[v^2]t + \frac{E_q[v^3]}{2!} t^2 + \frac{E_q[v^4]}{3!} t^3 + \dots = \sum_{k=1}^{\infty} \frac{E_q[v^k]}{(k-1)!} t^{k-1} ,$$

evaluando en $t = 0$ resulta $\frac{d}{dt} E_q[e^{tv}]|_{t=0} = E_q[v]$. De acuerdo con el teorema 3.12 tenemos

$$D^n K_p(u) \cdot v^n = \frac{d^n}{dt^n} \ln (E_q [e^{tv}]) |_{t=0} .$$

Para $n = 1$, resulta

$$DK_p(u) \cdot v = \frac{d}{dt} \ln (E_q[e^{tv}])|_{t=0} = \left(\frac{1}{E_q[e^{tv}]} \frac{d}{dt} E_q[e^{tv}] \right)_{t=0}$$

$$= \frac{1}{E_q[e^{tv}]}\Big|_{t=0} E_q[v] = \frac{E_q[v]}{E_q[1]}.$$

Como $E_q[1] = 1$ entonces se concluye que $DK_p(u) \cdot v = E_q[v]$. \square

Sea $p \in \mathfrak{M}_\mu$, si $q = e^{u-K_p(u)}p$ entonces $L^{\phi_1}(p \cdot \mu) = L^{\phi_1}(q \cdot \mu)$, por lo tanto si $u \in L^{\phi_1}(p \cdot \mu)$ entonces $u \in L^{\phi_1}(q \cdot \mu)$. De acuerdo con la elección de q se puede escribir como $q = \frac{e^u}{M_p(u)}p$, así para $v \in B_p$ se tiene

$$E_q[v] = \int_{\Omega} vq d\mu = \int_{\Omega} \frac{ve^u}{M_p(u)}p d\mu = \frac{1}{M_p(u)}E_p[ve^u].$$

Debido a que $DK_p(u) \cdot v = E_q[v]$ y $DM_p(u) \cdot v = E_p[ve^u]$ entonces en la expresión anterior se sigue que

$$DK_p(u) \cdot v = \frac{DM_p(u) \cdot v}{M_p(u)}. \quad (3.10)$$

3.4. Variedad de Información

De acuerdo con la definición 3.7 es posible escribir la igualdad (3.8) en términos de la función acumulante como $q = e^{u-K_p(u)}p$ siempre que $u \in B_p$. Esto permite definir el siguiente mapeo que induce las cartas sobre el espacio \mathfrak{M}_μ .

Definición 3.8. *Sea $p \in \mathfrak{M}_\mu$ se define el mapeo*

$$\begin{aligned} e_p(\cdot) : \mathcal{V}_p &\rightarrow \mathfrak{M}_\mu \\ u &\mapsto q = e_p(u) \end{aligned}$$

dado por $q = e_p(u) = e^{u-K_p(u)}p$.

En la definición 3.8 se tiene que $u \in \mathcal{V}_p$ lo cual garantiza $E_p[u] = 0$ y que se satisfaga el teorema 3.11 y así $q \in \mathfrak{M}_\mu$. Se denotará la imagen de \mathcal{V}_p respecto a $e_p(\cdot)$ como \mathcal{U}_p , es decir, $\mathcal{U}_p = e_p(\mathcal{V}_p)$ con lo que $e_p(\cdot) : \mathcal{V}_p \rightarrow \mathcal{U}_p$. Si tomamos $u = 0$ casi en todas partes de Ω entonces $u \in \mathcal{V}_p$ y $K_p(0) = \ln(E_p[e^0]) = 0$ por lo que $e_p(0) = e^{0-K_p(0)}p = p$, de allí que $p \in \mathcal{U}_p$. Luego p y q están conectados por un modelo exponencial no-paramétrico.

Teorema 3.13. *Sea $p \in \mathfrak{M}_\mu$. El mapeo $e_p(\cdot)$ es inyectivo.*

Demostración

Sean u_1 y u_2 variables aleatorias en \mathcal{V}_p tales que $e_p(u_1) = e_p(u_2)$ y probemos que $u_1 = u_2$. Debido a que $\mathcal{V}_p \subset B_p$ se tiene que $E_p[u_1] = E_p[u_2] = 0$, como $e_p(u_1) = e_p(u_2)$, por la definición 3.8 se tiene

$$e^{u_1 - K_p(u_1)} p = e^{u_2 - K_p(u_2)} p .$$

Como $p > 0$ por estar en \mathfrak{M}_μ , de la anterior igualdad se obtiene que $e^{u_1 - K_p(u_1)} = e^{u_2 - K_p(u_2)}$, debido a que la función exponencial es inyectiva entonces

$$u_1 - K_p(u_1) = u_2 - K_p(u_2) .$$

Aplicando $E_p[\cdot]$ resulta $E_p[u_1] - E_p[K_p(u_1)] = E_p[u_2] - E_p[K_p(u_2)]$, como el acumulante es constante respecto de la medida $p \cdot \mu$ y $E_p[u_1] = E_p[u_2] = 0$ entonces $K_p(u_1) = K_p(u_2)$. Ya que $u_1 - K_p(u_1) = u_2 - K_p(u_2)$ y $K_p(u_1) = K_p(u_2)$ entonces $u_1 = u_2$. El mapeo $e_p(\cdot)$ es inyectivo. \square

Sea $p \in \mathfrak{M}_\mu$, existe $q \in \mathcal{U}_p$ tal que $q = e_p(u)$ con $u \in \mathcal{V}_p$. Procedamos ahora a despejar u en función de q . De acuerdo con la definición del mapeo $e_p(\cdot)$ dado en (3.8) tenemos $q = e^{u - K_p(u)} p$ puesto que $p > 0$ y $q > 0$ entonces $\ln\left(\frac{q}{p}\right)$ está bien definido y está dado por

$$u - K_p(u) = \ln\left(\frac{q}{p}\right) . \quad (3.11)$$

Como $u \in \mathcal{V}_p \subset L^{\phi_1}(p \cdot \mu)$ y $K_p(u)$ es un número real entonces $(u - K_p(u)) \in L^{\phi_1}(p \cdot \mu)$, por tanto $\ln\left(\frac{q}{p}\right) \in L^{\phi_1}(p \cdot \mu)$, es decir, el logaritmo de $\frac{q}{p}$ es $p \cdot \mu$ integrable. Puesto que p y q están conectados por un modelo exponencial unidimensional no paramétrico entonces $L^{\phi_1}(p \cdot \mu) = L^{\phi_1}(q \cdot \mu)$ por el teorema 3.10 y así $\ln\left(\frac{q}{p}\right)$ es también $q \cdot \mu$ integrable. Ahora bien, aplicando el operador integral $E_p[\cdot]$ en (3.11) se tiene

$$E_p[u] - E_p[K_p(u)] = E_p\left[\ln\left(\frac{q}{p}\right)\right] .$$

Como K_p es constante respecto de la medida $p \cdot \mu$ y $E_p[u] = 0$ entonces, $-K_p(u) = E_p\left[\ln\left(\frac{q}{p}\right)\right]$. Al sustituir en (3.11) se tiene

$$u = \ln\left(\frac{q}{p}\right) - E_p\left[\ln\left(\frac{q}{p}\right)\right] . \quad (3.12)$$

Debido a que $u \in \mathcal{V}_p$ entonces por la igualdad (3.12) resulta

$$\left(\ln\left(\frac{q}{p}\right) - E_p\left[\ln\left(\frac{q}{p}\right)\right]\right) \in \mathcal{V}_p ,$$

con base en esta misma expresión el mapeo s_p definido a continuación está bien definido.

Definición 3.9. Sea $p \in \mathfrak{M}_\mu$ se define el mapeo

$$\begin{aligned} s_p(\cdot) : \mathcal{U}_p &\rightarrow \mathcal{V}_p \\ q &\mapsto u = s_p(q) \end{aligned}$$

dado por $s_p(q) = \ln\left(\frac{q}{p}\right) - E_p\left[\ln\left(\frac{q}{p}\right)\right]$.

El mapeo $e_p(\cdot)$ es inyectivo de acuerdo con el teorema 3.13, además $e_p(\mathcal{V}_p) = \mathcal{U}_p$, entonces e_p es sobreyectivo y por tanto posee una inversa, $e_p^{-1}(\cdot)$. Se demuestra en el siguiente teorema que $e_p^{-1} = s_p$.

Teorema 3.14. Sea $p \in \mathfrak{M}_\mu$. El mapeo $s_p(\cdot)$ es el inverso del mapeo $e_p(\cdot)$

Demostración

Sea $p \in \mathfrak{M}_\mu$. Para lograr el objetivo de la demostración debe garantizarse que $(s_p \circ e_p)(u) = u$ y $(e_p \circ s_p)(q) = q$ para $q \in \mathcal{U}_p$ y $u \in \mathcal{V}_p$. Por la definición de la función compuesta y por la definición del mapeo s_p en (3.9), resulta

$$(s_p \circ e_p)(u) = s_p(e_p(u)) = \ln\left(\frac{e_p(u)}{p}\right) - E_p\left[\ln\left(\frac{e_p(u)}{p}\right)\right].$$

Como $e_p(u) = e^{u-K_p(u)}p$ y $p > 0$ entonces $\frac{e_p(u)}{p} = e^{u-K_p(u)}$ y por tanto $\ln\left(\frac{e_p(u)}{p}\right) = u - K_p(u)$. Ya que $E_p[u] = 0$ por ser $u \in \mathcal{V}_p$,

$$\begin{aligned} (s_p \circ e_p)(u) &= u - K_p(u) - E_p[u - K_p(u)] \\ &= u - K_p(u) - E_p[u] + E_p[K_p(u)] \\ &= u - K_p(u) + K_p(u) = u. \end{aligned}$$

Veamos ahora que si $q \in \mathcal{U}_p$ entonces $(e_p \circ s_p)(q) = q$. Como $q \in \mathcal{U}_p$ entonces $q \in \mathfrak{M}_\mu$, por lo que $E_q[1] = 1$ y se verifica

$$E_p\left[\frac{q}{p}\right] = \int_{\Omega} \frac{q}{p} p d\mu = \int_{\Omega} q d\mu = 1.$$

Ahora, de acuerdo con la definición de la función $s_p(\cdot)$ aplicada a q tenemos

$$e^{s_p(q)} = e^{\ln\left(\frac{q}{p}\right) - E_p\left[\ln\left(\frac{q}{p}\right)\right]} = e^{\ln\left(\frac{q}{p}\right)} e^{-E_p\left[\ln\left(\frac{q}{p}\right)\right]} = \frac{q}{p e^{E_p\left[\ln\left(\frac{q}{p}\right)\right]}}.$$

Por la definición de la función acumulante $K_p(u) = \ln(E_p[e^u])$ aplicada a la $s_p(q)$ resulta

$$\begin{aligned} K_p(s_p(q)) &= \ln \left(E_p \left[e^{s_p(q)} \right] \right) = \ln \left(E_p \left[e^{\ln \left(\frac{q}{p} \right) - E_p \left[\ln \left(\frac{q}{p} \right) \right]} \right] \right) \\ &= \ln \left(E_p \left[\frac{q}{p} e^{-E_p \left[\ln \left(\frac{q}{p} \right) \right]} \right] \right) = \ln \left(e^{-E_p \left[\ln \left(\frac{q}{p} \right) \right]} E_p \left[\frac{q}{p} \right] \right). \end{aligned}$$

En este último paso se hizo uso de que $e^{-E_p \left[\ln \left(\frac{q}{p} \right) \right]}$ es un número real y por lo tanto es una constante respecto del operador $E_p[\cdot]$. Ya que $E_p \left[\frac{q}{p} \right] = 1$, la ecuación anterior es equivalente a $K_p(s_p(q)) = \ln \left(e^{-E_p \left[\ln \left(\frac{q}{p} \right) \right]} \right)$, que aplicando el operador exponencial a ambos lados resulta

$$e^{-K_p(s_p(q))} = e^{-\ln \left(e^{-E_p \left[\ln \left(\frac{q}{p} \right) \right]} \right)} = \frac{1}{e^{-E_p \left[\ln \left(\frac{q}{p} \right) \right]}} = e^{E_p \left[\ln \left(\frac{q}{p} \right) \right]}.$$

Con base en la definición de la función compuesta se escribe $(e_p \circ s_p)(q) = e_p(s_p(q))$, que de acuerdo con la definición 3.8 respecto de la función e_p resulta

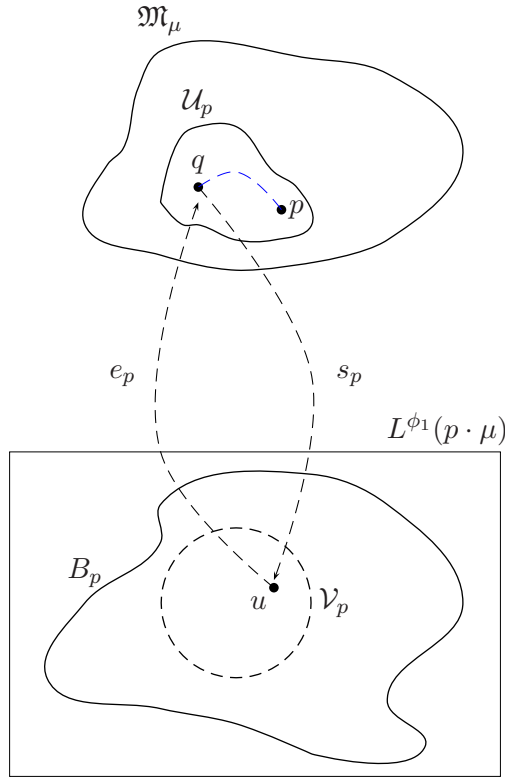
$$\begin{aligned} (e_p \circ s_p)(q) &= e^{s_p(q) - K_p(s_p(q))} p = e^{s_p(q)} e^{-K_p(s_p(q))} p = \\ &= (e_p \circ s_p)(q) = \frac{q}{p e^{E_p \left[\ln \left(\frac{q}{p} \right) \right]}} e^{E_p \left[\ln \left(\frac{q}{p} \right) \right]} p = q. \end{aligned}$$

De acuerdo con los resultados encontrados $(e_p \circ s_p)(q) = q$ y $(s_p \circ e_p)(u) = u$ puede concluirse que s_p es el mapeo inverso del mapeo e_p , lo cual se escribe como $s_p^{-1} = e_p \circ s_p = e_p^{-1}$. \square

Por el teorema 3.14, la familia de pares (\mathcal{U}_p, s_p) con $p \in \mathfrak{M}_\mu$ es una carta en \mathfrak{M}_μ , ya que la unión de los conjuntos \mathcal{U}_p cubre a \mathfrak{M}_μ , en efecto, si $q \in \bigcup_{p \in \mathfrak{M}_\mu} \mathcal{U}_p$ entonces q pertenece a alguno

de los conjuntos \mathcal{U}_p por lo que $q = e^{u - K_p(u)} p$ y así por el teorema 3.11 $q \in \mathfrak{M}_\mu$ presentándose la inclusión $\bigcup_{p \in \mathfrak{M}_\mu} \mathcal{U}_p \subset \mathfrak{M}_\mu$. En el caso en que $q \in \mathfrak{M}_\mu$ entonces $q \in \mathcal{U}_q \subset \bigcup_{p \in \mathfrak{M}_\mu} \mathcal{U}_p$, de donde

$\mathfrak{M}_\mu \subset \bigcup_{p \in \mathfrak{M}_\mu} \mathcal{U}_p$ y se presenta la igualdad $\mathfrak{M}_\mu = \bigcup_{p \in \mathfrak{M}_\mu} \mathcal{U}_p$; es decir, los conjuntos \mathcal{U}_p cubren a \mathfrak{M}_μ . A continuación se hace una ilustración de la situación



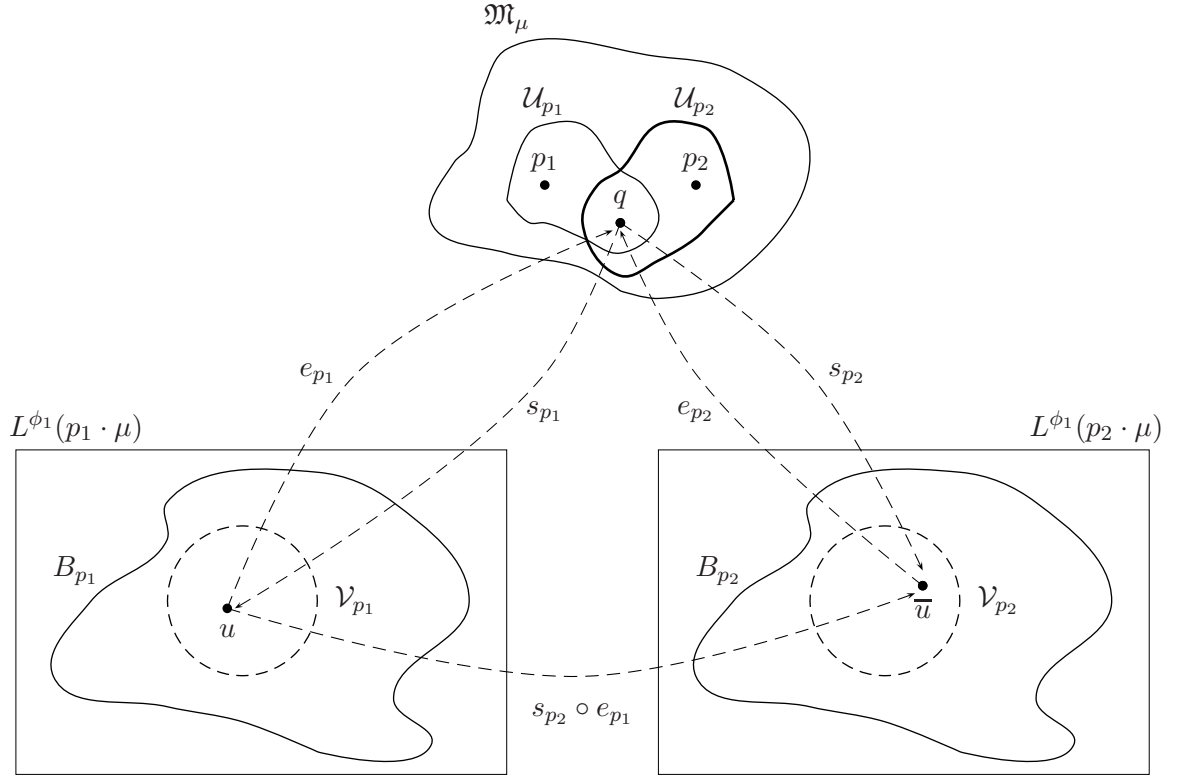
Consideremos ahora dos densidades de probabilidad en \mathfrak{M}_μ denotados por p_1 y p_2 donde $\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2} \neq \emptyset$. En el siguiente gráfico se ilustran los mapeos transición o cambios de coordenadas, donde $q \in (\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2})$ es otra densidad de probabilidad.

En este caso, el cambio de coordenadas es la composición $s_{p_2} \circ e_{p_1}$ y su inversa $s_{p_1} \circ e_{p_2}$. Sea $u \in \mathcal{V}_{p_1}$ tal que $e_{p_1}(u) = q$ y $\bar{u} \in \mathcal{V}_{p_2}$ tal que $s_{p_2}(q) = \bar{u}$ así $(s_{p_2} \circ e_{p_1})(u) = \bar{u}$. Como $(s_{p_2} \circ e_{p_1})(u) = s_{p_2}(e_{p_1}(u))$ entonces, por las definiciones de las funciones s_p y e_p dadas en (3.9) y (3.8), se tiene

$$\begin{aligned}
 (s_{p_2} \circ e_{p_1})(u) &= s_{p_2}(e_{p_1}(u)) = \ln \left(\frac{e_{p_1}(u)}{p_2} \right) - E_{p_2} \left[\ln \left(\frac{e_{p_1}(u)}{p_2} \right) \right] \\
 &= \ln \left(\frac{e^{u-K_{p_1}(u)} p_1}{p_2} \right) - E_{p_2} \left[\ln \left(\frac{e^{u-K_{p_1}(u)} p_1}{p_2} \right) \right] \\
 &= \ln \left(e^{u-K_{p_1}(u)} \right) + \ln \left(\frac{p_1}{p_2} \right) - E_{p_2} \left[\ln \left(e^{u-K_{p_1}(u)} \right) + \ln \left(\frac{p_1}{p_2} \right) \right] \\
 &= u - K_{p_1}(u) + \ln \left(\frac{p_1}{p_2} \right) - E_{p_2} \left[u - K_{p_1}(u) + \ln \left(\frac{p_1}{p_2} \right) \right].
 \end{aligned}$$

Como el acumulante es constante respecto de la medida $p \cdot \mu$ se sigue que

$$(s_{p_2} \circ e_{p_1})(u) = u + \ln \left(\frac{p_1}{p_2} \right) - E_{p_2} \left[u + \ln \left(\frac{p_1}{p_2} \right) \right]. \quad (3.13)$$



De acuerdo con la expresión (3.13) es posible definir los mapeos transiciones como

$$(s_{p_2} \circ e_{p_1})(\cdot) : s_{p_1}(\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2}) \rightarrow s_{p_2}(\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2})$$

$$u \mapsto \bar{u} = u + \ln \left(\frac{p_1}{p_2} \right) - E_{p_2} \left[u + \ln \left(\frac{p_1}{p_2} \right) \right].$$

Por (3.13), los mapeos transición son funciones afines, es decir, difieren de una función lineal en una constante dada por $\ln \left(\frac{p_1}{p_2} \right) - E_{p_2} \left[u + \ln \left(\frac{p_1}{p_2} \right) \right]$ y es por ello que dichos mapeos son continuos y de clase C^∞ . El siguiente teorema tomado del artículo de [9] presenta la continuidad de los espacios $s_{p_i}(\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2})$ con $i = 1, 2$ a través de las propiedades topológicas inducidas por la norma de Luxemburgo.

Teorema 3.15. *Sean p_1, p_2 densidades en \mathfrak{M}_μ . El conjunto $s_{p_2}(\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2})$ es abierto en la topología de B_{p_1} .*

Demostración

Sea $q \in (\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2})$, de allí que $q \in \mathcal{U}_{p_1}$ luego existe una variable aleatoria $u \in \mathcal{V}_{p_1}$ tal que $q = e_{p_1}(u) = e^{-u - K_{p_1}(u)} p_1$. Sea $\bar{u} = s_{p_2}(q) = s_{p_2}(e_{p_1}(u))$ que de acuerdo con (3.13) se sigue

$$s_{p_2}(q) = u + \ln \left(\frac{p_1}{p_2} \right) - E_{p_2} \left[u + \ln \left(\frac{p_1}{p_2} \right) \right]. \quad (1)$$

Como \bar{u} es un elemento en \mathcal{V}_{p_2} entonces la norma en el espacio de funciones de Orlicz $L^{\phi_1}(p_2 \cdot \mu)$ es menor que 1, $\|\bar{u}\|_{\phi_1, p_2} < 1$ (2). Consideremos ahora una bola de radio r entorno de la variable aleatoria $u \in s_{p_1}(\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2}) \subset \mathcal{V}_{p_1}$ en la topología de B_{p_1} , así

$$A_r = \{v \in B_{p_1} : \|u - v\|_{\phi_1, p_1} < r\}. \quad (3)$$

Con base en la construcción, A_r es un conjunto abierto en la topología de B_{p_1} . Se hace r lo suficientemente pequeño para que $A_r \subset \mathcal{V}_{p_1}$. Sean $v \in A_r$ y $\tilde{q} = e_{p_1}(v) = e^{v - K_{p_1}(v)} p_1$. Bajo la consideración de que r sea lo suficientemente pequeño se tiene $\tilde{q} \in (\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2})$, por lo que $s_{p_2}(\tilde{q})$ existe; con base en la composición de funciones $s_{p_2}(\tilde{q}) = s_{p_2}(e_{p_1}(v))$ y por el mapeo transición

$$s_{p_2}(\tilde{q}) = v + \ln\left(\frac{p_1}{p_2}\right) - E_{p_2}\left[v + \ln\left(\frac{p_1}{p_2}\right)\right]. \quad (4)$$

Sumando y restando los términos u y $E_{p_2}\left[u + \ln\left(\frac{p_1}{p_2}\right)\right]$ en el lado derecho de la igualdad (4) y asociando adecuadamente se tienen las igualdades

$$\begin{aligned} s_{p_2}(\tilde{q}) &= v + \ln\left(\frac{p_1}{p_2}\right) - E_{p_2}\left[v + \ln\left(\frac{p_1}{p_2}\right)\right] + (u - u) + \left(E_{p_2}\left[u + \ln\left(\frac{p_1}{p_2}\right)\right] - E_{p_2}\left[u + \ln\left(\frac{p_1}{p_2}\right)\right]\right) \\ &= (v - u) + \left[u + \ln\left(\frac{p_1}{p_2}\right) - E_{p_2}\left[u + \ln\left(\frac{p_1}{p_2}\right)\right]\right] + E_{p_2}[u - v] \\ &= (v - u) + \bar{u} + E_{p_2}[u - v]. \end{aligned}$$

Ya que $s_{p_2}(\tilde{q})$ es un elemento en $L^{\phi_1}(p_2 \cdot \mu)$ entonces, de acuerdo con la desigualdad triangular, respecto de la norma de Luxemburgo, resulta

$$\|s_{p_2}(\tilde{q})\|_{\phi_1, p_2} \leq \|v - u\|_{\phi_1, p_2} + \|\bar{u}\|_{\phi_1, p_2} + \|E_{p_2}[u - v]\|_{\phi_1, p_2}. \quad (5)$$

Como $E_{p_2}[u - v]$ es una constante entonces $\|E_{p_2}[u - v]\|_{\phi_1, p_2} = |E_{p_2}[u - v]| \cdot \|1\|_{\phi_1, p_2}$. Como $|E_{p_2}[u - v]| \leq E_{p_2}[|u - v|]$ resulta la desigualdad $\|E_{p_2}[u - v]\|_{\phi_1, p_2} \leq K E_{p_2}[|u - v|]$ donde $K = \|1\|_{\phi_1, p_2}$. El número $E_{p_2}[|u - v|]$ es precisamente la norma de $u - v$ en el espacio de funciones de Lebesgue L^1 bajo la medida de densidad $p_2 \cdot \mu$, así

$$\|E_{p_2}[u - v]\|_{\phi_1, p_2} \leq K \|v - u\|_{L^1(p_2 \cdot \mu)}. \quad (6)$$

Debido que $L^{\phi_1}(p_2 \cdot \mu) \subset L^1(p_2 \cdot \mu)$, existe una constante C_1 positiva tal que $\|\cdot\|_{L^1(p_2 \cdot \mu)} \leq C_1 \|\cdot\|_{\phi_1, p_2}$; por lo que en la desigualdad (6) se escribe

$$\|E_{p_2}[u - v]\|_{\phi_1, p_2} \leq K C_1 \|v - u\|_{\phi_1, p_2}. \quad (7)$$

Puesto que $q \in \mathcal{U}_{p_1}$ entonces q está conectado con p_1 por un modelo exponencial, por lo que se presenta la igualdad entre los espacios de Orlicz (ver (3.10)) asociados $L^{\phi_1}(p_1 \cdot \mu) =$

$L^{\phi_1}(q \cdot \mu)$; por un razonamiento similar $L^{\phi_1}(p_2 \cdot \mu) = L^{\phi_1}(q \cdot \mu)$ por lo que se presenta la igualdad entre los espacios $L^{\phi_1}(p_1 \cdot \mu) = L^{\phi_1}(p_2 \cdot \mu)$ y por ende la equivalencia entre las normas, es decir, existe una constante positiva C_2 tal que $\|\cdot\|_{\phi_1, p_2} \leq C_2 \|\cdot\|_{\phi_1, p_1}$, con base en esto y en la desigualdad (7), en la expresión (5) se tiene

$$\begin{aligned} \|s_{p_2}(\tilde{q})\|_{\phi_1, p_2} &\leq C_2 \|v - u\|_{\phi_1, p_1} + \|\bar{u}\|_{\phi_1, p_2} + KC_1 \|v - u\|_{\phi_1, p_2} \\ &\leq C_2 \|v - u\|_{\phi_1, p_1} + \|\bar{u}\|_{\phi_1, p_2} + KC_1 C_2 \|v - u\|_{\phi_1, p_1} \\ &\leq C_2(1 + KC_1) \|v - u\|_{\phi_1, p_1} + \|\bar{u}\|_{\phi_1, p_2}. \end{aligned}$$

Como $v \in A_r$ entonces $\|v - u\|_{\phi_1, p_1} < r$. Por lo tanto, en la desigualdad anterior $\|s_{p_2}(\tilde{q})\|_{\phi_1, p_2} < C_2(1 + KC_1)r + \|\bar{u}\|_{\phi_1, p_2}$; para que $\|s_{p_2}(\tilde{q})\|_{\phi_1, p_2} < 1$ es necesario que r cumpla la desigualdad

$$r < \frac{1 - \|\bar{u}\|_{\phi_1, p_2}}{C_2(1 + KC_1)}. \quad (8)$$

Por lo tanto, en el caso de que r satisfaga la desigualdad 8, $\|s_{p_2}(\tilde{q})\|_{\phi_1, p_2} < 1$. De allí que $s_{p_2}(\tilde{q})$ es un elemento en $s_{p_2}(\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2}) \subset \mathcal{V}_{p_2}$, debido a que los mapeos de transición son continuos entonces el conjunto $s_{p_1}(\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2})$ es abierto en la topología inducida por la norma de Luxemburgo en B_{p_1} . \square

A continuación se define la convergencia exponencial que servirá para demostrar que los conjuntos $s_{p_i}(\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2})$ con $i = 1, 2$ son abiertos también por medio de la topología que induce la convergencia.

Definición 3.10. La sucesión $\{g_n\}_{n \in \mathbb{N}}$ en \mathfrak{M}_μ converge exponencialmente (o es *e-convergente*) a $g \in \mathfrak{M}_\mu$, si $\{g_n\}_{n \in \mathbb{N}}$ tiende a g en μ probabilidad y además $\left\{\frac{g_n}{g}\right\}$ y $\left\{\frac{g}{g_n}\right\}$ son eventualmente convergentes en el siguiente sentido:

$$\liminf_{n \rightarrow \infty} E_g \left[\left(\frac{g_n}{g} \right)^p \right] < \infty \quad y \quad \liminf_{n \rightarrow \infty} E_g \left[\left(\frac{g}{g_n} \right)^p \right] < \infty,$$

para todo $p > 1$.

La *e-convergencia* de $\{g_n\}$ a g en \mathfrak{M}_μ es equivalente a la convergencia

$$\frac{g_n}{g} \rightarrow 1 \quad y \quad \frac{g}{g_n} \rightarrow 1,$$

con respecto a todas las seminormas $f \rightarrow E_g[|f|^h]$ para todo $h > 1$, esto es

$$E_p \left[\left| \frac{g_n}{g} \right|^h \right] \rightarrow 1 \quad y \quad E_p \left[\left| \frac{g}{g_n} \right|^h \right] \rightarrow 1, \quad (3.14)$$

que es equivalente a

$$\frac{1}{2} \left(E_p \left[\left| \frac{g_n}{g} \right|^h + \left| \frac{g}{g_n} \right|^h \right] \right) - 1 = \frac{1}{2} \left(E_p \left[\left| \frac{g_n}{g} \right|^h \right] + E_p \left[\left| \frac{g}{g_n} \right|^h \right] \right) - 1 .$$

Por la expresión (3.14) se concluye que la e -convergencia de g_n a g es equivalente a

$$\left\{ \frac{1}{2} \left(E_p \left[\left| \frac{g_n}{g} \right|^h + \left| \frac{g}{g_n} \right|^h \right] \right) - 1 \right\} \rightarrow 0 .$$

El siguiente teorema permitirá demostrar que los mapeos e_p y s_p son continuos en el sentido que preservan la convergencia de las sucesiones, sólo que esta convergencia es la e -convergencia definida en (3.10).

Teorema 3.16. *Sea $\{g_n\}_{n \in \mathbb{N}}$ una sucesión e -convergente a g y supongamos que $g \in \mathcal{U}_p$. La sucesión $\{g_n\}_{n \in \mathbb{N}}$ está eventualmente en \mathcal{U}_p y la sucesión $u_n = s_p(g_n)$ converge a $u = s_p(g)$ en \mathcal{V}_p .*

Demostración

Debe demostrarse que existe una sucesión $u_n \in \mathcal{V}_p$ tal que $g_n = e_p(u_n)$ está eventualmente en \mathcal{U}_p y $u_n \rightarrow u$ converge en $p \cdot \mu$ probabilidad en la bola unidad \mathcal{V}_p . Como $u = s_p(g)$ entonces $g = e_p(u) = e^{u - K_p(u)} p$. Consideremos ahora la sucesión $\{w_n\}_{n \in \mathbb{N}} \in L^{\phi_1}(p \cdot \mu)$ dada por

$$g_n = e^{w_n - K_p(u)} p .$$

De la e -convergencia de g_n a g se tiene que

$$\left\{ \frac{1}{2} \left(E_p \left[\left| \frac{g_n}{g} \right|^h + \left| \frac{g}{g_n} \right|^h \right] \right) - 1 \right\} \rightarrow 0 . \quad (3.15)$$

Utilizando la definición de la sucesión g_n y de la densidad g obtenemos

$$\begin{aligned} \frac{1}{2} \left(E_p \left[\left| \frac{g_n}{g} \right|^h + \left| \frac{g}{g_n} \right|^h \right] \right) - 1 &= \frac{1}{2} \left(E_p \left[\left| \frac{e^{w_n - K_p(u)}}{e^{u - K_p(u)}} \right|^h + \left| \frac{e^{u - K_p(u)}}{e^{w_n - K_p(u)}} \right|^h \right] \right) - 1 \\ &= \frac{1}{2} E_p \left[(e^{w_n - u})^h + \frac{1}{(e^{w_n - u})^h} \right] - 1 \\ &= E_p \left[\frac{1}{2} \left(e^{h(w_n - u)} + e^{-h(w_n - u)} \right) \right] - 1 . \end{aligned}$$

Puesto que $\phi_1(x) = \cosh x - 1$ entonces, la expresión anterior se puede escribir como

$$\frac{1}{2} \left(E_p \left[\left| \frac{g_n}{g} \right|^h + \left| \frac{g}{g_n} \right|^h \right] \right) - 1 = E_p[\cosh(h(w_n - u)) - 1] = E_p[\phi_1(h(w_n - u))] .$$

Por (3.15) se sigue que $E_p[\phi_1(h(w_n - u))] \rightarrow 0$ para todo $h > 1$. Así

$$\limsup_{n \rightarrow \infty} E_p[\phi_1(h(w_n - u))] \leq 1 ,$$

Y por lo tanto $w_n \rightarrow u$, donde $u \in \mathcal{V}_p$. Tomemos ahora $u_n = w_n - E_p[w_n]$, luego $E_p[u_n] = E_p[w_n - E_p[w_n]] = E_p[w_n] - E_p[w_n] = 0$, por lo que para cada $n \in \mathbb{N}$, $u_n \in B_p$. Veamos ahora quien es $K_p(u_n)$. Por la definición de la función acumulante

$$\begin{aligned} K_p(u_n) &= \ln(E_p[e^{u_n}]) = \ln\left(E_p\left[e^{w_n - E_p[w_n]}\right]\right) = \ln\left(\frac{E_p[e^{w_n}]}{e^{E_p[w_n]}}\right) \\ &= \ln(E_p[e^{w_n}]) - \ln\left(e^{E_p[w_n]}\right) = \ln(E_p[e^{w_n}]) - E_p[w_n] . \end{aligned}$$

En términos del acumulante se sigue que $K_p(u_n) = K_p(w_n) - E_p[w_n]$ y por tanto

$$u_n - K_p(u_n) = (w_n - E_p[w_n]) - (K_p(w_n) - w_n) = w_n - K_p(w_n) .$$

Como w_n converge a u y el operador integral es continuo, entonces $w_n - E_p[w_n]$ converge a $u - E_p[u]$. Además $u \in B_p$ por lo que $E_p[u] = 0$ y así $(w_n - E_p[w_n]) \rightarrow u$ y por tanto $u_n \rightarrow u$. \square

En la demostración, w_n y u_n son dos sucesiones que convergen a u , pero difieren en una constante dada por $E_p[w_n]$, es decir, se hace una traslación de todo el espacio $L^{\phi_1}(p \cdot \mu)$ al subespacio abierto \mathcal{V}_p . Además, se está demostrando que el mapeo s_p es continuo, así como e_p ya que se preserva la e -convergencia de las sucesiones g_n en \mathfrak{M}_μ . También, como $s_p : \mathcal{U}_p \rightarrow \mathcal{V}_p$ es continua y \mathcal{V}_p es un subconjunto abierto en $L^{\phi_1}(p \cdot \mu)$ por la construcción, entonces \mathcal{U}_p es abierto también en \mathfrak{M}_μ . Por lo que, dadas dos densidades p_1 y p_2 en \mathfrak{M}_μ , los conjuntos \mathcal{U}_{p_1} y \mathcal{U}_{p_2} son abiertos y por ende $\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2}$ también lo es.

Teorema 3.17. *Para las densidades p_1 y p_2 en \mathfrak{M}_μ , la función $s_{p_2} \circ e_{p_1}$ es un homeomorfismo.*

Demostración

Como los mapeos e_{p_1} y s_{p_2} son continuos entonces $s_{p_2} \circ e_{p_1}$ es otro mapeo continuo. Ya que $(s_{p_2} \circ e_{p_1})^{-1} = e_{p_1}^{-1} \circ s_{p_2}^{-1} = s_{p_1} \circ e_{p_2}$ entonces por un razonamiento análogo, el mapeo

$(s_{p_2} \circ e_{p_1})^{-1}$ es continuo. Por lo que el mapeo $s_{p_2} \circ e_{p_1}$ y su inversa son continuos, esto es, $s_{p_2} \circ e_{p_1}$ es un homeomorfismo. \square

En el teorema 3.17 se ha probado que los mapeos transición son continuos, en el siguiente teorema se caracteriza la derivada de dicho mapeo en la dirección de $v \in \mathcal{V}_p$.

Teorema 3.18. *La derivada del mapeo transición*

$$\begin{aligned} (s_{p_2} \circ e_{p_1})(\cdot) : \quad s_{p_1}(\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2}) &\rightarrow s_{p_2}(\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2}) \\ u &\mapsto u + \ln \left(\frac{p_1}{p_2} \right) - E_{p_2} \left[u + \ln \left(\frac{p_1}{p_2} \right) \right] \end{aligned}$$

en la dirección de $v \in \mathcal{V}_p$ es $D(s_{p_2} \circ e_{p_1})(u) \cdot v = v - E_{p_2}[v]$.

Demostración

Haciendo uso de la definición del operador diferencial, se sigue que para el mapeo $s_{p_2} \circ e_{p_1}$ en la dirección de v se tiene (donde $r(u-v)$ es tal que $\lim_{v \rightarrow 0} \frac{r(u-v)}{\|v\|} = 0$ y $r(u) = 0$)

$$\begin{aligned} D(s_{p_2} \circ e_{p_1})(u) \cdot v &= (s_{p_2} \circ e_{p_1})(u+v) - (s_{p_2} \circ e_{p_1})(u) + r(u-v) \\ &= u+v + \ln \left(\frac{p_1}{p_2} \right) - E_{p_2} \left[u+v + \ln \left(\frac{p_1}{p_2} \right) \right] - u - \ln \left(\frac{p_1}{p_2} \right) + \\ &\quad E_{p_2} \left[u + \ln \left(\frac{p_1}{p_2} \right) \right] + r(u-v) \\ &= v - E_{p_2}[v] + r(u-v) . \end{aligned}$$

Así que la derivada del mapeo transición, en la dirección de v , es $v - E_{p_2}[v]$. \square

Teorema 3.19. *La colección de pares $\{(\mathcal{U}_p, s_p)\}_{p \in \mathfrak{M}_\mu}$ es una atlas de clase C^∞ modelado sobre \mathfrak{M}_μ .*

Demostración

Debido a que $p \in \mathcal{U}_p$ entonces \mathcal{U}_p cubre a \mathfrak{M}_μ , esto es, $\bigcup_{p \in \mathfrak{M}_\mu} \mathcal{U}_p = \mathfrak{M}_\mu$; además cada \mathcal{U}_p es abierto y s_p es una biyección con lo que (\mathcal{U}_p, s_p) es una carta. Como los mapeos transición son afines, entonces se sigue que $s_{p_2} \circ e_{p_1}$ es un mapeo de clase C^∞ , por lo que la colección $\{(\mathcal{U}_p, s_p)\}_{p \in \mathfrak{M}_\mu}$ es una atlas de clase C^∞ . \square

De acuerdo con el teorema 3.19 se ha construido una variedad de Banach de clase C^∞ para el espacio \mathfrak{M}_μ

Definición 3.11. *La variedad exponencial estadística es la variedad definida por el atlas en el teorema 3.19.*

Definición 3.12. *Para una densidad $p \in \mathfrak{M}_\mu$, el modelo exponencial maximal en p está definido por la familia de densidades*

$$\mathcal{E}_p := \{e^{u-K_p(u)}p : u \in \text{dom}^0 K_p\} .$$

Se presenta las inclusiones conjuntistas $\mathcal{V}_p \subseteq \mathcal{E}_p \subseteq \mathfrak{M}_\mu$. La demostración del siguiente teorema se puede ver en [22].

Teorema 3.20. *El modelo exponencial maximal es la componente conexa que contiene a p en la variedad \mathfrak{M}_μ .*

3.5. Entropía Relativa

Definición 3.13. *Sean $p, q \in \mathfrak{M}_\mu$. Si $\frac{q}{p} \ln \left(\frac{q}{p}\right)$ es $p \cdot \mu$ integrable entonces la entropía relativa de Kullback-Leibler de q con respecto a p es el número*

$$K(q|p) = E_p \left[\frac{q}{p} \ln \left(\frac{q}{p} \right) \right] .$$

En la definición, si la expresión $\frac{q}{p} \ln \left(\frac{q}{p}\right)$ es $p \cdot \mu$ integrable entonces $\ln \left(\frac{q}{p}\right)$ es $q \cdot \mu$ integrable y la entropía relativa se puede escribir como

$$K(q|p) = E_p \left[\frac{q}{p} \ln \left(\frac{q}{p} \right) \right] = E_q \left[\ln \left(\frac{q}{p} \right) \right] . \quad (3.16)$$

En efecto, de acuerdo con la definición del funcional $E_p[\cdot]$ se tiene

$$K(q|p) = E_p \left[\frac{q}{p} \ln \left(\frac{q}{p} \right) \right] = \int_{\Omega} \frac{q}{p} \ln \left(\frac{q}{p} \right) p d\mu = \int_{\Omega} \ln \left(\frac{q}{p} \right) q d\mu = E_q \left[\ln \left(\frac{q}{p} \right) \right] .$$

Debido a que la función $-\ln(\cdot)$ es convexa entonces por la desigualdad de Jensen

$$-\ln \left(\int_{\Omega} \frac{p}{q} q d\mu \right) \leq \int_{\Omega} -\ln \left(\frac{p}{q} \right) q d\mu ,$$

como $E_q \left[\frac{p}{q} \right] = E_p[1] = 1$ entonces $0 \leq E_q \left[\ln \left(\frac{q}{p} \right) \right]$ y por (3.16) se concluye

$$K(q|p) \geq 0. \quad (3.17)$$

En el caso en que $p = q$ casi en todas partes de Ω (lo que implica que $u = 0$), entonces $K(q|p) = E_p[\ln 1] = 0$. De manera recíproca, si $K(p|q) = 0$ se tiene que $p = q$.

Sean $u \in \mathcal{V}_p$ y $p \in \mathfrak{M}_\mu$ de tal manera que $q = e^{u - K_p(u)}p$, por lo que q es una densidad de probabilidad y además $\ln \left(\frac{q}{p} \right) = u - K_p(u)$. Debido a que $u - K_p(u) \in L^{\phi_1}(p \cdot \mu)$ entonces $u - K_p(u) \in L^{\phi_1}(p \cdot \mu) \subseteq L^1(p \cdot \mu)$, es decir, $u - K_p(u)$ es $p \cdot \mu$ integrable y por tanto $\ln \left(\frac{q}{p} \right)$ es $p \cdot \mu$ integrable e implica la existencia de la entropía relativa $K(q|p) < \infty$.

Integrando la igualdad $\ln \left(\frac{q}{p} \right) = u - K_p(u)$ respecto de la medida $q \cdot \mu$ se tiene

$$E_q \left[\ln \left(\frac{q}{p} \right) \right] = E_q[u] - E_q[K_p(u)] = E_q[u] - K_p(u), \quad (3.18)$$

donde la función acumulante es un número real y por tanto es constante respecto de la medida $q \cdot \mu$. Como la entropía relativa de q respecto de p está dada por $K(q|p) = E_q \left[\ln \left(\frac{q}{p} \right) \right]$ entonces, por transitividad con la igualdad 3.18 se tiene

$$K(q|p) = E_q[u] - K_p(u) \quad (3.19)$$

En el corolario 3.1 se demostró que $DK_p(u) \cdot v = E_q[v]$ con $v \in \mathcal{V}_p$, si $u = v$ entonces $DK_p(u) \cdot u = E_q[u]$; al sustituir en la igualdad 3.19 se obtiene

$$K(q|p) = DK_p(u) \cdot u - K_p(u). \quad (3.20)$$

Esta expresión muestra la relación directa entre la entropía relativa de q con respecto a p , el funcional acumulante y la variable aleatoria u , la cual es un elemento de la bola unitaria de B_p . Además, la entropía relativa es un número real ($K(q|p) < \infty$). Este análisis motiva que este bien definida la siguiente función

$$\begin{aligned} \mathcal{V}_p &\rightarrow \mathbb{R} \\ u &\mapsto K(q|p) = DK_p(u) \cdot u - K_p(u). \end{aligned} \quad (3.21)$$

De acuerdo con (3.21) podemos hallar la derivada de la entropía relativa de q con respecto a p en la dirección de $v \in \mathcal{V}_p$, la cual está dada por

$$DK(q|p) \cdot v = D^2K_p(u)(u, v) - DK_p(u) \cdot v. \quad (3.22)$$

En el caso particular en que $v = u$ se tiene $DK(q|p) \cdot u = D^2K_p(u)(u, u) - DK_p(u) \cdot u$. El siguiente teorema relaciona la entropía relativa de q con respecto a p y el funcional acumulante, garantizando que la entropía es una extensión del funcional acumulante.

Teorema 3.21. *Sea $p \in \mathfrak{M}_\mu$ y $q = e^{u-K_p(u)}p$. Si $u \in \mathcal{V}_p$ entonces $K(p|q) = K_p(u)$.*

Demostración

Por hipótesis $q = e^{u-K_p(u)}p$, donde $q \in \mathfrak{M}_\mu$, $q > 0$, luego $\frac{p}{q} = e^{-u+K_p(u)}$ o equivalente a $\ln\left(\frac{p}{q}\right) = -u + K_p(u)$. Por definición, la entropía relativa de p con respecto a q es

$$K(p|q) = E_q \left[\frac{p}{q} \ln \left(\frac{p}{q} \right) \right] = E_p \left[\ln \left(\frac{p}{q} \right) \right].$$

Ya que $\ln\left(\frac{p}{q}\right) = -u + K_p(u)$ se obtiene $K(p|q) = E_p[-u + K_p(u)]$, ya que u es una variable aleatoria centrada respecto de $p \cdot \mu$ y el acumulante es constante respecto de esta medida, se concluye $K(p|q) = E_p[-u] + E_p[K_p(u)] = K_p(u)$. \square

Si se combinan los resultados del teorema 3.21 con la igualdad dada en 3.21 se tiene que para cada $u \in \mathcal{V}_p$

$$E_q[u] = DK_p(u) \cdot u = K(q|p) + K(p|q). \quad (3.23)$$

3.6. B_p como un espacio de Hilbert

El conjunto B_p tiene la estructura de espacio lineal convexo y cerrado, que es un espacio de Banach con la norma de Luxemburgo inducida por el espacio de Orlicz $L^{\phi_1}(p \cdot \mu)$. Al espacio B_p se le dotará de la estructura de espacio de Hilbert con la norma inducida por el producto interno y $(B_p, \|\cdot\|_{\phi_1, p})$ se identifica con un subespacio del espacio lineal $(B_p, \|\cdot\|_{B_p})$.

Teorema 3.22. *Se verifica la inclusión continua $B_p \subseteq L^n(p \cdot \mu)$ con $n \in \mathbb{N}$.*

Demostración

Sea $u \in B_p$. Como B_p es un subespacio de la clase de Cramer entonces $u \in \mathcal{C}_p$, por lo que la función generadora de momentos existe, $\hat{u}_p(t) < \infty$ para t en una vecindad de cero. Como $\hat{u}_p(t) < \infty$ entonces existen los momentos de todos los órdenes, esto es, $E_p[u^n] < \infty$, por lo que $u \in L^n(p \cdot \mu)$ para $n \in \mathbb{N}$. Y se cumple que $B_p \subseteq L^n(p \cdot \mu)$, donde el mapeo continuo

entre dichos espacios es $u \mapsto E_p[u^n]$ por ser el operador integral continuo. \square

De acuerdo con el teorema se sigue que B_p es un subespacio de $L^2(p \cdot \mu)$, donde $L^2(p \cdot \mu)$ es un espacio de Hilbert. Como B_p es completo, se requiere definir un producto interno sobre B_p para concluir así mismo que éste es un espacio de Hilbert.

Definición 3.14. Sobre $B_p \times B_p$ se define el operador

$$\begin{aligned} \langle \cdot, \cdot \rangle_p : B_p \times B_p &\rightarrow \mathbb{R} \\ (u, v) &\mapsto \langle u, v \rangle_p = E_p[uv]. \end{aligned}$$

Debido a que la integral es un operador lineal continuo entonces el operador $\langle \cdot, \cdot \rangle_p$ también es continuo. Sean u y v variables aleatorias en B_p ($E_p[u] = E_p[v] = 0$), el operador definido en (3.14) se puede escribir como

$$\langle u, v \rangle_p = E_p[uv] - E_p[u]E_p[v].$$

En probabilidad, la **covarianza** relativa a la medida $p \cdot \mu$ de las variables aleatorias u y v está dada por,

$$cov_p[u, v] = E_p[(u - E_p[u])(v - E_p[v])] = E_p[uv] - E_p[u]E_p[v]. \quad (3.24)$$

En el caso en que las variables aleatorias sean centradas, la covarianza se escribe como $cov_p[u, v] = E_p[uv]$ y así el operador de la definición (3.14) se puede escribir como

$$\langle u, v \rangle_p = cov_p[u, v]. \quad (3.25)$$

Por definición, la **varianza** de la variable aleatoria u respecto de la medida de probabilidad $p \cdot \mu$ está dada por $var_p(u) = E_p[(u - E_p[u])^2]$. Si u es una variable aleatoria centrada en $L^{\phi_1}(p \cdot \mu)$ entonces, la varianza se escribe como $var_p[u] = E_p[u^2]$, es decir, el segundo momento central alrededor del cero. En el caso en que $u = v$ en la definición (3.14) se tiene

$$\langle u, u \rangle_p = E_p[u^2] = var_p[u]. \quad (3.26)$$

La **desviación típica** es la raíz cuadrada positiva de la varianza. Para la variable aleatoria u , bajo la medida $p \cdot \mu$, se tiene

$$\sigma_p(u) = \sqrt{var_p[u]}. \quad (3.27)$$

En el siguiente teorema se demuestra que el operador $\langle \cdot, \cdot \rangle_p$ es un producto escalar sobre el espacio de Banach B_p .

Teorema 3.23. *El operador $\langle \cdot, \cdot \rangle_p$ es un producto escalar en B_p .*

Demostración

Se deben demostrar las cuatro condiciones para que $\langle \cdot, \cdot \rangle_p$ sea un producto escalar sobre B_p

- i. Sean u, v, w en B_p , de acuerdo con la definición 3.14 se tiene

$$\langle u + v, w \rangle_p = E_p[(u + v)w] = E_p[uw + vw] = E_p[uw] + E_p[vw] = \langle u, w \rangle_p + \langle v, w \rangle_p .$$

De igual forma se demuestra que $\langle u, v + w \rangle_p = \langle u, v \rangle_p + \langle u, w \rangle_p$ para concluir así que dicho operador es bilineal.

- ii. Sea $u \in B_p$, veamos que $\langle u, u \rangle_p \geq 0$. De acuerdo con (3.26) se tiene $\langle u, u \rangle_p = var_p[u]$, ya que por la definición de la varianza $var_p[u] \geq 0$ entonces se sigue que $\langle u, u \rangle_p \geq 0$ para todo $u \in B_p$.
- iii. Sean u y v en B_p . Por definición tenemos que $\langle u, v \rangle_p = cov_p[u, v]$, ya que la covarianza es conmutativa se sigue que $\langle u, v \rangle_p = cov_p[v, u] = cov_p[u, v] = \langle v, u \rangle_p$.
- iv. Sea $u \in B_p$ veamos que $\langle u, u \rangle_p = 0$ si y sólo si $u = 0$. Supongamos que $\langle u, u \rangle_p = 0$, en términos de la varianza resulta $var_p[u] = 0$, por lo que la variable aleatoria u es constante, es decir, $u = \alpha$ con $\alpha \in \mathbb{R}$, igualdad que es equivalente a $E_p[u] = E_p[\alpha]$, como $E_p[u] = 0$ y α es constante entonces $\alpha = 0$ y por tanto $u = 0$. Si $u = 0$ entonces $\langle u, u \rangle_p = E_p[u^2] = E_p[0] = 0$.

Con estas cuatro propiedades se concluye que el operador $\langle \cdot, \cdot \rangle_p$ definido sobre B_p es un producto escalar real. \square

Ahora bien, como B_p es completo y está definido sobre éste el producto escalar $\langle \cdot, \cdot \rangle_p$ se concluye que B_p es en sí un espacio de Hilbert contenido en $L^2(p \cdot \mu)$. El producto escalar induce una norma sobre B_p , la cual denotaremos como $\| \cdot \|_{B_p}$, la cual está dada por

$$\begin{aligned} \| \cdot \|_{B_p} : B_p &\rightarrow \mathbb{R} \\ u &\mapsto \|u\|_{B_p} = \sqrt{\langle u, u \rangle_p} . \end{aligned} \tag{3.28}$$

En este caso se hizo alusión a la palabra norma para hacer alusión a $\| \cdot \|_{B_p}$, en el corolario 3.2 se demostrará que efectivamente cumple las propiedades de la norma, además que se demostrará la equivalencia con la norma de Luxemburgo $\| \cdot \|_{\phi_1, p}$ definida sobre el espacio de Orlicz $L^{\phi_1}(p \cdot \mu)$.

Si u es una variable aleatoria en B_p , entonces la norma bajo B_p se puede escribir en términos de la varianza como

$$\|u\|_{B_p}^2 = \langle u, u \rangle_p = E_p[u^2] = \text{var}_p[u]. \quad (3.29)$$

Esto permite concluir que la desviación típica σ_p es la norma bajo B_p de la variable aleatoria u es decir, $\sigma_p(u) = \|u\|_{B_p}$. Según el teorema 3.22, si $u \in B_p$ entonces $E_p[u^2] < \infty$, es decir, el segundo momento central siempre existe por lo que la varianza está definida para todos los elementos de B_p , así para cada $u \in B_p$, la norma $\|u\|_{B_p}$ siempre existe ($\|u\|_{B_p} < \infty$).

Los siguientes teoremas son las propiedades clásicas de los espacios de Hilbert, los cuales se cumplirán para B_p en términos de la norma B_p . Entre estos teoremas se encuentra la desigualdad de Cauchy-Schwarz y la desigualdad triangular.

En el siguiente teorema se escribe $(p \cdot \mu)(\alpha u + \beta v = 0) = 1$ lo cual indica que la medida bajo $p \cdot \mu$ del evento $\alpha u + \beta v = 0$ es 1, es decir, es un evento seguro; ya que $(\Omega, \mathcal{A}, p \cdot \mu)$ es un espacio de probabilidad.

Teorema 3.24. Desigualdad de Cauchy-Schwarz Sean u y v elementos en B_p entonces $uv \in L^1(p \cdot \mu)$ y

$$|\langle u, v \rangle_p|^2 \leq \|u\|_{B_p} \|v\|_{B_p}.$$

Donde la igualdad ocurre si y sólo si existen constantes α y β no ambas nulas tales que $(p\mu)(\alpha u + \beta v = 0) = 1$.

Demostración

Considere la función real valuada

$$\begin{aligned} f: \mathbb{R} &\rightarrow \mathbb{R} \\ t &\mapsto f(t) = E_p[(u + tv)^2], \end{aligned} \quad (3.30)$$

que es no-negativa ($f(t) \geq 0$) para todo t en \mathbb{R} . Como t es constante respecto a la medida $p \cdot \mu$ entonces

$$\begin{aligned} f(t) &= E_p[(u + tv)^2] = E_p[u^2 + 2uvt + v^2t^2] \\ &= E_p[u^2] + 2E_p[uv]t + E_p[v^2]t^2. \end{aligned}$$

Luego f es un polinomio cuadrático no negativo, por lo que el discriminante $\Delta = B^2 - 4AC$ (con la igualdad $f(t) = At^2 + Bt + C$) es no-positivo $\Delta \leq 0$, así

$$4(E_p[uv])^2 - 4E_p[u^2]E_p[v^2] \leq 0,$$

equivalente a $(E_p[uv])^2 \leq E_p[u^2]E_p[v^2]$ como $E_p[u^2]$ y $E_p[v^2]$ son finitos entonces $E_p[uv]$ es finito y así $uv \in L^1(p \cdot \mu)$ y además

$$|E_p[uv]|^2 = (E_p[uv])^2 \leq E_p[u^2]E_p[v^2].$$

En términos de la norma B_p y del producto escalar se tiene

$$|\langle u, v \rangle_p|^2 \leq \|u\|_{B_p}^2 \|v\|_{B_p}^2,$$

de donde $|\langle u, v \rangle_p| \leq \|u\|_{B_p} \|v\|_{B_p}$. Expresión que se puede escribir como

$$|cov_p[u, v]| \leq \sqrt{var_p[u]} \sqrt{var_p[v]}.$$

“ \Rightarrow ” Consideremos que $\alpha u + \beta v = 0$ con probabilidad igual 1 respecto de $p \cdot \mu$, de donde $\alpha u = -\beta v$. Si $\alpha \neq 0$ y $v \neq 0$ casi en todas partes de Ω entonces $\alpha^2 u^2 = \beta^2 v^2$ y así $\alpha^2 E_p[u^2] = \beta^2 E_p[v^2]$ y por tanto

$$\frac{\beta^2}{\alpha^2} = \frac{E_p[u^2]}{E_p[v^2]}. \quad (1)$$

Como $\alpha \neq 0$ entonces $u = -\frac{\beta}{\alpha}v$, lo cual, junto con la expresión (1), se obtienen las igualdades

$$\begin{aligned} |\langle u, v \rangle_p|^2 &= (\langle u, v \rangle_p)^2 = (E_p[uv])^2 = \left(E_p \left[\left(-\frac{\beta}{\alpha}v \right) v \right] \right)^2 = \left(E_p \left[-\frac{\beta}{\alpha}v^2 \right] \right)^2 \\ &= \frac{\beta^2}{\alpha^2} (E_p[v^2])^2 = \frac{E_p[u^2]}{E_p[v^2]} (E_p[v^2])^2 = E_p[u^2]E_p[v^2]. \end{aligned}$$

En términos de la norma y el producto escalar se satisface que $|\langle u, v \rangle_p|^2 = \|u\|_{B_p}^2 \|v\|_{B_p}^2$, así $|\langle u, v \rangle_p| = \|u\|_{B_p} \|v\|_{B_p}$ siempre que $\alpha u + \beta v = 0$ tiene probabilidad 1 respecto de $p \cdot \mu$.

“ \Leftarrow ” Supongamos que $|\langle u, v \rangle_p|^2 = \|u\|_{B_p}^2 \|v\|_{B_p}^2$ (2). Probemos que $\alpha u + \beta v = 0$ con probabilidad 1 bajo $p \cdot \mu$. La expresión (2) se puede escribir como $(E_p[uv])^2 - E_p[u^2]E_p[v^2] = 0$ (3).

Si $\alpha = E_p[v^2]$ y $\beta = -E_p[uv]$ entonces $\alpha \geq 0$, de donde $\alpha > 0$ ya que si $\alpha = 0$ entonces $v = 0$ casi en todas partes de Ω y así $\beta = 0$ por lo que $\alpha u + \beta v = 0$ casi en todas partes de Ω . Supongamos que $\alpha > 0$, por las propiedades del operador $E_p[\cdot]$ se tiene

$$\begin{aligned} E_p[(\alpha u + \beta v)^2] &= \alpha^2 E_p[u^2] + 2\alpha\beta E_p[uv] + \beta^2 E_p[v^2] \\ &= (E_p[v^2])^2 E_p[u^2] - 2(E_p[uv])^2 E_p[v^2] + (E_p[uv])^2 E_p[v^2] \\ &= (E_p[v^2])^2 E_p[u^2] - (E_p[uv])^2 E_p[v^2] \\ &= E_p[v^2](E_p[v^2]E_p[u^2] - (E_p[uv])^2). \end{aligned}$$

Por (3) se concluye que $E_p[(\alpha u + \beta v)^2] = 0$, equivalente a $\text{var}_p[\alpha u + \beta v] = 0$; luego la variable aleatoria es constante, llámese r , $\alpha u + \beta v = r$, como u y v son variables aleatorias centradas entonces $r = 0$ y así $\alpha u + \beta v = 0$ con probabilidad 1 respecto de la medida $p \cdot \mu$. \square

Teorema 3.25. Sean u y v elementos en B_p entonces $\|u + v\|_{B_p} < \infty$ y

$$\|u + v\|_{B_p}^2 = \|u\|_{B_p}^2 + 2\langle u, v \rangle_p + \|v\|_{B_p}^2 .$$

Demostración

Con base en la definición de la norma B_p se tiene

$$\|u + v\|_{B_p}^2 = E_p[(u + v)^2] = E_p[u^2] + 2E_p[uv] + E_p[v^2] .$$

Por el teorema 3.24 se tiene que $uv \in L^1(p \cdot \mu)$, es decir, $E_p[uv] < \infty$, luego $\|u + v\|_{B_p} < \infty$ debido a que $E_p[u^2]$ y $E_p[v^2]$ son finitas al igual que $E_p[uv]$. Ahora bien, en términos de la norma B_p y el producto escalar se concluye $\|u + v\|_{B_p}^2 = \|u\|_{B_p}^2 + 2\langle u, v \rangle_p + \|v\|_{B_p}^2$. \square

Teorema 3.26. Desigualdad Triangular Sean u y v en B_p entonces

$$\|u + v\|_{B_p} \leq \|u\|_{B_p} + \|v\|_{B_p} .$$

Demostración

Por el teorema 3.25 se tiene que $\|u + v\|_{B_p} < \infty$ y

$$\|u + v\|_{B_p}^2 = \|u\|_{B_p}^2 + 2\langle u, v \rangle_p + \|v\|_{B_p}^2 .$$

Como la imagen del producto escalar es un número real entonces, $\langle u, v \rangle_p \leq |\langle u, v \rangle_p|$. Por la desigualdad de Cauchy-Schwarz resulta $\langle u, v \rangle_p \leq \|u\|_{B_p} \|v\|_{B_p}$. Por tanto

$$\|u + v\|_{B_p}^2 \leq \|u\|_{B_p}^2 + 2\|u\|_{B_p} \|v\|_{B_p} + \|v\|_{B_p}^2 \leq (\|u\|_{B_p} + \|v\|_{B_p})^2 .$$

Para concluir que $\|u + v\|_{B_p} \leq \|u\|_{B_p} + \|v\|_{B_p}$, para todo u, v en B_p . \square

El siguiente teorema permite demostrar que la norma B_p es, en efecto, una norma sobre B_p .

Corolario 3.2. La aplicación $\|\cdot\|_{B_p}$ es una norma en B_p .

Demostración

Por el teorema 3.26 se sabe que $\|\cdot\|_{B_p}$ satisface la desigualdad triangular. Además, si $u \in B_p$ entonces $\|u\|_{B_p} \geq 0$ ya que $\|u\|_{B_p}^2 = \text{var}_p[u] \geq 0$ para cada $u \in B_p$. Ahora bien, por las propiedades de la varianza se tiene que para $\alpha \in \mathbb{R}$, $\|\alpha u\|_{B_p}^2 = \text{var}_p[\alpha u] = \alpha^2 \text{var}_p[u] = \alpha^2 \|u\|_{B_p}^2$, con lo cual se puede concluir que $\|\alpha u\|_{B_p} = |\alpha| \|u\|_{B_p}$. En el caso en que $u = 0$ resulta $\|u\|_{B_p}^2 = 0$ y así $\text{var}_p[u] = 0$, de allí que la variable aleatoria u sea constante, llámese r , así $u = r$, pero al ser u centrada resulta que $r = 0$ y por tanto $u = 0$. Se concluye que la aplicación $\|\cdot\|_{B_p}$ es una norma sobre los elementos de B_p . \square

La norma B_p no es invariante por traslación; ya que si $u \in B_p$ y b es una constante entonces $E_p[u + b] = b$ y así $u + b$ no es un elemento de B_p por lo que no está definida la norma B_p .

La ecuación (3.29) se puede escribir como

$$\begin{aligned} \|u\|_{B_p} &= \sqrt{\text{var}_p[u]} = \sqrt{E_p[u^2]} \\ &= \sqrt{\int_{\Omega} u^2 p d\mu} = \left(\int_{\Omega} u^2 p d\mu \right)^{1/2}. \end{aligned}$$

Por medio de la norma definida sobre el espacio de Banach $L^2(p \cdot \mu)$ se tiene que la norma B_p se puede escribir como

$$\|u\|_{B_p} = \|u\|_{L^2}. \quad (3.31)$$

El siguiente teorema permite establecer que el espacio $(B_p, \|\cdot\|_{\phi_1, p})$ se identifica con un subespacio del espacio lineal $(B_p, \|\cdot\|_{B_p})$.

Teorema 3.27. *Para $u \in B_p$ se verifica*

$$\|u\|_{B_p} \leq k \|u\|_{\phi_1, p}.$$

Demostración

La función coseno hiperbólico se puede escribir en términos de series de potencias como

$$\cosh x = \sum_{n=0}^{\infty} \frac{x^{2n}}{(2n)!} = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \dots$$

Por lo que la serie de potencias de la función $\phi_1(x)$ se puede escribir como

$$\phi_1(x) = \sum_{n=1}^{\infty} \frac{x^{2n}}{(2n)!} = \frac{x^2}{2!} + \frac{x^4}{4!} + \frac{x^6}{6!} + \dots,$$

con base en esta serie de potencias, resulta la desigualdad $\phi_1(x) \geq \frac{x^2}{2}$. Si $\phi_p(x) = x^2$ entonces $2\phi_1(x) \geq \phi_p(x)$ esto se verifica para todo $x \in \mathbb{R}$. Para una variable aleatoria $u \in B_p$ y un $r > 0$ se tiene la desigualdad $2\phi_1\left(\frac{u}{r}\right) \geq \phi_p\left(\frac{u}{r}\right)$ y al aplicar el operador $E_p[\cdot]$ resulta

$$2E_p\left[\phi_1\left(\frac{u}{r}\right)\right] \geq E_p\left[\phi_p\left(\frac{u}{r}\right)\right]. \quad (1)$$

Consideremos los conjuntos $A = \{r > 0 : E_p[\phi_p\left(\frac{u}{r}\right)] \leq 1\}$ y $B = \{r > 0 : 2E_p[\phi_1\left(\frac{u}{r}\right)] \leq 1\}$, por lo tanto si $r_1 \in B$ entonces $2E_p\left[\phi_1\left(\frac{u}{r_1}\right)\right] \leq 1$, por (1) se sigue que $E_p\left[\phi_p\left(\frac{u}{r_1}\right)\right] \leq 1$ de donde $r_1 \in A$ y se logra la inclusión conjuntista $B \subset A$, para la cual los ínfimos se relacionan como $\inf A \leq \inf B$ en este caso $\inf A$ representa la norma de Luxemburgo del espacio de funciones de Orlicz generado por $\phi_p(x) = x^2$ el cual es equivalente a $L^2(p \cdot \mu)$, es decir, $\inf A = \|u\|_{L^2}$; mientras que el $\inf B = k\|u\|_{\phi_{1,p}}$ por lo tanto se obtiene la desigualdad $\|u\|_{L^2} \leq k\|u\|_{\phi_{1,p}}$ y por 3.31 se concluye $\|u\|_{B_p} \leq k\|u\|_{\phi_{1,p}}$. \square

La siguiente definición permite extender el concepto de ortogonalidad entre vectores a variables aleatorias centradas.

Definición 3.15. Sean u y v en B_p . Se dice que u y v son ortogonales en B_p si y sólo si $\langle u, v \rangle_p = 0$ y se escribe $u \perp_p v$.

Corolario 3.3. Teorema de Pitágoras Si u y v en B_p son ortogonales entonces

$$\|u + v\|_{B_p}^2 = \|u\|_{B_p}^2 + \|v\|_{B_p}^2.$$

Demostración

Sean u y v en B_p variables aleatorias ortogonales luego $\langle u, v \rangle_p = 0$. Haciendo uso del teorema 3.25 se tiene la igualdad $\|u + v\|_{B_p}^2 = \|u\|_{B_p}^2 + 2\langle u, v \rangle_p + \|v\|_{B_p}^2$. Al sustituir el valor del producto escalar se concluye que $\|u + v\|_{B_p}^2 = \|u\|_{B_p}^2 + \|v\|_{B_p}^2$. \square

El coeficiente de correlación ρ ayuda a predecir los resultados posibles de una variable aleatoria u en términos de otra variable aleatoria v . Este coeficiente se escribe en términos de la covarianza y la varianza, bajo la medida $p \cdot \mu$, como

$$\rho_p(u, v) = \frac{\text{cov}_p[u, v]}{\sqrt{\text{var}_p[u]} \sqrt{\text{var}_p[v]}}.$$

En la siguiente definición escribimos dicho coeficiente de correlación en términos del producto escalar y la norma B_p .

Definición 3.16. Sean u y v variables aleatorias en $B_p - \{0\}$. El coeficiente de correlación es una función de $(B_p - \{0\}) \times (B_p - \{0\})$ en \mathbb{R} denotado por ρ_p y dado por

$$\begin{aligned} \rho_p : (B_p - \{0\}) \times (B_p - \{0\}) &\rightarrow \mathbb{R} \\ (u, v) &\mapsto \rho_p(u, v) = \frac{\langle u, v \rangle_p}{\|u\|_{B_p} \|v\|_{B_p}} . \end{aligned}$$

Diremos que dos variables aleatorias u y v están en la misma dirección si existe un escalar α tal que $u = \alpha v$, esto es, si u es un múltiplo escalar de v , también se dirá que las variables aleatorias u y v son linealmente dependientes.

Teorema 3.28. Sean u y v variables aleatorias en $B_p - \{0\}$ entonces

- i. $\rho_p(u, v) = \rho_p(v, u)$.
- ii. Si \hat{u} y \hat{v} son variables aleatorias de norma 1 en la dirección de u y v respectivamente, entonces el coeficiente de correlación no cambia

$$\rho_p(\hat{u}, \hat{v}) = \rho_p(u, v) .$$

- iii. $|\rho_p(u, v)| \leq 1$.
- iv. $\rho_p(u, u) = 1$.
- v. $\rho_p(u, -u) = -1$.
- vi. Si $a \in \mathbb{R}$ entonces $\rho_p(au, v) = \frac{a}{|a|} \rho_p(u, v)$.

Demostración

- i. Como el producto escalar es conmutativo resulta

$$\rho_p(u, v) = \frac{\langle u, v \rangle_p}{\|u\|_{B_p} \|v\|_{B_p}} = \frac{\langle v, u \rangle_p}{\|v\|_{B_p} \|u\|_{B_p}} = \rho_p(v, u) .$$

- ii. Sean u y v en $B_p - \{0\}$, así $\|u\|_{B_p} \neq 0$ y $\|v\|_{B_p} \neq 0$. Consideremos las variables aleatorias $\hat{u} = \frac{1}{\|u\|_{B_p}} u$ y $\hat{v} = \frac{1}{\|v\|_{B_p}} v$. Tanto \hat{u} como \hat{v} están en B_p ya que $E_p[\hat{u}] = \frac{1}{\|u\|_{B_p}} E_p[u] = 0$, de forma análoga $E_p[\hat{v}] = 0$. Veamos ahora que $\|\hat{u}\|_{B_p} = 1$ y $\|\hat{v}\|_{B_p} = 1$, por definición de la norma B_p se tiene

$$\|\hat{v}\|_{B_p}^2 = E_p[\hat{v}^2] = E_p \left[\frac{1}{\|v\|_{B_p}^2} v^2 \right] = \frac{1}{\|v\|_{B_p}^2} E_p[v^2] = \frac{1}{\|v\|_{B_p}^2} \|v\|_{B_p}^2 = 1 ,$$

luego $\|\hat{v}\|_{B_p} = 1$, de igual forma se procede para concluir $\|\hat{u}\|_{B_p} = 1$. Se encontraron así dos variables aleatorias \hat{u} y \hat{v} de norma 1 bajo B_p y que están en la misma dirección de las variables u y v . Como $\|\hat{u}\|_{B_p} = 1 = \|\hat{v}\|_{B_p}$ entonces $\rho_p(\hat{u}, \hat{v}) = \langle \hat{u}, \hat{v} \rangle_p$. Con base en la definición del producto escalar en B_p resulta

$$\begin{aligned} \langle \hat{u}, \hat{v} \rangle_p &= E_p[\hat{u}\hat{v}] = E_p \left[\left(\frac{1}{\|u\|_{B_p}} u \right) \left(\frac{1}{\|v\|_{B_p}} v \right) \right] \\ &= \frac{1}{\|u\|_{B_p} \|v\|_{B_p}} E_p[uv] = \frac{\langle u, v \rangle_p}{\|u\|_{B_p} \|v\|_{B_p}}. \end{aligned}$$

Como $\rho_p(\hat{u}, \hat{v}) = \langle \hat{u}, \hat{v} \rangle_p$ se concluye que $\rho_p(\hat{u}, \hat{v}) = \rho_p(u, v)$.

iii. Por la desigualdad de Cauchy-Schwarz (3.24) se obtiene que

$$|\rho_p(u, v)| = \frac{|\langle u, v \rangle_p|}{\|u\|_{B_p} \|v\|_{B_p}} \leq \frac{\|u\|_{B_p} \|v\|_{B_p}}{\|u\|_{B_p} \|v\|_{B_p}} \leq 1.$$

iv. Para $u \in B_p - \{0\}$ se tiene por la ecuación dada en (3.29) que

$$\rho_p(u, u) = \frac{\langle u, u \rangle_p}{\|u\|_{B_p} \|u\|_{B_p}} = \frac{\|u\|_{B_p}^2}{\|u\|_{B_p}^2} = 1.$$

v. Sea $u \in B_p$, por las propiedades de la norma y el producto escalar,

$$\rho_p(u, -u) = \frac{\langle u, -u \rangle_p}{\|u\|_{B_p} \|-u\|_{B_p}} = \frac{-\langle u, u \rangle_p}{\|u\|_{B_p} \|-u\|_{B_p}} = \frac{-\|u\|_{B_p}^2}{\|u\|_{B_p} \|u\|_{B_p}} = \frac{-\|u\|_{B_p}^2}{\|u\|_{B_p}^2} = -1.$$

vi. Sean $a \in \mathbb{R}$ y $u \in B_p - \{0\}$. Por las propiedades de la norma y el producto escalar resulta

$$\rho_p(au, v) = \frac{\langle au, v \rangle_p}{\|au\|_{B_p} \|v\|_{B_p}} = \frac{a}{|a|} \frac{\langle u, v \rangle_p}{\|u\|_{B_p} \|v\|_{B_p}} = \frac{a}{|a|} \rho_p(u, v).$$

Si a es un real positivo entonces $\rho_p(au, v) = \rho_p(u, v)$ y si a sea un real negativo entonces $\rho_p(au, v) = -\rho_p(u, v)$. \square

El teorema 3.28 permite vislumbrar una correspondencia geométrica con el coeficiente de correlación, para ello consideremos dos vectores x, y en \mathbb{R}^m con $m \in \mathbb{N}$, si θ es el ángulo comprendido entre estos vectores entonces

$$\cos \theta = \frac{x \cdot y}{\|x\| \|y\|}.$$

Donde la norma de la igualdad anterior equivale a la norma euclídea en \mathbb{R}^m . En el teorema 3.28 se tiene que la función coseno está bien definido ya que su rango es el intervalo $[-1, 1]$ ($|\rho_p(u, v)| \leq 1$). Esto motiva la siguiente definición.

Definición 3.17. Sean u y v en $B_p - \{0\}$. Denotamos por θ_p el ángulo comprendido entre las variables aleatorias u y v y para la cual se cumple

$$\cos \theta_p = \rho_p(u, v) = \frac{\langle u, v \rangle_p}{\|u\|_{B_p} \|v\|_{B_p}} .$$

3.7. Gradiente del funcional Acumulante

La función $\phi_3(x) = (1+|x|) \ln(1+|x|) - |x|$ forma un par conjugado con $\phi_2(x) = e^{|x|} - |x| - 1$. Además, la función ϕ_3 cumple la condición Δ_2 (ver preliminares). Sobre el espacio de Orlicz $L^{\phi_3}(p \cdot \mu)$ se define el espacio ${}^*B_p = \{u \in L^{\phi_3}(p \cdot \mu) : E_p[u] = 0\}$, el cual es cerrado y por tanto un espacio Banach. Se relacionan los espacios B_p y *B_p , como ${}^*B_p \subseteq B_p^*$, es decir, el espacio *B_p está contenido en el dual algebraico de B_p (ver [21]).

Por la relación ${}^*B_p \subseteq B_p^*$, el operador $\langle \cdot, \cdot \rangle_{*,p}$, que se define a continuación, tiene la estructura de un producto escalar, sin embargo no lo es ya que está definido sobre espacios de Banach distintos.

Definición 3.18. Se denotará por $\langle \cdot, \cdot \rangle_{*,p}$ la forma bilineal entre los espacios de Banach B_p y su espacio dual B_p^* , donde

$$\begin{aligned} \langle \cdot, \cdot \rangle_{*,p} : B_p^* \times B_p &\rightarrow \mathbb{R} \\ (u^*, u) &\mapsto \langle u^*, u \rangle_{*,p} = u^*(u) . \end{aligned}$$

Se dice que $u^* \in B_p^*$ es ortogonal a $u \in B_p$ si $\langle u^*, u \rangle_{*,p} = 0$.

En los espacios de Hilbert se satisface que si $\langle u, v \rangle = \langle u_1, v \rangle$ para todo v entonces $u = u_1$, donde $\langle \cdot, \cdot \rangle$ es el producto escalar definido sobre dicho espacio. El siguiente lema muestra como esta propiedad se satisface sobre B_p a pesar que el operador $\langle \cdot, \cdot \rangle_{*,p}$ no es un producto escalar.

Lema 3.2. Sean u^* y v^* funcionales tales que $\langle u^*, u \rangle_{*,p} = \langle v^*, u \rangle_{*,p}$ para todo $u \in B_p$ entonces $u^* = v^*$.

Demostración

Sea $u \in B_p$, por hipótesis $\langle u^*, u \rangle_{*,p} = \langle v^*, u \rangle_{*,p}$, luego

$$\langle u^*, u \rangle_{*,p} - \langle v^*, u \rangle_{*,p} = 0 ,$$

por la linealidad respecto de la primera componente se tiene $\langle u^* - v^*, u \rangle_{*,p} = 0$; con base en la definición 3.18 se tiene $(u^* - v^*)(u) = 0$ de donde $u^*(u) = v^*(u)$. Como $u \in B_p$ es arbitrario entonces se cumple que $u^* = v^*$ para todo $u \in B_p$. \square .

Teorema 3.29. *Sea $u \in \mathcal{V}_p$ tal que $q = e^{u-K_p(u)}p$. Si v_1 y v_2 están en B_p entonces*

$$D^2K_p(u)(v_1, v_2) = \text{cov}_q[v_1, v_2].$$

Demostración

Sea $u \in \text{dom}(K_p)$ tal que $q = e^{u-K_p(u)}p$. Para $v_1 \in B_p$ se hace uso de (3.10) para obtener

$$DK_p(u) \cdot v_1 = \frac{DM_p(u) \cdot v_1}{M_p(u)}. \quad (1)$$

La diferencial de orden 2 (D^2) se escribe como

$$D^2K_p(u)(v_1, v_2) = D(DK_p(u) \cdot v_1) \cdot v_2. \quad (2)$$

Combinando (1) y (2) resulta

$$D^2K_p(u)(v_1, v_2) = D \left[\frac{DM_p(u) \cdot v_1}{M_p(u)} \right] \cdot v_2. \quad (3)$$

Como $DM_p(u) \cdot v_1 = E_q[v_1 e^u]$ entonces $DM_p(u) \cdot v_1$ es una función real y así en la expresión (3) se puede derivar como un cociente para tener

$$\begin{aligned} D^2K_p(u)(v_1, v_2) &= \frac{D(DM_p(u) \cdot v_1) \cdot v_2 \cdot M_p(u) - [DM_p(u) \cdot v_1][DM_p(u) \cdot v_2]}{M_p^2(u)} \\ &= \frac{[D^2M_p(u)(v_1, v_2)] \cdot M_p(u) - [DM_p(u) \cdot v_1][DM_p(u) \cdot v_2]}{M_p^2(u)}. \end{aligned}$$

Debido a que $D^2M_p(u)(v_1, v_2) = E_p[v_1 v_2 e^u]$ (ver teorema 3.6) y $DM_p(u) \cdot v_i = M_p(u)E_q[v_i]$ para $i = 1, 2$ entonces, la igualdad anterior se puede escribir como

$$D^2K_p(u)(v_1, v_2) = \frac{E_p[v_1 v_2 e^u] M_p(u) - M_p^2(u) E_q[v_1] E_q[v_2]}{M_p^2(u)}. \quad (4)$$

Puesto que $q = e^{u-K_p(u)}p$ entonces $e^u p = M_p(u)q$. Así

$$E_p[v_1 v_2 e^u] = \int_{\Omega} v_1 v_2 e^u p d\mu = \int_{\Omega} v_1 v_2 M_p(u) q d\mu = M_p(u) \int_{\Omega} v_1 v_2 q d\mu = M_p(u) E_q[v_1 v_2],$$

sustituyendo en (4) se obtiene

$$D^2K_p(u)(v_1, v_2) = \frac{E_q[v_1v_2]M_p^2(u) - M_p^2(u)E_q[v_1]E_q[v_2]}{M_p^2(u)} = E_q[v_1v_2] - E_q[v_1]E_q[v_2].$$

Haciendo uso de la igualdad 3.24 se puede concluir que $D^2K_p(u)(v_1, v_2) = cov_q[v_1, v_2]$. \square

El siguiente teorema permite deducir el gradiente del funcional acumulante, el cual se denotará $\nabla K_p(u)$; por definición dicho funcional es lineal y además satisface que $Df(u)v = \nabla f(u) \cdot v$, en términos del operador $\langle \cdot, \cdot \rangle_{*,p}$

$$DK_p(u) \cdot v = \langle \nabla K_p(u), v \rangle_{*,p}, \quad (3.32)$$

donde $v \in B_p$ y ∇K_p es un elemento B_p^* ; por lo que el operador $\langle \cdot, \cdot \rangle_{*,p}$ está bien definido.

Teorema 3.30. *Para todo $u \in \mathcal{V}_p$ y $q = e^{u-K_p(u)}p$ se tiene que*

$$DK_p(u) \cdot v = E_p \left[\left(\frac{q}{p} - 1 \right) v \right].$$

En otras palabras, el diferencial de K_p en u , $DK_p(u)$ está en B_p^* identificado con un elemento de *B_p denotado por $\nabla K_p(u)$ y dado por

$$\nabla K_p(u) = \frac{q}{p} - 1 = e^{u-K_p(u)} - 1.$$

Demostración

Sea $u \in \mathcal{V}_p$, por hipótesis $q = e^{u-K_p(u)}p$. En [5] se demuestra que $\frac{q}{p} - 1$ es un elemento de *B_p . Por las propiedades de la integral resulta

$$E_p \left[\left(\frac{q}{p} - 1 \right) v \right] = E_p \left[\frac{q}{p} v \right] - E_p[v] = \int_{\Omega} \frac{q}{p} v p d\mu = \int_{\Omega} v q d\mu = E_q[v].$$

Como $DK_p(u) \cdot v = E_q[v]$ entonces por transitividad con la expresión antes encontrada se tiene

$$DK_p(u) \cdot v = E_p \left[\left(\frac{q}{p} - 1 \right) v \right]. \quad (1)$$

Ya que $\left(\frac{q}{p} - 1 \right) \in {}^*B_p \subset B_p^*$ entonces la igualdad (1) se puede escribir como

$$DK_p(u) \cdot v = \left\langle \frac{q}{p} - 1, v \right\rangle_{*,p}.$$

Por transitividad con la igualdad dada en (3.32) se tiene

$$\langle \nabla K_p(u), v \rangle_{*,p} = \left\langle \frac{q}{p} - 1, v \right\rangle_{*,p}.$$

Debido a que $v \in B_p$ es arbitrario entonces, por el lema 3.2 se tiene que $\nabla K_p(u) = \frac{q}{p} - 1 = e^{u-K_p(u)} - 1$. \square

Como consecuencia del teorema 3.30 tenemos la igualdad

$$DK_p(u) \cdot v = E_p[\nabla K_p(u)v]. \quad (3.33)$$

El siguiente teorema, cuya demostración se encuentra en [21], indica que el gradiente del funcional acumulante es uno a uno, se obtiene también la derivada del gradiente, deducción que se encuentra en la misma referencia bibliográfica.

Teorema 3.31. *El operador gradiente*

$$\begin{aligned} \nabla K_p(\cdot) : B_p &\rightarrow B_p^* \\ u &\mapsto \nabla K_p(u) = e^{u-K_p(u)} - 1 \end{aligned}$$

es uno a uno.

Es posible derivar el gradiente del funcional acumulante en la dirección $w \in B_p$, $D(\nabla K_p(u)) \cdot w$, esta derivada es uno a uno también.

Teorema 3.32. *La derivada del mapeo gradiente aplicado en $w \in B_p$ está dado por*

$$D(\nabla K_p(u)) \cdot w = \frac{q}{p}(w - E_q[w]).$$

Teorema 3.33. *El diferencial de $\nabla K_p(u)$ es uno a uno.*

Demostración

Sean $u \in \text{dom}(K_p)$ y w, w_1 en B_p tales que $D(\nabla K_p(u)) \cdot w = D(\nabla K_p(u)) \cdot w_1$; por el teorema 3.32 se escriben ambos diferenciales como

$$\frac{q}{p}(w - E_q[w]) = \frac{q}{p}(w_1 - E_q[w_1]),$$

de donde $w - w_1 = E_q[w] - E_q[w_1]$, como la diferencia del lado derecho de la igualdad anterior es constante, escribimos $w - w_1 = r$, aplicando el operador $E_p[\cdot]$ a ambos lados se tiene $E_p[w] - E_p[w_1] = E_p[r]$; ya que r es constante respecto de la medida $p \cdot \mu$ y $E_p[w] = E_p[w_1] = 0$ entonces $r = 0$. Como $w - w_1 = r$ entonces $w = w_1$. El diferencial de $\nabla K_p(u)$ es un mapeo uno a uno. \square

3.8. Espacio Tangente

Sean $p \in \mathfrak{M}_\mu$ y $p(t)$ una curva a través de p que es diferenciable en \mathfrak{M}_μ , esto es, un modelo estadístico unidimensional paramétrico $p(t)$, para $t \in I$ donde I es un abierto en \mathbb{R} y $p(t_0) = p$. Cada curva tiene un modelo tangente de la forma exponencial $e^{tu - K_p(tu)} p$. Esto identifica el espacio tangente con el conjunto de modelos exponenciales unidimensionales.

Si (\mathcal{U}_q, s_q) es una carta para p , entonces

$$p(t) = e^{u(t) - K_q(u(t))} q,$$

donde $u(t) = s_q(p(t))$. Con respecto a esta carta, el vector tangente a la curva en p es $\dot{u}(t_0)$. En una carta diferente $(\mathcal{U}_{q_1}, s_{q_1})$, el vector tangente es $\dot{u}_1(t_0)$. Por el mapeo transición, las variables $u(t)$ y $u_1(t)$ en B_q y B_{q_1} respectivamente (donde $u(t_0) = s_q(p)$ y $u_1(t_0) = s_{q_1}(p)$), están relacionadas a través de la expresión $(s_{q_1} \circ e_q)(u(t_0)) = u_1(t)$, y al derivar resulta $(s_{q_1} \circ e_q)'(\dot{u}(t_0)) = \dot{u}_1(t_0)$ que determina una relación de equivalencia.

La colección de los vectores tangentes en p , a través de la relación de equivalencia $(s_{q_1} \circ e_q)'(\dot{u}(t_0)) = \dot{u}_1(t_0)$, es llamado el espacio tangente en p de la variedad \mathfrak{M}_μ , que se denota $T_p(\mathfrak{M}_\mu)$. El espacio tangente es un espacio vectorial, la definición no depende de la escogencia de las cartas.

Considere una familia parametrizada de funciones de densidad de probabilidad y la aplicación

$$x \mapsto f(x | \theta)$$

donde θ es el vector de parámetros y x los observables. Si se fijan los observables x entonces la función de verosimilitud es

$$\theta \mapsto f(x | \theta),$$

denotada, generalmente, como $l(\theta)$, $l(\theta) = f(x | \theta)$. La función soporte es el logaritmo, usualmente natural, de la función de verosimilitud, se escribe $L(\theta) = \ln(l(\theta))$. La función score es el gradiente de la función soporte, denotada V ,

$$V = \frac{\partial}{\partial \theta} L(\theta) = \frac{\partial}{\partial \theta} \ln(l(\theta)) = \frac{\partial}{\partial \theta} \ln(f(x | \theta))$$

Teorema 3.34. Espacio tangente Sea $p(t)$ una curva regular en \mathfrak{M}_μ con $p(t_0) = p$ y sea $u(t) \in B_q$ su representación en una carta s_q , donde $t \in \{t : u(t) \in (\text{dom}(K_q))^\circ\}$. Entonces $p(t) = e^{u(t)-K_q(u(t))}q$.

1. La relación entre $\dot{u}(t_0)$, el espacio tangente a $u(t)$ en t_0 y la función score de $p(t)$ con respecto a la densidad p es

$$\frac{d}{dt} \left[\ln \left(\frac{p(t)}{p} \right) \right]_{t=t_0} = \dot{u}(t_0) - E_p[\dot{u}(t_0)].$$

Si $p = q$, es decir, si la carta está centrada en el mismo punto donde es calculada la log-probabilidad entonces

$$\frac{d}{dt} \left[\ln \left(\frac{p(t)}{p} \right) \right]_{t=t_0} = \dot{u}(t_0).$$

2. La curva

$$t \mapsto \frac{p(t)}{p} - 1$$

está en *B_p y su derivada en t_0 es $\dot{u}(t_0)$.

3. La función score de cualquier modelo exponencial unidimensional a través de p , es decir,

$$e^{tu-K_p(tu)}p,$$

en $t = t_0$, es u . Y recíprocamente, cualquier $u \in B_p$ tiene una correspondencia con un modelo exponencial.

Demostración

1. Por hipótesis $p(t) = e^{u(t)-K_q(u(t))}q$. Como $p(t_0) = p$ entonces $p = e^{u(t_0)-K_q(u(t_0))}q$, de donde

$$\frac{p(t)}{p} = \frac{e^{u(t)-K_q(u(t))}}{e^{u(t_0)-K_q(u(t_0))}} = e^{u(t)-u(t_0)-K_q(u(t))+K_q(u(t_0))},$$

lo cual implica

$$\ln \left(\frac{p(t)}{p} \right) = u(t) - u(t_0) - K_q(u(t)) + K_q(u(t_0)). \quad (1)$$

Como t es fijo entonces $u(t_0)$ y $K_q(u(t_0))$ son constante respecto a la derivada de t . Además, $\frac{d}{dt}K_q(u(t)) = DK_q(u(t)) \cdot \dot{u}(t)$. Al derivar la expresión (1) y evaluar en t_0 , se tiene

$$\frac{d}{dt} \left[\ln \left(\frac{p(t)}{p} \right) \right]_{t=t_0} = \frac{d}{dt} [u(t) - u(t_0) - K_q(u(t)) + K_q(u(t_0))]_{t=t_0}$$

$$=[\dot{u}(t) - DK_q(u(t)) \cdot \dot{u}(t)]_{t=t_0} = \dot{u}(t_0) - DK_q(u(t_0)) \cdot \dot{u}(t_0) .$$

En el corolario 3.1 se demostró que si $q = e_p(u)$ entonces $DK_p(u)v = E_q[v]$ donde $v \in B_p$, en esta situación se tiene que $p = p(t_0) = e_q(u(t_0))$ por lo que $DK_q(u(t_0)) \cdot \dot{u}(t_0) = E_p[\dot{u}(t_0)]$. Se concluye que

$$\frac{d}{dt} \left[\ln \left(\frac{p(t)}{p} \right) \right]_{t=t_0} = \dot{u}(t_0) - E_p[\dot{u}(t_0)] ,$$

donde $\dot{u}(t) \in B_q$. Veamos que $\dot{u}(t) \in B_q$, como $u(t) \in B_q$ entonces $E_q[u(t)] = 0$, por la regla de Leibnitz se concluye que $E_q[\dot{u}(t)] = 0$, en particular para $t = t_0$, $E_q[\dot{u}(t_0)] = 0$. En el caso en que $q = p$ entonces $E_p[\dot{u}(t_0)] = E_q[\dot{u}(t_0)] = 0$ y por lo tanto

$$\frac{d}{dt} \left[\ln \left(\frac{p(t)}{p} \right) \right]_{t=t_0} = \dot{u}(t_0) .$$

2. En el teorema 3.30 se concluyó que si $q = e^{u-K_p(u)}p$ entonces $\frac{q}{p} - 1$ es un elemento en *B_p que se identifica con el $\nabla K_p(u)$ que es un elemento en B_p^* . De manera análoga, $\frac{p(t)}{p} - 1$ se identifica con $\nabla K_p(u(t))$, se escribe $\frac{p(t)}{p} - 1 = \nabla K_p(u(t))$. La derivada respecto de t en t_0 es

$$\frac{d}{dt} \left(\frac{p(t)}{p} - 1 \right)_{t=t_0} = D\nabla K_p(u(t)) \cdot \dot{u}(t) |_{t=t_0} .$$

Por el teorema 3.32 resulta

$$\frac{d}{dt} \left(\frac{p(t)}{p} - 1 \right)_{t=t_0} = \left[\frac{p(t)}{p} (\dot{u}(t) - E_{p(t)}[\dot{u}(t)]) \right]_{t=t_0} = \frac{p(t_0)}{p} (\dot{u}(t_0) - E_{p(t_0)}[\dot{u}(t_0)]) .$$

Como $p(t_0) = p$ y $E_p[\dot{u}(t_0)] = 0$ entonces

$$\frac{d}{dt} \left(\frac{p(t)}{p} - 1 \right)_{t=t_0} = \dot{u}(t_0) .$$

3. Derivando respecto de t la expresión $tu - K_p(tu)$ resulta

$$\frac{d}{dt} (tu - K_p(tu))_{t=t_0} = \left[u - \frac{d}{dt} K_p(tu) \right]_{t=t_0} = [u - DK_p(tu) \cdot u]_{t=t_0} = u - DK_p(t_0 u) \cdot u$$

Por el corolario 3.1 se llega a $\frac{d}{dt} (tu - K_p(tu))_{t=t_0} = u - E_q[u] = u$. \square

Como consecuencia del literal 3 del teorema 3.34, el espacio tangente del modelo exponencial unidimensional es otra representación del espacio tangente $T_p(\mathfrak{M}_\mu)$.

El espacio tangente hereda la estructura de espacio vectorial y la topología desde B_p . Como el producto escalar está definido en B_p , por la definición 3.14, está definido el producto escalar en $T_p(\mathfrak{M}_\mu)$ junto con la ortogonalidad.

Capítulo 4

Modelos k-exponenciales según G. Pistone

4.1. Introducción

En este capítulo se presentan los resultados expuestos por Giovanni Pistone en el artículo “*k-exponential models from the geometrical viewpoint*” [23], en el cual se construye una estructura de variedad diferenciable para \mathfrak{M}_μ , donde los modelos *k*-exponenciales son de la forma $q = e_k^{u - D_k(q||p)} p$, siendo $D_k(p || q)$ la *k*-divergencia entre las densidades p y q según la teoría de G. Kaniadakis [14, 15]. La construcción presentada por Pistone se hace a través de espacios de Lebesgue de la forma $L^{1/k}$, donde $0 < k < 1$.

4.2. Cartas

Sea $0 < k < 1$ fijo. Con el fin de construir las cartas del atlas, se construyen biyecciones desde el conjunto de densidades estrictamente positivas \mathfrak{M}_μ a algún espacio vectorial modelador; cada carta será asociada a una densidad $p \in \mathfrak{M}_\mu$. Tal densidad p se utiliza como referencia para cualquier otra densidad q de un subconjunto adecuado de \mathfrak{M}_μ . La teoría no está restringida a espacios de dimensión finita.

Primero se define una *k*-divergencia. Si $p, q \in \mathfrak{M}_\mu$ satisfacen la condición

$$\left(\frac{q}{p}\right)^k, \left(\frac{p}{q}\right)^k \in L^1(p \cdot \mu), \quad (4.1)$$

entonces la *k*-divergencia de p respecto de q , está definida como

$$D_k(p || q) = E_p \left[\ln_k \left(\frac{p}{q} \right) \right]. \quad (4.2)$$

Por la definición del logaritmo k -deformado, la divergencia se escribe como

$$D_k(p \parallel q) = \frac{1}{2k} E_p \left[\left(\frac{p}{q} \right)^k - \left(\frac{q}{p} \right)^k \right]. \quad (4.3)$$

La segunda de las dos condiciones en (4.1) siempre se cumple; en efecto, la función x^k es cóncava para $x > 0$ y $0 < k < 1$, ya que si $f(x) = x^k$ entonces $f''(x) = k(k-1)x^{k-2} < 0$ puesto que $k-1 < 0$. De la cóncavidad de la función x^k se sigue que $-x^k$ es convexa, por lo que la desigualdad de Jensen se verifica y así

$$- \left(\int_{\Omega} \frac{q}{p} p d\mu \right)^k \leq \int_{\Omega} - \left(\frac{q}{p} \right)^k p d\mu.$$

Al multiplicar por -1 , resulta

$$\int_{\Omega} \left(\frac{q}{p} \right)^k p d\mu \leq \left(\int_{\Omega} \frac{q}{p} p d\mu \right)^k = \left(\int_{\Omega} q d\mu \right)^k = 1,$$

así $E_p \left[\left(\frac{q}{p} \right)^k \right] < \infty$ y por tanto $\left(\frac{q}{p} \right)^k \in L^1(p \cdot \mu)$. La primer condición no es trivial, por que se debe asumir que

$$E_p \left[\left(\frac{p}{q} \right)^k \right] = \int_{\Omega} \left(\frac{p}{q} \right)^k p d\mu = \int_{\Omega} \frac{p^{k+1}}{q^k} d\mu < \infty. \quad (4.4)$$

Cuando tal condición no se satisface, el valor esperado en (4.3) es infinito, que no es el interés. Ahora bien, por la convexidad de la función $-\ln_k(\cdot)$ y por la desigualdad de Jensen, se sigue que

$$-\ln_k \left(\int_{\Omega} \frac{q}{p} p d\mu \right) \leq \int_{\Omega} -\ln_k \left(\frac{q}{p} \right) p d\mu,$$

equivalente a $-\ln_k(1) \leq E_p \left[\ln_k \left(\frac{p}{q} \right) \right]$. Por la definición de la k -divergencia en (4.2), se tiene $D_k(p \parallel q) \geq 0$, donde la igualdad ocurre si y sólo si $p = q$.

La variedad se define a través de espacios de Lebesgue centrados de la forma $(\frac{1}{k} - p)$ -integrable, donde una variable aleatoria pertenece a $L_0^{1/k}(p)$ si y sólo si

1. $E_p [|v|^{1/k}] = \int |v|^{1/k} p d\mu < \infty$,
2. $E_p[v] = 0$.

Para cada densidad p hay un espacio modelador diferente, de modo que debe proporcionarse un isomorfismo para dos densidades en \mathfrak{M}_μ . La más simple identificación isométrica entre $L^{1/k}(p_1 \cdot \mu)$ y $L^{1/k}(p_2 \cdot \mu)$ es

$$\begin{aligned} L^{1/k}(p_1 \cdot \mu) &\rightarrow L^{1/k}(p_2 \cdot \mu) \\ u &\mapsto \left(\frac{p_1}{p_2}\right)^k u, \end{aligned} \quad (4.5)$$

donde p_1 y p_2 son dos densidades de probabilidad en \mathfrak{M}_μ . Veamos que el mapeo está bien definido, consideremos una variable aleatoria u en $L^{1/k}(p_1 \cdot \mu)$, lo cual implica que $E_{p_1} [|u|^{1/k}] < \infty$. Por la definición del mapeo (4.5) resulta

$$\begin{aligned} E_{p_2} \left[\left| \left(\frac{p_1}{p_2}\right)^k u \right|^{1/k} \right] &= \int_{\Omega} \left| \left(\frac{p_1}{p_2}\right)^k u \right|^{1/k} p_2 d\mu = \int_{\Omega} \frac{p_1}{p_2} |u|^{1/k} p_2 d\mu \\ &= \int_{\Omega} |u|^{1/k} p_1 d\mu = E_{p_1} [|u|^{1/k}]. \end{aligned}$$

Se concluye que $E_{p_2} \left[\left| \left(\frac{p_1}{p_2}\right)^k u \right|^{1/k} \right] < \infty$, y así la variable aleatoria $\left(\frac{p_1}{p_2}\right)^k \cdot u$ es un elemento en $L^{1/k}(p_2 \cdot \mu)$; y así el mapeo (4.5) está bien definido. El mapeo inverso de (4.5) está dado por

$$\begin{aligned} L^{1/k}(p_2 \cdot \mu) &\rightarrow L^{1/k}(p_1 \cdot \mu) \\ u &\mapsto \left(\frac{p_2}{p_1}\right)^k u. \end{aligned} \quad (4.6)$$

Por la definición de la norma en los espacios de Lebesgue, resulta la expresión

$$\begin{aligned} \left\| \left(\frac{p_1}{p_2}\right)^k u \right\|_{L_{p_2}^{1/k}} &= \left(\int_{\Omega} \left| \left(\frac{p_1}{p_2}\right)^k u \right|^{1/k} p_2 d\mu \right)^k \\ &= \left(\int_{\Omega} |u|^{1/k} p_1 d\mu \right)^k = \|u\|_{L_{p_1}^{1/k}}^k. \end{aligned}$$

De la igualdad anterior entre las normas, se sigue que la identificación definida en (4.5) es isométrica.

Se hará una variación del formalismo usado en el caso exponencial usual [22]. La nueva variedad será llamada *La variedad k-estadística*. Se define el subconjunto \mathcal{E}_p de \mathfrak{M}_μ por

$$\mathcal{E}_p = \left\{ q \in \mathfrak{M}_\mu : \left(\frac{q}{p}\right)^k, \left(\frac{p}{q}\right)^k \in L^{1/k}(p \cdot \mu) \right\}. \quad (4.7)$$

Las condiciones $\left(\frac{q}{p}\right)^k, \left(\frac{p}{q}\right)^k \in L^{1/k}(p \cdot \mu)$ son equivalentes a $\frac{q}{p}, \frac{p}{q} \in L^1(p \cdot \mu)$, en efecto, si consideramos que $\left(\frac{q}{p}\right)^k \in L^{1/k}(p \cdot \mu)$ entonces, por la definición de los espacios de Lebesgue se escribe

$$E_p \left[\left| \left(\frac{q}{p}\right)^k \right|^{1/k} \right] = \int_{\Omega} \left| \left(\frac{q}{p}\right)^k \right|^{1/k} p \cdot d\mu = \int_{\Omega} \left| \frac{q}{p} \right| p \cdot d\mu = E_p \left[\left| \frac{q}{p} \right| \right] < \infty ,$$

lo que implica que $\frac{q}{p} \in L^1(p \cdot \mu)$, situación análoga para $\frac{p}{q} \in L^1(p \cdot \mu)$. Con base en estos resultados, el conjunto \mathcal{E}_p definido en 4.7 se representa también como

$$\mathcal{E}_p = \left\{ q \in \mathfrak{M}_{\mu} : \frac{q}{p}, \frac{p}{q} \in L^1(p \cdot \mu) \right\} .$$

La expresión $\frac{q}{p} \in L^1(p \cdot \mu)$ siempre es cierta ya que $E_p \left[\frac{q}{p} \right] = E_q[1] = 1 < \infty$ entonces el conjunto \mathcal{E}_p puede restringirse a

$$\mathcal{E}_p = \left\{ q \in \mathfrak{M}_{\mu} : \frac{p}{q} \in L^1(p \cdot \mu) \right\} .$$

Luego, una densidad q está en \mathcal{E}_p siempre que $E_p \left[\frac{p}{q} \right] < \infty$, esto es, si

$$E_p \left[\frac{p}{q} \right] = \int_{\Omega} \frac{p^2}{q} d\mu < \infty . \quad (4.8)$$

La condición (4.8) es más fuerte que la condición (4.4); por la desigualdad de Jensen se tiene que

$$E_p \left[\left(\frac{p}{q}\right)^k \right] \leq \left(E_p \left[\frac{p}{q} \right] \right)^k ,$$

esto debido a la convexidad de la función $-x^k$. Por propiedades algebraicas, la desigualdad anterior se escribe como

$$\int_{\Omega} \frac{p^{k+1}}{q^k} d\mu \leq \left(\int_{\Omega} \frac{p^2}{q} d\mu \right)^k .$$

Luego, si $\int \frac{p^2}{q} d\mu < \infty$ entonces $\int \frac{p^{k+1}}{q^k} d\mu < \infty$; por lo que la condición (4.8) implica la condición (4.4) y por ello que sea más fuerte, además que esta condición (4.8) no depende de k . Por lo tanto

$$\mathcal{E}_p = \left\{ q \in \mathfrak{M}_{\mu} : \frac{p}{q} \in L^1(p \cdot \mu) \right\} , \quad (4.9)$$

es independiente de k . De hecho, se está asumiendo que la k -divergencia $D_k(p \parallel q)$ exista. Cada uno de los subconjuntos \mathcal{E}_p , para $p \in \mathfrak{M}_{\mu}$, va a ser el dominio de una carta, y éstas

definirán un atlas de cartas debido al cubrimiento $\mathfrak{M}_\mu = \bigcup_{p \in \mathfrak{M}_\mu} \mathcal{E}_p$, una componente conexa de la variedad será la unión de superposiciones de conjuntos \mathcal{E}_p .

Sea $q \in \mathcal{E}_p$, entonces q es positiva casi en todas partes de Ω y puede ser escrita como

$$q = \exp_{\{k\}}(v) p = e_k^v p. \quad (4.10)$$

Despejando v resulta

$$v = \ln_k \left(\frac{q}{p} \right) = \frac{1}{2k} \left[\left(\frac{q}{p} \right)^k - \left(\frac{p}{q} \right)^k \right]. \quad (4.11)$$

Como $q \in \mathcal{E}_p$, entonces $\left(\frac{q}{p} \right)^k$ y $\left(\frac{p}{q} \right)^k$ son elementos en $L^{1/k}(p \cdot \mu)$ por (4.7), esto implica que $\ln_k \left(\frac{q}{p} \right)$ es también un elemento en $L^{1/k}(p \cdot \mu)$ y, por la igualdad (4.11), $v \in L^{1/k}(p \cdot \mu)$. El valor esperado de v en p es

$$E_p[v] = E_p \left[\ln_k \left(\frac{q}{p} \right) \right] = -E_p \left[\ln_k \left(\frac{p}{q} \right) \right].$$

En términos de la k -divergencia $E_p[v] = -D_k(p \parallel q)$; de modo que la variable aleatoria $u = v + D_k(p \parallel q)$ es un elemento en $L^{1/k}(p \cdot \mu)$ y es centrada debido a que

$$\begin{aligned} E_p[u] &= E_p[v + D_k(p \parallel q)] = E_p[v] + D_k(p \parallel q) \\ &= -D_k(p \parallel q) + D_k(p \parallel q) = 0. \end{aligned}$$

Luego, $v = u - D_k(p \parallel q)$, y por tanto en (4.10) se tiene la igualdad

$$q = e_k^{u - D_k(p \parallel q)} p, \quad (4.12)$$

donde u es un elemento definido de forma única en el conjunto de variables aleatorias centradas $L_0^{1/k}(p \cdot \mu)$. Además, v equivale a $v = \ln_k \left(\frac{q}{p} \right)$, de donde

$$u = \ln_k \left(\frac{q}{p} \right) + D_k(p \parallel q). \quad (4.13)$$

Como el valor esperado de v es igual a $-D_k(p \parallel q)$ entonces, $D_k(p \parallel q) = -E_p[v] = -E_p \left[\ln_k \left(\frac{q}{p} \right) \right]$, luego la expresión (4.13) se escribe como

$$u = \ln_k \left(\frac{q}{p} \right) - E_p \left[\ln_k \left(\frac{q}{p} \right) \right], \quad (4.14)$$

que permite definir los mapeos inversos entre \mathfrak{M}_μ y el espacio modelador que tiene la estructura $L^{1/k}$. En esta situación se ha logrado hacer una identificación de una densidad

$q \in \mathcal{E}_p$ con una variable aleatoria centrada $u \in L_0^{1/k}(p \cdot \mu)$, por medio de la expresión $q = e_k^{u - D_k(p \| q)} p$. Se considera ahora el caso contrario. Sea u una variable aleatoria centrada; la función real valuada

$$\psi \longmapsto E_p \left[e_k^{u - \psi} \right], \quad (4.15)$$

es continua por ser la composición de funciones continuas, además es decreciente, ya que si $\psi_1 < \psi_2$ entonces, $-\psi_2 < -\psi_1$ por lo que $u - \psi_2 < u - \psi_1$ y al ser la función k -exponencial creciente y la integral de Lebesgue monótona resulta $E_p \left[e_k^{u - \psi_2} \right] < E_p \left[e_k^{u - \psi_1} \right]$. Luego, el mapeo (4.15) es decreciente y es continuo, por lo que es uno a uno. Así, existe un único valor de ψ , llámese $\psi_{k,p}(u)$, tal que

$$E_p \left[e_k^{u - \psi_{k,p}(u)} \right] = 1. \quad (4.16)$$

Si se elige

$$q = e_k^{u - \psi_{k,p}(u)} p, \quad (4.17)$$

entonces $\int q d\mu = 1$; de allí que q es una densidad de probabilidad, $q \in \mathcal{E}_p \subseteq \mathfrak{M}_\mu$. El mapeo uno a uno

$$\begin{aligned} \mathcal{E}_p &\rightarrow L_0^{1/k}(p \cdot \mu) \\ q &\mapsto u \end{aligned} \quad (4.18)$$

es una carta para \mathfrak{M}_u . Por comparación entre las expresiones 4.12 y 4.17 se obtiene que $\psi_{k,p}(u) = D_k(p \| q)$, el dominio de la función $\psi_{k,p}$ es todo el espacio de las variables aleatorias centradas, $dom(\psi_{k,p}) = L_0^{1/k}(p \cdot \mu)$. Donde u es la imagen de q en la carta definida para p . La divergencia $D_k(q \| p)$ de q con respecto a p resulta

$$D_k(q \| p) = E_q \left[\ln_k \left(\frac{q}{p} \right) \right] = E_q[u - \psi_{k,p}(u)] = E_q[u] - \psi_{k,p}(u).$$

Al ser $\psi_{k,p}(u) = D_k(p \| q)$, se obtiene la expresión

$$D_k(p \| q) + D_k(q \| p) = E_q[u]. \quad (4.19)$$

El funcional $\psi_{k,p}$ es de mucha importancia cuando $k = 0$ donde es el funcional acumulante $K_p(u)$ de la variable aleatoria u y está dado por

$$\psi_p(u) = K_p(u) = \ln(E_p[e^u]).$$

En el caso en que $k = 0$ se logra que $\psi_{0,p} = \psi_p$, en efecto, cuando $k = 0$ la exponencial k -deformada es una exponencial usual, en (4.17) resulta $q = e^{u - \psi_{0,p}(u)} p$; como q es una densidad de probabilidad entonces $E_p[e^{u - \psi_{0,p}(u)}] = 1$, que es equivalente a

$$E_p \left[\frac{e^u}{e^{\psi_{0,p}(u)}} \right] = 1 \quad \Leftrightarrow \quad \frac{E_p[e^u]}{e^{\psi_{0,p}(u)}} = 1.$$

Por lo que $e^{\psi_{0,p}(u)} = E_p[e^u]$ y al aplicar logaritmo usual se concluye $\psi_{0,p}(u) = \ln(E_p[e^u])$, que es el acumulante usual $\psi_{0,p}(u) = \psi_p(u) = K_p(u)$.

Consideremos que $e_{p,k}(u) = e_k^{u-\psi_{k,p}(u)} p$ es el mapeo de $L_0^{1/k}(p \cdot \mu)$ en \mathcal{E}_p , mientras que su mapeo inverso está representado como $s_{p,k}(q) = \ln_k \left(\frac{q}{p} \right) - E_p \left[\ln_k \left(\frac{q}{p} \right) \right]$. Para el cambio de cartas se asume que $u \in L_0^{1/k}(p_1 \cdot \mu)$ y $\bar{u} \in L_0^{1/k}(p_2 \cdot \mu)$ donde, p_1 y p_2 son dos densidades en \mathfrak{M}_μ , además $\mathcal{E}_{p_1} \cap \mathcal{E}_{p_2} \neq \emptyset$. Sea $q \in \mathcal{E}_{p_1} \cap \mathcal{E}_{p_2}$, el mapeo transición se expresa como

$$\begin{aligned} L_0^{1/k}(p_1 \cdot \mu) &\rightarrow \mathcal{E}_{p_1} \cap \mathcal{E}_{p_2} \rightarrow L_0^{1/k}(p_2 \cdot \mu) \\ u &\mapsto q \mapsto \bar{u}, \end{aligned}$$

donde $\bar{u} = (s_{p_2,k} \circ e_{p_1,k})(u)$. Por las propiedades del logaritmo k -deformado se logra

$$\begin{aligned} \bar{u} &= \ln_k \left(\frac{e_{p_1,k}(u)}{p_2} \right) - E_{p_2} \left[\ln_k \left(\frac{e_{p_1,k}(u)}{p_2} \right) \right] \\ &= \ln_k \left(\frac{e_k^{u-\psi_{k,p_1}(u)} p_1}{p_2} \right) - E_{p_2} \left[\ln_k \left(\frac{e_k^{u-\psi_{k,p_1}(u)} p_1}{p_2} \right) \right] \\ &= \ln_k \left(e_k^{u-\psi_{k,p_1}(u)} \right) \oplus^k \ln_k \left(\frac{p_1}{p_2} \right) - E_{p_2} \left[\ln_k \left(e_k^{u-\psi_{k,p_1}(u)} \right) \oplus^k \ln_k \left(\frac{p_1}{p_2} \right) \right]; \end{aligned}$$

con esto se concluye que el mapeo transición es

$$\bar{u} = (u - \psi_{k,p_1}(u)) \oplus^k \ln_k \left(\frac{p_1}{p_2} \right) - E_{p_2} \left[(u - \psi_{k,p_1}(u)) \oplus^k \ln_k \left(\frac{p_1}{p_2} \right) \right]. \quad (4.20)$$

En el caso que $k = 0$, el logaritmo k -deformado, así como la suma k -deformada se transforman en logaritmo y suma usuales, además que $\psi_{0,p}$ equivale al funcional acumulante, por lo que

$$\begin{aligned} \bar{u} &= u - K_p(u) + \ln \left(\frac{p_1}{p_2} \right) - E_{p_2} \left[u - K_p(u) + \ln \left(\frac{p_1}{p_2} \right) \right] \\ &= u + \ln \left(\frac{p_1}{p_2} \right) - E_{p_2} \left[u + \ln \left(\frac{p_1}{p_2} \right) \right], \end{aligned}$$

que son los mapeos transición en la variedad de Pistone usual, los cuales son de clase \mathcal{C}^∞ por ser mapeos afines.

Para hallar la derivada del mapeo transición obtenido en (4.20), se recuerda que

$$\frac{\partial}{\partial x_1} (x_1 \oplus^k x_2) = \sqrt{1 + k^2 x_2^2} + \frac{k^2 x_1 x_2}{\sqrt{1 + k^2 x_1^2}}.$$

Además, la derivada de $u - \psi_{k,p}(u)$ en la dirección de una variable aleatoria centrada v es $D(u - \psi_{k,p}(u))v = v - D\psi_{k,p}(u)v$. En efecto, por la definición de la derivada de Fréchet, se tiene

$$\begin{aligned} D(u - \psi_{k,p}(u))v &= (u + v - \psi_{k,p}(u + v)) - (u - \psi_{k,p}(u)) + R(u - v) \\ &= v - (\psi_{k,p}(u + v) - \psi_{k,p}(u) - R(u - v)) \\ &= v - D\psi_{k,p}(u)v . \end{aligned}$$

Con base en estos resultados es posible derivar $(u - \psi_{k,p}(u)) \oplus_k \ln_k \left(\frac{p_1}{p_2} \right)$ en la dirección de v (variable aleatoria centrada) para tener

$$D \left((u - \psi_{k,p}(u)) \oplus_k \ln_k \left(\frac{p_1}{p_2} \right) \right) v = \sqrt{1 + k^2 \ln_k^2 \left(\frac{p_1}{p_2} \right)} + \frac{k^2 \ln_k \left(\frac{p_1}{p_2} \right) (u - \psi_{k,p}(u))(v - D\psi_{k,p}(u)v)}{\sqrt{1 + k^2 (u - \psi_{k,p}(u))^2}} . \quad (4.21)$$

La expresión del lado derecho en (4.21) se denotará como A . Haciendo uso de la regla de Leibnitz se halla la derivada del mapeo transición definido en (4.20) como

$$D(s_{p_2,k} \circ e_{p_1,k})(u)v = A - E_{p_2}[A] . \quad (4.22)$$

Recordemos en este punto que la derivada de la exponencial k -deformada es

$$\frac{d}{dx} e_k^x = \frac{1}{\sqrt{1 + k^2 x^2}} e_k^x . \quad (4.23)$$

Sea q una densidad expresada como $q = e_k^{u - \psi_{k,p}(u)} p$, de donde $E_p \left[e_k^{u - \psi_{k,p}(u)} \right] = 1$; al diferenciar esta expresión en la dirección de $v \in L_0^{1/k}(p \cdot \mu)$ y hacer uso de la regla de Leibnitz resulta

$$E_p \left[\frac{e_k^{u - \psi_{k,p}(u)}}{\sqrt{1 + k^2 (u - \psi_{k,p}(u))^2}} (v - D\psi_{k,p}(u)v) \right] = 0 .$$

En términos de la integral, resulta la expresión equivalente

$$\int_{\Omega} \frac{e_k^{u - \psi_{k,p}(u)}}{\sqrt{1 + k^2 (u - \psi_{k,p}(u))^2}} (v - D\psi_{k,p}(u)v) p \, d\mu = 0 .$$

Debido a que $q = e_k^{u - \psi_{k,p}(u)} p$, entonces la integral anterior se escribe en términos de q como

$$E_q \left[\frac{v - D\psi_{k,p}(u)v}{\sqrt{1 + k^2 (u - \psi_{k,p}(u))^2}} \right] = 0 . \quad (4.24)$$

Para $u = 0$ se tiene que $\psi_{k,p}(0) = 0$, por lo que en la igualdad (4.24) resulta que $E_q[D\psi_{k,p}(0)v] = 0$; como $D\psi_{k,p}(0)v$ es constante respecto de la medida $q \cdot \mu$ entonces $D\psi_{k,p}(0)v = 0$.

Con base en la ecuación (4.23) se escribe

$$e_k^{u-\psi_{k,p}(u)} = \sqrt{1 + k^2(u - \psi_{k,p}(u))^2} \left(\frac{d}{du} e_k^{u-\psi_{k,p}(u)} \right).$$

El lado izquierdo equivale a $\frac{q}{p}$, además $\ln_k \left(\frac{q}{p} \right) = u - \psi_{k,p}(u)$. La igualdad anterior se escribe

$$\frac{q}{p} = \sqrt{1 + k^2 \ln_k^2 \left(\frac{q}{p} \right)} \left(\frac{d}{du} e_k^{u-\psi_{k,p}(u)} \right),$$

de donde

$$\left(\frac{d}{du} e_k^{u-\psi_{k,p}(u)} \right) p = \frac{q}{\sqrt{1 + k^2 \ln_k^2 \left(\frac{q}{p} \right)}}. \quad (4.25)$$

Para efectos de notación se escribirá $E_{q|p}[\cdot]$ de acuerdo con las expresiones siguientes, las cuales se basan en el resultado hallado en la ecuación (4.25)

$$E_{q|p}[\bar{u}] = \int_{\Omega} \bar{u} \left(\frac{d}{du} e_k^{u-\psi_{k,p}(u)} \right) p d\mu = \int_{\Omega} \bar{u} \frac{q}{\sqrt{1 + k^2 \ln_k^2 \left(\frac{q}{p} \right)}} d\mu. \quad (4.26)$$

A la densidad $q|p$ se le llamará *probabilidad score*, la igualdad (4.24) se escribe a través de la probabilidad score como $E_{q|p}[v - D\psi_{k,p}(u)] = 0$ que es equivalente

$$D\psi_{k,p}(u)v = E_{q|p}[v]. \quad (4.27)$$

En el caso en que $k = 0$ la derivada de la función exponencial es ella misma, por lo que la probabilidad score para una variable aleatoria \bar{u} es

$$E_{q|p}[v] = \int_{\Omega} v e^{u-\psi_{0,p}(u)} p d\mu.$$

Como $q = e^{u-\psi_{0,p}(u)}p$ y además $\psi_{0,p}(u) = K_p(u)$ entonces en la igualdad anterior se llega a $E_{q|p}[v] = \int_{\Omega} v q d\mu = E_q[v]$ y por tanto en (4.27) se obtiene la igualdad $D\psi_{0,p}(u)v = DK_p(u)v = E_q[v]$ que es el resultado obtenido en la variedad de Pistone usual, por lo que la probabilidad score para el caso usual es la derivada del acumulante en la dirección de una variable aleatoria centrada v .

La segunda derivada del mapeo

$$u \longmapsto e_k^{u-\psi_{k,p}(u)} \quad (4.28)$$

en las direcciones de v y w (variables aleatorias centradas) es la primer derivada del mapeo

$$u \longmapsto \left(\frac{d}{du} e_k^{u-\psi_{k,p}(u)} \right) (v - D\psi_{k,p}(u)v), \quad (4.29)$$

que es la derivada en la dirección de v del mapeo definido en (4.28), ahora para derivar el mapeo (4.29) se tiene en cuenta la regla del producto para derivadas en \mathbb{R} , además que $D(v - D\psi_{k,p}(u)v)w = -D(D\psi_{k,p}(u)v)w$ (ya que v es una variable aleatoria constante respecto de la variable u sobre la que se deriva) esta expresión se escribe como una derivada de orden superior de la forma $D(v - D\psi_{k,p}(u)v)w = -D^2\psi_{k,p}(u)(v, w)$, con estas situaciones la primer derivada del mapeo (4.29) es

$$u \longmapsto \left(\frac{d^2}{du^2} e_k^{u-\psi_{k,p}(u)} \right) (w - D\psi_{k,p}(u)w)(v - D\psi_{k,p}(u)v) - \left(\frac{d}{du} e_k^{u-\psi_{k,p}(u)} \right) D^2\psi_{k,p}(u)(v, w).$$

Como u es una variable aleatoria centrada, entonces su valor esperado es cero, de allí que

$$E_p \left[\left(\frac{d^2}{du^2} e_k^{u-\psi_{k,p}(u)} \right) (w - D\psi_{k,p}(u)w)(v - D\psi_{k,p}(u)v) - \left(\frac{d}{du} e_k^{u-\psi_{k,p}(u)} \right) D^2\psi_{k,p}(u)(v, w) \right] = 0.$$

Por las propiedades de la integral de Lebesgue, esta igualdad equivale a

$$E_p \left[\left(\frac{d^2}{du^2} e_k^{u-\psi_{k,p}(u)} \right) (w - D\psi_{k,p}(u)w)(v - D\psi_{k,p}(u)v) \right] = D^2\psi_{k,p}(u)(v, w) E_p \left[\frac{d}{du} e_k^{u-\psi_{k,p}(u)} \right],$$

despejando $D^2\psi_{k,p}(u)(v, w)$ se logra la expresión

$$D^2\psi_{k,p}(u)(v, w) = \frac{E_p \left[\left(\frac{d^2}{du^2} e_k^{u-\psi_{k,p}(u)} \right) (w - D\psi_{k,p}(u)w)(v - D\psi_{k,p}(u)v) \right]}{E_p \left[\frac{d}{du} e_k^{u-\psi_{k,p}(u)} \right]}. \quad (4.30)$$

Si $w = v \neq 0$ entonces la igualdad (4.30) se escribe

$$D^2\psi_{k,p}(v, v) = D^2\psi_{k,p}(u)v^2 = \frac{E_p \left[\left(\frac{d^2}{du^2} e_k^{u-\psi_{k,p}(u)} \right) (v - D\psi_{k,p}(u)v)^2 \right]}{E_p \left[\frac{d}{du} e_k^{u-\psi_{k,p}(u)} \right]}.$$

Como las expresiones $\frac{d^2}{du^2} e_k^{u-\psi_{k,p}(u)}$, $(v - D\psi_{k,p}(u)v)^2$ y $\frac{d}{du} e_k^{u-\psi_{k,p}(u)}$ son todas positivas entonces $D^2\psi_{k,p}(v, v) > 0$ y así la función $\psi_{k,p}(u)$ es estrictamente convexa para cualquier variable aleatoria centrada u . En el caso en que $u = 0$ se tiene que $\psi_{k,p}(0) = 0$ y $D\psi_{k,p}(0)v = 0$ para cualquier variable aleatoria centrada, en (4.30) se logra que

$$D^2\psi_{k,p}(0)(v, w) = E_p[vw] .$$

Debido que v y w son variables aleatorias en $L_0^{1/k}(p \cdot \mu)$ entonces $E_p[v] = 0 = E_p[w]$, por lo que la igualdad anterior implica

$$D^2\psi_{k,p}(0)(v, w) = E_p[vw] - E_p[v]E_p[w] = cov_p[v, w] \quad (4.31)$$

siempre que $u = 0$.

4.3. Espacio Tangente

Se estudia ahora el espacio tangente. Sea p_θ , con $\theta \in (-1, 1)$. Una curva en \mathcal{E}_p está definida como

$$p_\theta = e_k^{u_\theta - \psi_{k,p_0}(u_\theta)} p_0 , \quad (4.32)$$

donde p_0 es una densidad en \mathfrak{M}_μ y $u_0 = 0$ para $\theta = 0$. En la carta para p , el vector velocidad está dado por

$$\dot{u}_\theta \in L_0^{1/k}(p \cdot \mu) . \quad (4.33)$$

Se identifica el espacio tangente en p_0 con el espacio de variable aleatorias \dot{u}_0 . En general, el espacio tangente de $p \in \mathfrak{M}_\mu$ está identificado por $T_p = L_0^{1/k}(p \cdot \mu)$. Derivando la curva p_θ definido en 4.32, respecto de θ se tiene

$$\dot{p}_\theta = \frac{e_k^{u_\theta - \psi_{k,p_0}(u_\theta)} p_0}{\sqrt{1 + k^2(u_\theta - \psi_{k,p_0}(u_\theta))^2}} (\dot{u}_\theta - D\psi_{k,p_0}(u_\theta)\dot{u}_\theta) = \frac{p_\theta(\dot{u}_\theta - D\psi_{k,p_0}(u_\theta)\dot{u}_\theta)}{\sqrt{1 + k^2(u_\theta - \psi_{k,p_0}(u_\theta))^2}} ,$$

de donde resulta la expresión

$$\frac{\dot{p}_\theta}{p_\theta} = \frac{\dot{u}_\theta - D\psi_{k,p_0}(u_\theta)\dot{u}_\theta}{\sqrt{1 + k^2(u_\theta - \psi_{k,p_0}(u_\theta))^2}} . \quad (4.34)$$

Para $\theta = 0$, $u_0 = 0$ así como $\psi_{k,p}(0) = 0$ y $D\psi_{k,p}(0) = 0$, en la ecuación (4.34) se tiene $\frac{\dot{p}_0}{p_0} = \dot{u}_0$. En general, si $u = u_\theta$ y $v = \dot{u}_\theta$ entonces, por la expresión (4.24) se tiene

$$E_{p_\theta} \left[\frac{\dot{p}_\theta}{p_\theta} \right] = E_{p_\theta} \left[\frac{\dot{u}_\theta - D\psi_{k,p_0}(u_\theta)\dot{u}_\theta}{\sqrt{1 + k^2(u_\theta - \psi_{k,p_0}(u_\theta))^2}} \right] = 0 . \quad (4.35)$$

Por la definición de la curva p_θ se sigue que $u - \psi_{k,p_0}(u) = \ln_k \left(\frac{p_\theta}{p_0} \right)$; y por la ecuación (4.27), $\frac{\dot{p}_\theta}{p_\theta}$ se escribe como

$$\frac{\dot{p}_\theta}{p_\theta} = \frac{\dot{u}_\theta - E_{p_\theta|p_0}[\dot{u}_\theta]}{\sqrt{1 + k^2 \ln_k^2 \left(\frac{p_\theta}{p_0} \right)}}, \quad (4.36)$$

de donde

$$E_{p_\theta} \left[\frac{\dot{u}_\theta - E_{p_\theta|p_0}[\dot{u}_\theta]}{\sqrt{1 + k^2 \ln_k^2 \left(\frac{p_\theta}{p_0} \right)}} \right] = 0. \quad (4.37)$$

Se define el transporte paralelo $U_{p,\bar{p}}^k$ mapeando desde el espacio tangente en p , es decir, $T_p = L_0^{1/k}(p \cdot \mu)$, sobre el espacio tangente en \bar{p} , $T_{\bar{p}} = L_0^{1/k}(\bar{p} \cdot \mu)$ como sigue.

$$\begin{aligned} L_0^{1/k}(p \cdot \mu) &\rightarrow L_0^{1/k}(\bar{p} \cdot \mu) \\ u &\mapsto \bar{u} = \frac{u - E_{\bar{p},p}[u]}{\sqrt{1 + k^2 \ln_k^2 \left(\frac{\bar{p}}{p} \right)}}, \end{aligned} \quad (4.38)$$

se escribe $\bar{u} = U_{p,\bar{p}}^k$. El mapeo está bien definido y en forma análoga a la expresión (4.37) se tiene que

$$E_{\bar{p}}[\bar{u}] = E_{\bar{p}} \left[\frac{u - E_{\bar{p},p}[u]}{\sqrt{1 + k^2 \ln_k^2 \left(\frac{\bar{p}}{p} \right)}} \right] = 0.$$

Donde p y \bar{p} se conectan a través de la isometría definida en (4.6). Ahora bien, por la desigualdad

$$\frac{1}{\sqrt{1 + k^2 \ln_k^2 \left(\frac{\bar{p}}{p} \right)}} = \frac{1}{\sqrt{1 + k^2 \ln_k^2 \left(\frac{p}{\bar{p}} \right)}} \leq 2 \left(\frac{p}{\bar{p}} \right)^k,$$

se tiene

$$\begin{aligned} E_{\bar{p}} \left[|U_{p,\bar{p}}^k(u)|^{1/k} \right] &= E_{\bar{p}} \left[\left| \frac{u - E_{\bar{p},p}[u]}{\sqrt{1 + k^2 \ln_k^2 \left(\frac{\bar{p}}{p} \right)}} \right|^{1/k} \right] \leq E_{\bar{p}} \left[\left| 2 \left(\frac{p}{\bar{p}} \right)^k (u - E_{\bar{p},p}[u]) \right|^{1/k} \right] \\ &\leq 2^{1/k} E_{\bar{p}} \left[|u|^{1/k} \frac{p}{\bar{p}} \right] = 2^{1/k} E_p \left[|u|^{1/k} \right]. \end{aligned}$$

Como $u \in L_0^{1/k}(p \cdot \mu)$ entonces $E_p[|u|^{1/k}] < \infty$ y por tanto $E_{\bar{p}} \left[|U_{p,\bar{p}}^k(u)|^{1/k} \right] < \infty$. Luego, el mapeo (4.38) está bien definido.

Por la igualdad presentada en (4.34) y la definición del transporte paralelo en (4.38), se sigue que

$$\frac{\dot{p}_\theta}{p_\theta} = U_{p_0, p_\theta}^k(\dot{u}_\theta), \quad (4.39)$$

es decir, el vector velocidad en θ , transporta a el espacio tangente en p_θ . En este sentido, se puede definir los modelos k -exponenciales no-paramétricos como

$$q = e_k^{u - \psi_{k,p}(u)} p, \quad u \in V, \quad (4.40)$$

donde V es un subespacio lineal de $L_0^{1/k}(p \cdot \mu)$. Cada $v \in V$ es llamado una variable canónica del modelo k -exponencial. La representación implícita del modelo exponencial definido en (4.40) es

$$E_p \left[\ln_k \left(\frac{q}{p} \right) v \right] = 0, \quad v \in V^\perp, \quad (4.41)$$

donde V^\perp es un subconjunto del espacio dual de $L^{1/k}(p \cdot \mu)$ y representa el complemento ortogonal de V , como $L^{1/k}(p \cdot \mu)$ es un espacio de Lebesgue entonces su dual es $L_0^{1-1/k}(p \cdot \mu)$, así $V^\perp \subseteq L_0^{1-1/k}(p \cdot \mu)$.

Capítulo 5

Variedad de información k -deformada

En la discusión propuesta por G. Pistone [23], sobre posibles estructuras que permitan interpretaciones geométricas de aspectos relacionados con la información en el contexto del formalismo de la mecánica estadística según G. Kaniadakis, aparece una propuesta que plantea otro tipo de k -deformación para los modelos exponenciales. Precisamente propone la consideración de modelos tipo $q = e_k^{u \ominus \Psi(u)} p$, donde \ominus es la k -diferencia según G. Kaniadakis y $\Psi(u)$ es alguna función de una variable aleatoria u . En el presente capítulo se desarrolla dicha propuesta y, como se verá, se establece que $\Psi(u)$ es una k -deformación del mapeo acumulante definido para una variable aleatoria u por K_p y se denotará $\Psi(u) := K_{p,k}(u)$, donde $K_{p,k}$ se definirá posteriormente. Es por esto que en la sección 5.1 se presentan definiciones y propiedades sobre el espacio de Orlicz $L^{\phi_k}(p \cdot \mu)$, las secciones 5.2 y 5.3 están dedicadas a establecer la analiticidad del mapeo $K_{p,k}$ y a definir una k -deformación de los modelos exponenciales y en la sección 5.4 se definen los mapeos que determinan una nueva variedad topológica sobre \mathfrak{M}_μ . En el transcurso del capítulo se estudia la diferenciabilidad y propiedades de algunas de las funciones definidas, esto se hace con miras a la continuidad del trabajo (problemas abiertos), no para la construcción de una variedad diferenciable.

5.1. Generalidades sobre el espacio $L^{\phi_k}(p \cdot \mu)$

En el capítulo 2 se demostró que la función exponencial k deformada

$$\exp_{\{k\}}(x) = \left[\sqrt{1 + k^2 x^2} + kx \right]^{\frac{1}{k}} \quad \text{para } |k| < 1 \text{ y } k \neq 0 \quad \text{y} \quad \exp_{\{0\}}(x) = \exp(x),$$

es par respecto de la deformación, es decir, $\exp_{\{k\}}(x) = \exp_{\{-k\}}(x)$. Es por ello que el análisis en este capítulo se hará para $0 < k < 1$ y por ende las propiedades se cumplen para todo $-1 < k < 1$ con $k \neq 0$. Por efectos de notación se escribirá $\exp_{\{k\}}(x) = e_k^x$.

Sea $0 < k < 1$, se define la función $\phi_k(x) = \cosh_k(x) - 1$, la cual es una función de Young para cada k y representa una k -deformación de la función de Young $\phi_1(x) = \cosh x - 1$

definida en 2.12, recuperándola cuando k tiende a cero. En los preliminares se enunciaron algunas propiedades de la aplicación $\phi_1(\cdot)$.

Para una densidad $p \in \mathfrak{M}_\mu$, la función de Young ϕ_k determina el espacio de funciones de Orlicz dado por

$$\begin{aligned} L^{\phi_k}(p \cdot \mu) &= \{u \in L^1(p \cdot \mu) : E_p[\phi_k(\alpha u)] < \infty \text{ para alguna } \alpha > 0\} \\ &= \left\{ u \in L^1(p \cdot \mu) : \int_{\Omega} (\cosh_k(\alpha u) - 1) p d\mu < \infty \text{ para alguna } \alpha > 0 \right\}, \end{aligned} \quad (5.1)$$

donde la norma de Luxemburgo se presenta como

$$\begin{aligned} \|u\|_{\phi_k, p} &= \inf \left\{ r > 0 : E_p \left[\phi_k \left(\frac{u}{r} \right) \right] \leq 1 \right\} \\ &= \inf \left\{ r > 0 : \int_{\Omega} \left(\cosh_k \left(\frac{u}{r} \right) - 1 \right) p d\mu \leq 1 \right\}, \end{aligned} \quad (5.2)$$

caracterizando el espacio de Orlicz como

$$L^{\phi_k}(p \cdot \mu) = \{u \in L^1(p \cdot \mu) : \|u\|_{\phi_k, p} < \infty\}.$$

Con la norma de Luxemburgo, el espacio $L^{\phi_k}(p \cdot \mu)$ es un espacio de Banach convexo. Por $B_{\phi_k}(0, 1)$ o simplemente B_{ϕ_k} se denotará la bola unitaria en el espacio de Orlicz $L^{\phi_k}(p \cdot \mu)$, donde $u \in B_{\phi_k}$ si y sólo si $\|u\|_{\phi_k, p} < 1$. La esfera unitaria S_{ϕ_k} es el conjunto de variables aleatorias u para las cuales $\|u\|_{\phi_k, p} = 1$.

Debido a que $\phi_k(x) \leq \phi_1(x)$ para todo $x \in \mathbb{R}$ y $0 < k < 1$ entonces, se tiene que

$$L^{\phi_1}(p \cdot \mu) \subset L^{\phi_k}(p \cdot \mu) \subset L^1(p \cdot \mu).$$

En efecto, si $u \in L^{\phi_1}(p \cdot \mu)$, existe $\alpha > 0$ tal que $E_p[\phi_1(\alpha u)] < \infty$. Ya que $\phi_k(x) \leq \phi_1(x)$ entonces, $\phi_k(\alpha u) \leq \phi_1(\alpha u)$, de donde $E_p[\phi_k(\alpha u)] \leq E_p[\phi_1(\alpha u)]$, así $E_p[\phi_k(\alpha u)] < \infty$ para alguna $\alpha > 0$ y por tanto $L^{\phi_1}(p \cdot \mu) \subset L^{\phi_k}(p \cdot \mu)$; la segunda inclusión es inmediata.

Por la anterior inclusión conjuntista resultará que la variedad que se construya para \mathfrak{M}_μ a través de $L^{\phi_k}(p \cdot \mu)$, es una extensión para la variedad construida a través $L^{\phi_1}(p \cdot \mu)$ por Pistone, como se presentó en el capítulo 3.

Ahora, si $u \in L^{\phi_1}(p \cdot \mu)$ entonces, se satisface la desigualdad $\|u\|_{\phi_k, p} \leq \|u\|_{\phi_1, p}$, por lo que si $u \in B_{\phi_1}(0, 1)$ se sigue que $\|u\|_{\phi_k, p} < 1$. En consecuencia, se presenta la inclusión

$$B_{\phi_1}(0, 1) \subset B_{\phi_k}(0, 1), \quad (5.3)$$

es decir, la bola unitaria del espacio de Banach $L^{\phi_1}(p \cdot \mu)$ está contenida en la bola unitaria del espacio de Banach $L^{\phi_k}(p \cdot \mu)$.

5.2. Analiticidad

Sean k tal que $0 < k < 1$ y $p \in \mathfrak{M}_\mu$. Consideremos $1 \leq a < \frac{1}{k}$, por lo que el espacio de Lebesgue $L^a(p \cdot \mu)$ está bien definido.

Sea $p \in \mathfrak{M}_\mu$. El espacio de funciones de Orlicz $L^{\phi_1}(p \cdot \mu)$ es localmente convexo como consecuencia del teorema 2.4, por lo que existe una vecindad convexa \mathcal{O} del cero tal que $\mathcal{O} \subset B_{\phi_1}(0, 1)$ y por tanto se presentan, de acuerdo con (5.3), las inclusiones

$$\mathcal{O} \subset B_{\phi_1} \subset B_{\phi_k}. \quad (5.4)$$

Sean $u \in \mathcal{O}$ y $v_i \in \mathcal{O}$ para $i = 1, 2, \dots, n$. Por (5.4) u y cada v_i se encuentran en las bolas unitarias B_{ϕ_1} y B_{ϕ_k} . Con base en esto, las normas de Luxemburgo son tales que $\|u\|_{\phi_1, p} < 1$ y $\|u\|_{\phi_k, p} < 1$. Si se hace que $R = \frac{1 - \|u\|_{\phi_1, p}}{n}$ y $R_1 = \frac{1 - \|u\|_{\phi_k, p}}{n}$ entonces $R > 0$ y $R_1 > 0$. Procediendo en forma análoga que en la demostración del teorema 3.2 se sigue que

$$\left(u + R \sum_{i=1}^n |v_i| \right) \in \overline{B}_{\phi_1} \quad \text{y} \quad \left(u + R_1 \sum_{i=1}^n |v_i| \right) \in \overline{B}_{\phi_k}. \quad (5.5)$$

En el caso en que $r = \min\{R, R_1\}$, esto es, $r \leq R$ y $r \leq R_1$, entonces las dos inclusiones de (5.5) se verifican para r , es decir, $\left(u + r \sum_{i=1}^n |v_i| \right) \in \overline{B}_{\phi_1}$ y $\left(u + r \sum_{i=1}^n |v_i| \right) \in \overline{B}_{\phi_k}$, esto siempre y cuando $u \in \mathcal{O}$ y $v_i \in \mathcal{O}$. Por la condición de ser $u \in \mathcal{O} \subset B_{\phi_1}$ implica que $\|u\|_{\phi_1, p} < 1$ y por ende se satisface el lema 3.1. El siguiente lema relaciona las funciones polinómicas del tipo x^a para $a \geq 1$ con la exponencial k -deformada.

Lema 5.1. Sean $0 < k < 1$ y $1 \leq a < \frac{1}{k}$ entonces, para $u \in L^{\phi_k}(p \cdot \mu)$ se sigue que

$$\left| \frac{kru}{a} \right|^a \leq e_k^{|ru|},$$

siendo r un número real arbitrario.

Demostración

Para todo número real x se verifica que $|x| \leq \sqrt{1 + x^2}$ de allí que

$$\begin{aligned} |kru| &\leq \sqrt{1 + (kru)^2} \leq \sqrt{1 + (kru)^2} + |kru| \\ &\leq \sqrt{1 + k^2|ru|^2} + k|ru|. \end{aligned}$$

Por la elección de a se tiene que $1 \leq a$, equivalente a $0 < \frac{1}{a} \leq 1$, y por lo tanto al multiplicar por $|kru|$ resulta $\frac{|kru|}{a} \leq |kru|$; resulta así

$$\left| \frac{kru}{a} \right| = \frac{|kru|}{a} \leq \sqrt{1 + k^2|ru|^2} + k|ru|.$$

Debido a que $a < \frac{1}{k}$ entonces en la expresión anterior se tiene

$$\left| \frac{kru}{a} \right|^a \leq \left(\sqrt{1 + k^2|ru|^2} + k|ru| \right)^{\frac{1}{k}} = e_k^{|ru|},$$

lo cual era el objetivo de la demostración, donde r es un valor arbitrario. \square

En la siguiente definición se escribe $(e_k^u)^{\frac{1}{a}}$ para diferenciarlo del término $e_k^{\frac{u}{a}}$ que por las propiedades de la función exponencial k deformada son diferentes ($(e_k^u)^a = e_k^{au}$).

Definición 5.1. Sean $p \in \mathfrak{M}_\mu$, $0 < k < 1$, $1 \leq a < \frac{1}{k}$ y $u \in \mathcal{O} \subset B_{\phi_k}(0, 1)$. Se define el operador $\lambda_{a,n,k}(u)$ como

$$\begin{aligned} \lambda_{a,n,k}(u)(\cdot) : \quad L^{\phi_k}(p \cdot \mu) \times \dots \times L^{\phi_k}(p \cdot \mu) &\rightarrow L^a(p \cdot \mu) \\ (w_1, \dots, w_n) &\mapsto \frac{w_1 k}{a} \dots \frac{w_n k}{a} (e_k^u)^{\frac{1}{a}}. \end{aligned}$$

Para este operador que depende de cuatro parámetros (a , n , k y u), el siguiente teorema establece que dicho operador es multilineal, simétrico y continuo, es decir, $\lambda_{a,n,k}(u) \in \mathcal{L}_s^n(L^{\phi_k}(p \cdot \mu), L^a(p \cdot \mu))$. La demostración se asemeja en cuanto a estructura a la demostración del teorema 3.2.

Teorema 5.1. El operador $\lambda_{a,n,k}(u)$ es multilineal, simétrico y continuo.

Demostración

La multilinealidad se demuestra sobre la i -ésima componente y dejando fijas las $n - 1$ componentes restantes. El operador es simétrico ya que al cambiar cualquier componente de la n -tupla (w_1, w_2, \dots, w_n) , el valor del operador $\lambda_{a,n,k}(u)$ no cambia. Veamos ahora que es continuo. Debido a que $u \in \mathcal{O}$ entonces se satisface las expresiones 5.5, por lo que si se hace $r = \min\{R, R_1\}$ entonces

$$\left(u + r \sum_{i=1}^n |v_i| \right) \in \overline{B}_{\phi_1}(0, 1), \quad (1),$$

donde los v_i para $i = 1, 2, \dots, n$ son elementos en la vecindad convexa \mathcal{O} . De acuerdo con las propiedades de la integral de Lebesgue resulta

$$\int_{\Omega} \left| \frac{krv_1}{a} \dots \frac{krv_n}{a} (e_k^u)^{\frac{1}{a}} \right|^a p d\mu = \int_{\Omega} \left| \frac{krv_1}{a} \right|^a \dots \left| \frac{krv_n}{a} \right|^a e_k^u p d\mu. \quad (2)$$

En la igualdad (2) se hará uso del lema 5.1 donde se sabe que para cualquier $0 < k < 1$ se presenta la desigualdad $e_k^x \leq e^x$ para $x \geq 0$. Haciendo uso de estos dos hechos se tiene

$$\begin{aligned} \int_{\Omega} \left| \frac{krv_1}{a} \dots \frac{krv_n}{a} (e_k^u)^{\frac{1}{a}} \right|^a p d\mu &\leq \int_{\Omega} e_k^{|rv_1|} e_k^{|rv_2|} \dots e_k^{|rv_n|} e_k^u p d\mu \leq \int_{\Omega} e^{r|v_1|} e^{r|v_2|} \dots e^{r|v_n|} e^u p d\mu \\ &\leq \int_{\Omega} e^{u+r \sum_{i=1}^n |v_i|} p d\mu \leq 2 \int_{\Omega} \cosh \left(u + r \sum_{i=1}^n |v_i| \right) p d\mu. \end{aligned}$$

De (1) y el lema 3.1, la desigualdad anterior se expresa como

$$\int_{\Omega} |\lambda_{a,n,k}(u)(v_1, \dots, v_2)|^a p d\mu = \frac{1}{r^{na}} \int_{\Omega} \left| \frac{krv_1}{a} \dots \frac{krv_n}{a} (e_k^u)^{\frac{1}{a}} \right|^a p d\mu \leq \frac{4}{r^{na}}.$$

Con esto se concluye que $\lambda_{a,n,k}(u)(v_1, \dots, v_n)$ es un elemento en $L^a(p \cdot \mu)$, cuya norma está dada por

$$\|\lambda_{a,n,k}(u)(v_1, \dots, v_n)\|_{L^a} = \left(\int_{\Omega} |\lambda_{a,n,k}(u)(v_1, \dots, v_n)|^a p d\mu \right)^{\frac{1}{a}} \leq \left(\frac{4}{r^{na}} \right)^{\frac{1}{a}} = \frac{4^{\frac{1}{a}}}{r^n}.$$

Se sigue que para w_1, w_2, \dots, w_n en $L^{\phi_k}(p \cdot \mu)$

$$\int_{\Omega} |\lambda_{a,n,k}(u)(w_1, \dots, w_n)|^a p d\mu = \int_{\Omega} \left| \frac{w_1 k}{a} \dots \frac{w_n k}{a} (e_k^u)^{\frac{1}{a}} \right|^a p d\mu \leq \|w_1\|_{\phi_k}^a \dots \|w_n\|_{\phi_k}^a \frac{4}{r^{na}}.$$

Así $\lambda_{a,n,k}(u)(w_1, \dots, w_n) \in L^a(p \cdot \mu)$ y además

$$\|\lambda_{a,n,k}(u)\| = \sup_{w_i \neq 0} \frac{\|\lambda_{a,n,k}(w_1, \dots, w_n)\|_{L^a}}{\|w_1\|_{\phi_k} \dots \|w_n\|_{\phi_k}} \leq \frac{4^{\frac{1}{a}}}{r^n}, \quad (5.6)$$

por lo que el operador $\lambda_{a,n,k}(u)$ es acotado y por tanto continuo, así como multilineal y simétrico, lo cual permite concluir que $\lambda_{a,n,k}(u) \in \mathcal{L}_s^n(L^{\phi_k}(p \cdot \mu), L^a(p \cdot \mu))$. Nótese que dicho acotamiento no depende de la variable aleatoria u y de la k -deformación. \square

Como $\lambda_{a,n,k}(u)$ es un operador multilineal y simétrico, entonces, por la desigualdad (5.6) se escribe como

$$\|\lambda_{a,n,k}(u)\|_{\mathcal{L}_s^n} \leq \frac{4^{\frac{1}{a}}}{r^n}. \quad (5.7)$$

Sea $u \in \mathcal{O} \subset B_{\phi_1}(0, 1)$ entonces, por el lema 3.1 se sigue que

$$\int_{\Omega} \left| (e_k^u)^{\frac{1}{a}} \right|^a p d\mu = \int_{\Omega} |e_k^u|^a p d\mu \leq \int_{\Omega} e_k^{|u|} p d\mu \leq \int_{\Omega} e^{|u|} p d\mu \leq 2 \int_{\Omega} \cosh(|u|) p d\mu \leq 4,$$

donde $\cosh(|u|) = \cosh u$ por la simetría de esta función. La desigualdad anterior implica que $\left\| (e_k^u)^{\frac{1}{a}} \right\|_{L^a} \leq 4^{\frac{1}{a}}$ y así $(e_k^u)^{\frac{1}{a}} \in L^a(p \cdot \mu)$. Este resultado motiva la siguiente definición.

Definición 5.2. Sean $p \in \mathfrak{M}_{\mu}$, $0 < k < 1$ y $1 \leq a < \frac{1}{k}$ fijos, para $u \in \mathcal{O} \subset B_{\phi_k}(0, 1)$ se define el operador $\lambda_{a,0,k}(u)$ como $\lambda_{a,0,k}(u) = (e_k^u)^{\frac{1}{a}}$.

Sea $p \in \mathfrak{M}_{\mu}$, si hacemos que $w_i = v \in L^{\phi_k}(p \cdot \mu)$ para cada $i = 1, 2, \dots, n$ en la definición 5.1 se tiene que

$$\begin{aligned} \lambda_{a,n,k}(u)(v, v, \dots, v) &= \frac{kv}{a} \dots \frac{kv}{a} (e_k^u)^{\frac{1}{a}} \\ &= \left(\frac{kv}{a} \right)^n (e_k^u)^{\frac{1}{a}} = \left(\frac{kv}{a} \right)^n \lambda_{a,0,k}(u). \end{aligned}$$

Debido a que $(v, v, \dots, v) = v^n$ entonces la igualdad anterior se puede escribir como

$$\lambda_{a,n,k}(u)v^n = \left(\frac{kv}{a} \right)^n \lambda_{a,0,k}(u).$$

Se define ahora la función polinomial continua n -homogénea como

$$\hat{\lambda}_{a,n,k}(u)v = \lambda_{a,n,k}(u)v^n = \left(\frac{kv}{a} \right)^n (e_k^u)^{\frac{1}{a}} = \left(\frac{kv}{a} \right)^n \lambda_{a,0,k}(u). \quad (5.8)$$

Por lo que $\hat{\lambda}_{a,n,k}(u) : L^{\phi_k}(p \cdot \mu) \rightarrow L^a(p \cdot \mu)$. En el caso en que $u = 0$ se tiene que $\hat{\lambda}_{a,n,k}(0)v = \left(\frac{kv}{a} \right)^n$.

En los preliminares se expuso que la serie de Taylor para la función exponencial k -deformada está dada por

$$e_k^x = \sum_{n=0}^{\infty} a_n(k) \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + (1 - k^2) \frac{x^3}{3!} + \dots$$

Donde $k^2 x^2 < 1$, $a_0(k) = 1$, $a_1(k) = 1$, y para cualquier n , los coeficientes a_n están dados por

$$a_{2m}(k) = \prod_{j=0}^{m-1} [1 - (2j)^2 k^2] \quad \text{y} \quad a_{2m+1}(k) = \prod_{j=1}^m [1 - (2j-1)^2 k^2]. \quad (5.9)$$

Cada uno de los coeficientes cumple $a_n(k) = a_n(-k)$ es decir, son pares respecto a la deformación k para $-1 < k < 1$, esto se debe a que en cada caso se presenta el cuadrado de la deformación k , de allí que el siguiente lema se demuestre para $0 < k < 1$. En el caso en que $k = 0$ resulta que $a_{2m}(0) = 1$ y $a_{2m+1}(0) = 1$, por lo que la serie de potencia para $e_{\{0\}}^x$ se escribe como $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ que es la serie para la exponencial usual.

Lema 5.2. *Para todo $n \in \mathbb{N}$ se satisface que $a_n(k) \leq 1$, donde $0 < k < 1$.*

Demostración

Demostremos la propiedad para los coeficientes con índice par, es decir, $a_{2n}(k) \leq 1$ para todo $n \in \mathbb{N}$. Para $n = 1$ se obtiene $a_2(k) = 1$ y por tanto la propiedad se cumple. Supongamos ahora que la propiedad se cumple para m , es decir,

$$\prod_{j=0}^{m-1} [1 - (2j)^2 k^2] = a_{2m}(k) \leq 1 \quad \text{Hipótesis Inductiva .}$$

Demostremos que la propiedad se cumple para $m + 1$, es decir, $a_{2(m+1)}(k) \leq 1$. Por la definición de los coeficientes con índice par en (5.9) resulta

$$\begin{aligned} a_{2(m+1)}(k) &= \prod_{j=0}^{m+1-1} [1 - (2j)^2 k^2] = \prod_{j=0}^m [1 - (2j)^2 k^2] \\ &= \left(\prod_{j=0}^{m-1} [1 - (2j)^2 k^2] \right) [1 - (2m)^2 k^2] = a_{2m}(k) [1 - 4m^2 k^2] . \end{aligned}$$

Haciendo uso de la hipótesis inductiva resulta la desigualdad $a_{2(m+1)}(k) \leq (1 - 4m^2 k^2)$ (*). Debido a que $0 < k < 1$ entonces $0 < k^2 < 1$ de donde $-1 < -k^2 < 0$ que al multiplicar por $4m^2 > 0$ resulta $-4m^2 < -4m^2 k^2 < 0$ sumando 1 se logra $1 - 4m^2 < 1 - 4m^2 k^2 < 1$ y así por transitividad con la expresión (*) se tiene $a_{2(m+1)}(k) \leq 1$ y con esto concluir que $a_{2n}(k) \leq 1$ para todo n par. Procediendo de igual forma se demuestra que para n impar la propiedad se cumple y así $a_n(k) \leq 1$ para todo $n \in \mathbb{N}$. \square

Teorema 5.2. *Sean $p \in \mathfrak{M}_\mu$, $0 < k < 1$ y $1 \leq a < \frac{1}{k}$ fijos. El mapeo A está definido como*

$$\begin{aligned} A(\cdot) : L^{\phi_k}(p \cdot \mu) &\rightarrow L^a(p \cdot \mu) \\ v &\mapsto A(v) = \sum_{n=0}^{\infty} \frac{1}{n!} a_n(k) \left(\frac{v}{a}\right)^n \end{aligned}$$

Entonces $A(v)$ es una serie de potencia con radio de convergencia $\hat{\rho} > k > 0$.

Demostración

De acuerdo con la definición de $A(v)$ y de la función polinomial (5.8) resulta que

$$\begin{aligned} A(v) &= \sum_{n=0}^{\infty} \frac{1}{n!} a_n(k) \left(\frac{v}{a}\right)^n = \sum_{n=0}^{\infty} \frac{a_n(k)}{n!k^n} \left(\frac{kv}{a}\right)^n \\ &= \sum_{n=0}^{\infty} \frac{a_n(k)}{n!k^n} \lambda_{n,a,k}(0) v^n = \sum_{n=0}^{\infty} \frac{a_n(k)}{n!k^n} \hat{\lambda}_{n,a,k}(0) v. \end{aligned}$$

Con base en la definición 2.9 se sigue que $A(v) = \sum_{n=0}^{\infty} \frac{a_n(k)}{n!k^n} \hat{\lambda}_{n,a,k}(0) v$ es una serie de potencia de $L^{\phi_k}(p \cdot \mu)$ en $L^a(p \cdot \mu)$. Se determina ahora cual es su radio de convergencia, para ello sea $u \in \mathcal{O} \subset \overline{B}_{\phi_1}(0, 1)$, es decir, $\|u\|_{\phi_1} \leq 1$. Por el lema 5.2 tenemos que $a_n(k) \leq 1$ para todo $n \in \mathbb{N}$, como $n! \geq 1$ entonces $\frac{1}{n!} \leq 1$ por lo que el cociente $\frac{a_n(k)}{n!}$ no excede el 1. Por el lema 5.1 resulta

$$\left| \frac{a_n(k)}{n!k^n} \left(\frac{kv}{a}\right)^n \right|^a \leq \frac{1}{k^{an}} \left(\left| \frac{kv}{a} \right|^n \right)^a \leq \frac{1}{k^{an}} \left(e^{|v|} \right).$$

Como $0 < k < 1$ entonces $|kv| = k|v| \leq |v|$. Debido a que la función exponencial es creciente se sigue que $e^{|kv|} \leq e^{|v|}$, por transitividad con la expresión (1) resulta que

$$\left| \frac{a_n(k)}{n!k^n} \left(\frac{kv}{a}\right)^n \right|^a \leq \frac{1}{k^{an}} e^{|v|} \quad (2)$$

Haciendo uso de la definición de la norma en los espacios de Lebesgue y del lema 3.1 se tiene

$$\begin{aligned} \left\| \frac{a_n(k)}{n!k^n} \hat{\lambda}_{a,n,k}(0) v \right\|_{L^a} &= \left[\int_{\Omega} \left| \frac{a_n(k)}{n!k^n} \left(\frac{kv}{a}\right)^n \right|^a p d\mu \right]^{\frac{1}{a}} \\ &\leq \left[\int_{\Omega} \frac{1}{k^{an}} e^{|v|} p d\mu \right]^{\frac{1}{a}} \leq \frac{1}{k^n} \left[\int_{\Omega} e^{|v|} p d\mu \right]^{\frac{1}{a}} \\ &\leq \frac{1}{k^n} \left[2 \int_{\Omega} \cosh |v| p d\mu \right]^{\frac{1}{a}} \leq \frac{1}{k^n} 4^{\frac{1}{a}}, \end{aligned}$$

así $\left\| \frac{a_n(k)}{n!k^n} \hat{\lambda}_{a,n,k}(0) \right\|_{\mathcal{P}_n} \leq \frac{1}{k^n} 4^{\frac{1}{a}}$ que por el criterio de Cauchy-Hadamard 2.14 resulta

$$\frac{1}{\hat{\rho}} = \limsup_{n \rightarrow \infty} \left\| \frac{a_n(k)}{n!k^n} \hat{\lambda}_{a,n,k}(0) \right\|_{\mathcal{P}_n}^{\frac{1}{n}} \leq \limsup_{n \rightarrow \infty} \left(\frac{1}{k^n} 4^{\frac{1}{a}} \right)^{\frac{1}{n}} \leq \frac{1}{k} \limsup_{n \rightarrow \infty} \left[4^{\frac{1}{a}} \right]^{\frac{1}{n}}.$$

Por tanto $\frac{1}{\hat{\rho}} \leq \frac{1}{k}$ para obtener que $\hat{\rho} \geq k > 0$. \square

Teorema 5.3. Sean $p \in \mathfrak{M}_\mu$, $0 < k < 1$ y $1 \leq a < \frac{1}{k}$. Si $u \in \mathcal{O} \subset B_{\phi_k}(0, 1)$ entonces la serie

$$v \rightarrow \sum_{n=0}^{\infty} \frac{1}{n!} a_n(k) \left(\frac{v}{a}\right)^n$$

es analítica.

Demostración

Haciendo uso de la desigualdad presentada en (5.7) y por el hecho que $a_n(k) > 0$ para todo $n \in \mathbb{N}$ y $-1 < k < 1$ entonces

$$\left\| \frac{a_n(k)}{n!k^n} \lambda_{a,n,k}(0) \right\|_{\mathcal{L}_s^n} = \left| \frac{a_n(k)}{n!k^n} \right| \|\lambda_{a,n,k}(0)\|_{\mathcal{L}_s^n} \leq \frac{a_n(k)}{n!k^n} \frac{4^{\frac{1}{a}}}{r^n}, \quad (1)$$

donde $r = R = \frac{1-\|u\|_{\phi_1}}{n}$ o $r = R_1 = \frac{1-\|u\|_{\phi_k}}{n}$, siempre que $u \in \mathcal{O}$. Trabajemos con $r = R_1$ (de igual forma se hace el análisis para $r = R$) debido a que $u = 0$ entonces $\|u\|_{\phi_k} = 0$ por lo que $r = \frac{1}{n}$ y en la desigualdad presentada en (1) se sigue que

$$\left\| \frac{a_n(k)}{n!k^n} \lambda_{a,n,k}(0) \right\|_{\mathcal{L}_s^n} \leq \frac{a_n(k)4^{\frac{1}{a}}n^n}{n!k^n} \leq \frac{4^{\frac{1}{a}}n^n}{n!k^n},$$

en este caso también se aplicó el lema 5.2. Haciendo uso del criterio de Cauchy-Hadamard para el radio de convergencia restringido ver (2.14) se sigue

$$\frac{1}{\rho} = \limsup_{n \rightarrow \infty} \left\| \frac{a_n(k)}{n!k^n} \lambda_{a,n,k}(0) \right\|_{\mathcal{L}_s^n}^{\frac{1}{n}} \leq \limsup_{n \rightarrow \infty} \left(\frac{4^{\frac{1}{a}}n^n}{n!k^n} \right)^{\frac{1}{n}} = \frac{1}{k} \limsup_{n \rightarrow \infty} \left(\frac{4^{\frac{1}{a}}n^n}{n!} \right)^{\frac{1}{n}},$$

de donde $\frac{1}{\rho} \leq \frac{e}{k}$ y así $\rho \geq \frac{k}{e}$, por medio de la desigualdad presentada en (2.15) sobre la equivalencia de los radios de convergencia se tiene que $\hat{\rho} \geq \rho \geq \frac{e}{k} > 0$; de acuerdo con esto $\sum_{n=0}^{\infty} \frac{1}{n!} a_n(k) \left(\frac{v}{a}\right)^n = \sum_{n=0}^{\infty} \frac{a_n(k)}{n!k^n} \hat{\lambda}_{a,n,k}(0)(v)$ es una serie de potencia con radio de convergencia positivo por lo que es analítica en una vecindad del cero.

Supongamos ahora que $u_0 \in \mathcal{O}$ por lo que $u_0 \in B_{\phi_k}(0, 1)$ y así $r = \frac{1-\|u_0\|_{\phi_k}}{n}$, haciendo uso de nuevo la desigualdad (5.7) centrada en u_0 resulta que

$$\left\| \frac{a_n(k)}{n!k^n} \lambda_{a,n,k}(u_0) \right\|_{\mathcal{L}_s^n} \leq \frac{4^{\frac{1}{a}}n^n}{n!k^n(1-\|u_0\|_{\phi_k})^n}.$$

Y por tanto el radio de convergencia restringido está dado por

$$\frac{1}{\rho} = \limsup_{n \rightarrow \infty} \left\| \frac{a_n(k)}{n!k^n} \lambda_{a,n,k}(u_0) \right\|_{\mathcal{L}_s^n}^{\frac{1}{n}} \leq \frac{1}{k(1-\|u_0\|_{\phi_k})} \limsup_{n \rightarrow \infty} \left(\frac{4^{\frac{1}{a}}n^n}{n!} \right)^{\frac{1}{n}}.$$

Se sigue que $\frac{1}{\rho} \leq \frac{e}{k(1-\|u_0\|_{\phi_k})}$ y así $\hat{\rho} \geq \rho \geq \frac{k(1-\|u_0\|_{\phi_k})}{e} > 0$, por lo tanto la serie de potencia $\sum_{n=0}^{\infty} \frac{a_n(k)}{n!k^n} \hat{\lambda}_{a,n,k}(u_0)(v)$ tiene radio de convergencia positivo y por tanto es analítica en una vecindad de u_0 . \square

Con base en el teorema 5.3 se sigue que la serie de potencia e_k^x es analítica, la cual de forma puntual se tiene que

$$e_k^{\frac{v}{a}} = \sum_{n=0}^{\infty} \frac{a_n(k)}{n!} \left(\frac{v}{a}\right)^n,$$

convergencia que se cumple también en medida $p \cdot \mu$, en términos del operador $\lambda_{a,n,k}$ se puede escribir como

$$e_k^{\frac{v}{a}} = \sum_{n=0}^{\infty} \frac{a_n(k)}{n!k^n} \left(\frac{kv}{a}\right)^n = \sum_{n=0}^{\infty} \frac{a_n(k)}{n!k^n} \lambda_{a,n,k}(0)v^n. \quad (5.10)$$

Veamos ahora que la función e_k^u es integrable para $u \in \mathcal{O}$.

Teorema 5.4. Sean $p \in \mathfrak{M}_\mu$, $0 < k < 1$ y $u \in \mathcal{O}$. Si $q = \frac{e_k^u}{E_p[e_k^u]}p$ entonces la variable aleatoria e_k^u es $p \cdot \mu$ integrable y $q \in \mathfrak{M}_\mu$.

Demostración

Ya que $u \in \mathcal{O}$ entonces por las inclusiones presentadas en (5.4) resulta $u \in B_{\phi_1}(0, 1)$, por lo que $\|u\|_{\phi_1, p} < 1$, siguiendo la demostración del teorema 3.11 tenemos que $E_p[e^u] \leq 4$, ya que $e_k^x \leq e^x$ para todo $x \in \mathbb{R}^+ \cup \{0\}$ y $0 < k < 1$ entonces $E_p[e_k^u] \leq E_p[e^u]$. Como $E_p[e^u] \leq 4$ se sigue que $E_p[e_k^u] \leq 4$ para $u \in \mathcal{O}$ y así e_k^u es $p \cdot \mu$ integrable.

Por hipótesis $q = \frac{e_k^u}{E_p[e_k^u]}p$, de acuerdo con las propiedades de la integral se tiene

$$\int_{\Omega} q d\mu = \int_{\Omega} \frac{e_k^u}{E_p[e_k^u]} p d\mu = \frac{1}{E_p[e_k^u]} \int_{\Omega} e_k^u p d\mu = \frac{1}{E_p[e_k^u]} E_p[e_k^u] = 1,$$

esto permite concluir que $q \in \mathfrak{M}_\mu$. \square

Respecto de las propiedades de la exponencial k deformada, $\frac{e_k^u}{E_p[e_k^u]}$ se puede escribir como

$$q = \frac{e_k^u}{E_p[e_k^u]} p = \frac{e_k^u}{\ln_k(E_p[e_k^u])} p = e_k^{u \ominus \ln_k(E_p[e_k^u])} p,$$

así $q = e_k^{u \ominus \ln_k (E_p[e_k^u])}$ p es una densidad de probabilidad siempre que p sea una densidad de probabilidad. La integrabilidad respecto a la medida $p \cdot \mu$ de la variable aleatoria e_k^u para $u \in \mathcal{O}$, $0 < k < 1$ permite definir la siguiente función.

Definición 5.3. *El funcional generador de momentos k -deformado es el mapeo*

$$\begin{aligned} M_{p,k}(\cdot) : L^{\phi_k}(p \cdot \mu) &\rightarrow [0, \infty] \\ u &\mapsto M_{p,k}(u) = E_p[e_k^u] \end{aligned}$$

con $0 < k < 1$.

Si en la igualdad 5.10 se hace $a = 1$ entonces

$$e_k^u = \sum_{n=0}^{\infty} \frac{a_n(k)}{n!k^n} (ku)^n = \sum_{n=0}^{\infty} \frac{a_n(k)}{n!k^n} \lambda_{1,n,k}(0) u^n .$$

Al integrar, se obtiene una expresión del funcional generador de momentos como

$$M_{p,k}(u) = E_p[e_k^u] = E_p \left[\sum_{n=0}^{\infty} \frac{a_n(k)}{n!k^n} \lambda_{1,n,k}(0) u^n \right] . \quad (5.11)$$

En el siguiente teorema se enuncian las propiedades del funcional generador de momentos k -deformado, propiedades que fueron enunciadas para M_p en el teorema 3.6. La expresión \mathcal{O}^n representa el producto cartesiano $\mathcal{O} \times \dots \times \mathcal{O}$. Además, con base en la definición se sigue que

$$\mathcal{O} \subset B_{\phi_1}(0, 1) \subset B_{\phi_k}(0, 1) \subset \text{dom}M_{p,k} .$$

En efecto, si $u \in \mathcal{O}$ entonces $u \in B_{\phi_1}(0, 1)$ por lo que $\|u\|_{\phi_1,p} < 1$ y haciendo uso del lema (3.1) se tiene que $E_p[\cosh u] \leq 2$ de donde se desprende que $E_p[e^u] \leq 4$; debido a que $e_k^x \leq e^x$ para todo $x \geq 0$ y $0 < k < 1$ entonces, $E_p[e_k^u] \leq E_p[e^u] \leq 4$ y así $u \in \text{dom}M_{p,k}$ ya que $M_{p,k}(u) \leq 4 < \infty$.

Se verifica además que $\text{dom}M_p \subset \text{dom}M_{p,k}$, ya que si $u \in \text{dom}M_p$ entonces $M_p[u] = E_p[e^u] < \infty$, debido a que $e_k^u \leq e^u$ obtenemos $M_{p,k}(u) = E_p[e_k^u] \leq E_p[e^u] < \infty$ y así $u \in \text{dom}M_{p,k}$ y se presenta la inclusión conjuntista $\text{dom}M_p \subset \text{dom}M_{p,k}$.

Teorema 5.5. *El generador de momentos k -deformado $M_{p,k}$ satisface*

1. $M_{p,k}(0) = 1$, para $u \neq 0$ centrada $M_{p,k}(u) > 1$.
2. $M_{p,k}$ es convexa y continua inferiormente.

3. La n -ésima derivada en $u \in \mathcal{O}$ en la dirección de $v \in L^{\phi_k}(p \cdot \mu)$ está dada por

$$\frac{d^n}{dt^n} M_{p,k}(u \overset{k}{\oplus} tv)|_{t=0} = a_n(k) E_p[v^n e_k^u].$$

4. $M_{p,k}$ es analítico en una vecindad de $B_{\phi_k}(0, 1)$.

5. $M_{p,k}$ es infinitamente Frechet diferenciable y la n -ésima derivada en $u \in \mathcal{O}$ en la dirección de $(v_1, v_2, \dots, v_n) \in \mathcal{O}^n$ está dada por

$$D^n M_{p,k}(u)(v_1, v_2, \dots, v_n) = a_n(k) E_p[v_1 \dots v_n e_k^u].$$

Demostración

- i. Si $u = 0$ entonces $M_{p,k}(0) = E_p[e_k^0] = E_p[1] = 1$. Al ser u centrada con $u \neq 0$ se tiene que $E_p[u] = 0$, la exponencial k -deformada es convexa para todo $0 < k < 1$ y $x \in \mathbb{R}$, además $u \in \mathcal{O} \subset L^1(p \cdot \mu)$, lo cual implica que u es integrable, con base en esto se cumple la desigualdad de Jensen 2.2

$$E_p[e_k^u] = \int_{\Omega} e_k^u p d\mu \geq e_k^{\int_{\Omega} u p d\mu} = e_k^{E_p[u]} > 1.$$

Así para u centrada con $u \neq 0$, $M_{p,k}(u) > 1$.

- ii. La demostración se hace análoga a la del teorema (3.6) para lo que es necesario considerar que la función exponencial k -deformada también es convexa.

- iii. Sea $u \in \mathcal{O}$, para cada $v \in L^{\phi_k}(p \cdot \mu)$, $(u \overset{k}{\oplus} tv) \in \text{dom}^{\circ} M_{p,k}$ para t lo suficientemente pequeño, además

$$M_{p,k}(u \overset{k}{\oplus} tv) = E_p \left[e_k^{u \overset{k}{\oplus} tv} \right] = \int_{\Omega} e_k^{u \overset{k}{\oplus} tv} p d\mu = \int_{\Omega} e_k^u e_k^{tv} p d\mu. \quad (1)$$

Si hacemos $q = \frac{e_k^u}{M_{p,k}(u)} p$, entonces por la definición 5.3 se escribe $q = \frac{e_k^u}{E_p[e_k^u]} p$ que con base en el teorema 5.4 obtenemos $q \in \mathfrak{M}_{\mu}$. De acuerdo con la elección de q resulta $e_k^u p = M_{p,k}(u) q$, reemplazando en (1) resulta

$$\begin{aligned} M_{p,k}(u \overset{k}{\oplus} tv) &= \int_{\Omega} e_k^{tv} M_{p,k}(u) q d\mu = M_{p,k}(u) \int_{\Omega} e_k^{tv} q d\mu \\ &= M_{p,k}(u) E_q [e_k^{tv}]. \quad (2) \end{aligned}$$

Por medio de la igualdad presentada en (5.10) escribimos

$$e_k^{tv} = \sum_{j=0}^{\infty} \frac{a_j(k)}{j!} (tv)^j = \sum_{j=0}^{\infty} \frac{a_j(k)}{j!} t^j v^j ,$$

si se integra respecto de la variable aleatoria q , entonces t es constante respecto de la medida $q \cdot \mu$, de allí que

$$E_q [e_k^{tv}] = \sum_{j=0}^{\infty} \frac{a_j(k) E_q[v^j]}{j!} t^j .$$

Sustituyendo en (2)

$$\begin{aligned} M_{p,k}(u \oplus^k tv) &= M_{p,k}(u) \sum_{j=0}^{\infty} \frac{a_j(k) E_q[v^j]}{j!} t^j \\ &= M_{p,k}(u) \left[1 + E_q[v]t + \frac{E_q[v^2]}{2!} t^2 + (1 - k^2) \frac{E_q[v^3]}{3!} t^3 + \dots \right] . \end{aligned}$$

De allí que la n -ésima derivada es

$$\begin{aligned} \frac{d^n}{dt^n} M_{p,k}(u \oplus^k tv) &= M_{p,k}(u) \left[a_n(k) E_q[v^n] + a_{n+1} E_q[v^{n+1}] t + a_{n+2}(k) \frac{E_q[v^{n+2}]}{2!} t^2 + \dots \right] \\ &= M_{p,k}(u) \sum_{j=n}^{\infty} \frac{a_j(k) E_q[v^j]}{(j-n)!} t^{j-n} . \end{aligned}$$

Al evaluar en $t = 0$ y sustituyendo $e_k^u p = M_{p,k}(u) q$ se tiene

$$\begin{aligned} \frac{d^n}{dt^n} M_{p,k}(u \oplus^k tv)|_{t=0} &= M_{p,k}(u) a_n(k) E_q[v^n] = a_n(k) M_{p,k}(u) \int_{\Omega} v^n q d\mu \\ &= a_n(k) \int_{\Omega} M_{p,k}(u) v^n q d\mu = a_n(k) \int_{\Omega} v^n e_k^u p d\mu \\ &= a_k(n) E_p[v^n e_k^u] . \end{aligned}$$

- iv. Sea $u \in \mathcal{O}$, si hacemos que $a = 1$ y $u = 0$ por (5.1) resulta $\lambda_{1,n,k}(0) \in \mathcal{L}_s^n(L^{\phi_k}, L^1)$, por lo que $\hat{\lambda}_{1,n,k} \in \mathcal{P}^n(L^{\phi_k}, L^1)$. Por el teorema (5.2) se sigue que

$$\sum_{n=0}^{\infty} \frac{a_n(k)}{n!} v^n = \sum_{n=0}^{\infty} \frac{a_n(k)}{j! k^n} \lambda_{1,n,k}(0) v^n ,$$

es una serie de potencias con radio de convergencia positivo, donde

$$\sum_{n=0}^{\infty} \frac{a_n(k)}{n!} E_p[\hat{\lambda}_{1,n,k}(u)] \quad (1)$$

Es una serie de potencia con radio de convergencia positivo y es por ello que $\frac{a_n(k)}{n!} E_p[\hat{\lambda}_{1,n,k}(0)]$ pertenece a $\mathcal{L}_s^n(L^{\phi_k}, \mathbb{R})$. Ahora bien haciendo uso de la igualdad presentada en (5.11) y por el hecho de la analiticidad en u_0 resulta

$$\begin{aligned} M_{p,k}(u) &= E_p \left[\sum_{j=0}^{\infty} \frac{a_j(k)}{j!k^j} \lambda_{1,j,k}(u_0)(u - u_0)^j \right] = \sum_{j=0}^{\infty} \frac{a_j(k)}{j!k^j} E_p [\lambda_{1,j,k}(u_0)(u - u_0)^j] \\ &= \sum_{j=0}^{\infty} \frac{a_j(k)}{j!k^j} E_p [|k|^n (u - u_0)^j e_k^{u_0}] = \sum_{j=0}^{\infty} \frac{a_j(k)}{j!} E_p [(u - u_0)^j e_k^{u_0}]. \end{aligned}$$

De acuerdo con la expresión (1) se concluye la analiticidad de la función generador de momentos k -deformada $M_{p,k}$ para $u \in \mathcal{O}$ y $0 < k < 1$. \square

5.3. Funcional acumulante k -deformado

Sea $p \in \mathfrak{M}_\mu$ y $0 < k < 1$, la función generadora de momentos de la variable aleatoria $u \in L^{\phi_k}(p \cdot \mu)$ representada como $\hat{u}_{p,k}$ está dada por

$$\hat{u}_{p,k}(t) = \int_{\Omega} e_k^{tu} p d\mu = E_p[e_k^{tu}].$$

La clase de Cramer k -deformada $\mathcal{C}_{p,k}$ es el conjunto de las variables aleatorias $u \in L^1(p \cdot \mu)$ tales que $\hat{u}_{p,k}(t) < \infty$ en una vecindad del cero, llámese I , así

$$\mathcal{C}_{p,k} = \{u \in L^1(p \cdot \mu) : \hat{u}_{p,k}(t) < \infty \text{ para } t \in I\}.$$

La clase de Cramer k -deformada posee la estructura de espacio lineal el cual es convexo, la demostración se hace similar a la del teorema (3.7), donde la exponencial k -deformada e_k es una función convexa. El siguiente teorema es un análogo al teorema (3.8) en el que se demuestra que el espacio de funciones de Orlicz $L^{\phi_k}(p \cdot \mu)$ es igual, conjuntistamente, a la clase de Cramer k -deformada $\mathcal{C}_{p,k}$.

Teorema 5.6. *Sean $p \in \mathfrak{M}_\mu$ y $0 < k < 1$. La clase de Cramer k -deformada $\mathcal{C}_{p,k}$ es tal que $\mathcal{C}_{p,k} = L^{\phi_k}(p \cdot \mu)$.*

Demostración

Sea $u \in \mathcal{C}_{p,k}$, de allí que $\hat{u}_{p,k}(t) < \infty$ en una vecindad del cero I , por lo que, $E_p[e_k^{tu}] < \infty$ (1) para todo $t \in I$. Sea G una vecindad simétrica de I que contiene al cero por lo que existe $\delta > 0$ tal que $G = (-\delta, \delta) \subset I$. Sea $\frac{1}{r} \in G$ así $\frac{1}{r} \in I$ y por lo tanto satisface (1) de allí

que $E_p \left[e_k^{\frac{u}{r}} \right] < \infty$ (2). Como $-\frac{1}{r} \in G$ de la simetría de G entonces también satisface (1) y escribimos $E_p \left[e_k^{-\frac{u}{r}} \right] < \infty$ (3). Con base en la definición de la función ϕ_k resulta

$$E_p \left[\phi_k \left(\frac{u}{r} \right) \right] = E_p \left[\cosh_k \left(\frac{u}{r} \right) - 1 \right] = \frac{1}{2} E_p \left[e_k^{\frac{u}{r}} + e_k^{-\frac{u}{r}} \right] - 1.$$

Con base en las expresiones (2) y (3) resulta que $E_p \left[\phi_k \left(\frac{u}{r} \right) \right] < \infty$, de donde se sigue que $\|u\|_{\phi_k, p} < \infty$ y así $u \in L^{\phi_k}(p \cdot \mu)$. Resulta entonces la inclusión $\mathcal{C}_{p,k} \subset L^{\phi_k}(p \cdot \mu)$.

Tomemos ahora $u \in L^{\phi_k}(p \cdot \mu)$ entonces existe un $\alpha > 0$ tal que $E_p[\phi_k(\alpha u)] < \infty$, para lo que $E_p[\cosh_k(\alpha u)] < \infty$ teniendo así las expresiones $E_p[e_k^{\alpha u}] < \infty$ (4) y $E_p[e_k^{-\alpha u}] < \infty$ (5). Con base en la existencia de α se puede construir la vecindad $I = (-\alpha, \alpha)$ la cual contiene el cero, sea $t \in I$ entonces $-\alpha < t < \alpha$ (6).

Si $u > 0$ en la expresión (6) se tiene $-\alpha u < tu < \alpha u$ como la exponencial k -deformada es creciente entonces $e_k^{tu} < e_k^{\alpha u}$, de allí que $E_p[e_k^{tu}] < E_p[e_k^{\alpha u}]$ que por la expresión (4) se sigue que $E_p[e_k^{tu}] < \infty$ y así $u \in \mathcal{C}_{p,k}$. Si suponemos ahora que $u < 0$ entonces la expresión (6) se escribe como $\alpha u < tu < -\alpha u$ que razonando análogamente se tiene $E_p[e_k^{tu}] < E_p[e_k^{-\alpha u}] < \infty$ y por lo tanto $u \in \mathcal{C}_{p,k}$. Sin importar el signo de la variable u se sigue que $u \in \mathcal{C}_{p,k}$ y así $L^{\phi_k}(p \cdot \mu) \subset \mathcal{C}_{p,k}$. Con esto se concluye que $\mathcal{C}_{p,k} = L^{\phi_k}(p \cdot \mu)$. \square

Sean $p \in \mathfrak{M}_\mu$ y $0 < k < 1$; sobre la clase de Cramer k -deformada $\mathcal{C}_{p,k}$ se define el conjunto $B_{p,k}$ para el que la esperanza de las variables aleatorias respecto de la medida $p \cdot \mu$ es cero, por el teorema 5.6 se sigue

$$B_{p,k} = \{u \in \mathcal{C}_{p,k} : E_p[u] = 0\} = \{u \in L^{\phi_k}(p \cdot \mu) : E_p[u] = 0\}.$$

A $B_{p,k}$ se le llama clase de Cramer k -deformada centrada, el cual es un subespacio cerrado y convexo de $L^{\phi_k}(p \cdot \mu)$ en forma similar al espacio B_p .

El conjunto $\mathcal{V}_{p,k}$ representa la bola unitaria en el espacio $B_{p,k}$, es decir, $u \in \mathcal{V}_{p,k}$ si y solo si $\|u\|_{\phi_k, p} < 1$. Además, $u \in B_{p,k}$ y $u \in \mathcal{O}$, por lo que $E_p[u] = 0$ y $u \in B_{\phi_k}(0, 1)$ o también $u \in B_{\phi_1}(0, 1)$, esto por las inclusiones (5.4). De acuerdo con esto resulta

$$\mathcal{V}_{p,k} \subset B_{p,k} \subset \mathcal{C}_{p,k} = L^{\phi_k}(p \cdot \mu) \subset L^1(p \cdot \mu).$$

A continuación se define los modelos k -exponenciales unidimensionales de forma similar que en (3.5) considerando la resta k -deformada.

Definición 5.4. Sean $p \in \mathfrak{M}_\mu$ y $0 < k < 1$. Se dice que una función f es un modelo k -exponencial unidimensional si existen $u \in \mathcal{C}_{p,k}$, una función real de variable real ψ tales

que para todo t en un intervalo abierto I que contenga a cero, se cumple que

$$f(t) = e_k^{tu \overset{k}{\ominus} \psi(t)} p.$$

En la definición de modelo k -exponencial paramétrico, t es el parámetro y $f(t)$ es una densidad de probabilidad.

Definición 5.5. *Sea dice que dos densidades de probabilidad $p, z \in \mathfrak{M}_\mu$ están relacionadas por un modelo k -exponencial unidimensional, si existen $r \in \mathfrak{M}_\mu$, $u \in \mathcal{C}_{p,k}$, una función real de variable real ψ y $\lambda > 0$ tal que para todo $t \in (-\lambda, \lambda)$ la función definida por*

$$f(t) = e_k^{tu \overset{k}{\ominus} \psi(t)} r$$

es tal que existen t_0 y t_1 en $(-\lambda, \lambda)$ tales que $f(t_0) = p$ y $f(t_1) = z$.

Teorema 5.7. *Sean $p, z \in \mathfrak{M}_\mu$ dos densidades de probabilidad relacionadas por un modelo k -exponencial unidimensional. Entonces $L^{\phi_k}(p \cdot \mu) = L^{\phi_k}(z \cdot \mu)$.*

Demostración

Como p y z están relacionados por un modelo k -exponencial entonces existe una densidad de probabilidad $r \in \mathfrak{M}_\mu$, una variable aleatoria $u \in \mathcal{C}_{r,k}$, una función real de variable real ϕ tal que para todo $t \in I = (-\lambda, \lambda)$ se tiene

$$f(t) = e_k^{ut \overset{k}{\ominus} \phi(t)} r,$$

donde existen t_0 y t_1 en I tales que $f(t_0) = p$ y $f(t_1) = z$ esto es

$$p = e_k^{ut_0 \overset{k}{\ominus} \phi(t_0)} r \quad \text{y} \quad z = e_k^{ut_1 \overset{k}{\ominus} \phi(t_1)} r. \quad (1)$$

Sin pérdida de generalidad se asumirá que $t_0 < t_1$ por lo que $t_0 < t_1 < \lambda$. Demostremos que $L^{\phi_k}(p \cdot \mu) \subset L^{\phi_k}(z \cdot \mu)$, la inclusión $L^{\phi_k}(z \cdot \mu) \subset L^{\phi_k}(p \cdot \mu)$ se demuestra de forma análoga.

Supongamos ahora que $w \in L^{\phi_k}(p \cdot \mu)$, para concluir que $w \in L^{\phi_k}(z \cdot \mu)$ debe existir una constante positiva β tal que $E_z[\phi_k(\beta w)] < \infty$. Haciendo uso de la igualdad en (1) para β arbitrario y al ser $e_k^{\phi(t_1)}$ un número real entonces

$$E_z[e^{\beta w}] = \int_{\Omega} e_k^{\beta w} z d\mu = \int_{\Omega} e_k^{\beta w} e_k^{ut_1 \overset{k}{\ominus} \phi(t_1)} r d\mu$$

$$= \frac{1}{e_k^{\phi(t_1)}} \int_{\Omega} e_k^{\beta w} e_k^{ut_1} r d\mu = \frac{E_r[e_k^{\beta w} e_k^{ut_1}]}{e_k^{\phi(t_1)}}. \quad (2)$$

Debido a que $z = e_k^{ut_1} \stackrel{k}{\ominus} \phi(t_1) r$ entonces $e_k^{\phi(t_1)} = e_k^{ut_1} \frac{r}{z}$, esto es $E_z[e_k^{\phi(t_1)}] = E_z[e_k^{ut_1} \frac{r}{z}]$; como $e_k^{\phi(t_1)}$ es constante respecto de la medida $z \cdot \mu$ entonces

$$e_k^{\phi(t_1)} = E_z \left[e_k^{ut_1} \frac{r}{z} \right] = \int_{\Omega} \left(e_k^{ut_1} \frac{r}{z} \right) z d\mu = \int_{\Omega} e_k^{ut_1} r d\mu = E_r[e_k^{ut_1}].$$

Al sustituir en (2) se logra la expresión

$$E_z[e_k^{\beta w}] = \frac{E_r[e_k^{\beta w} e_k^{ut_1}]}{E_r[e_k^{ut_1}]},$$

así $E_r[e_k^{\beta w} e_k^{ut_1}] = E_z[e_k^{\beta w}] E_r[e_k^{ut_1}]$ (3). De igual forma se demuestra que $E_r[e_k^{\theta w} e_k^{ut_0}] = E_p[e_k^{\theta w}] E_r[e_k^{ut_0}]$ (4), donde θ es un real positivo. Con base en las expresiones (3) y (4) se construye la función g , dada por

$$\begin{aligned} g: \quad \mathbb{R} \times I &\rightarrow \mathbb{R} \\ (\theta, t) &\mapsto g(\theta, t) = E_r[e_k^{\theta w} e_k^{ut}]. \end{aligned}$$

La función g es convexa por las propiedades de la función exponencial k -deformada. Se analiza ahora el dominio de g .

1. Si $\theta = 0$ entonces $g(0, t) = E_r[e_k^{ut}]$. Como $u \in \mathcal{C}_{r,k}$ entonces $\tilde{u}_{r,k}(t) < \infty$ para $t \in I$; esto es $E_r[e_k^{ut}] < \infty$. Por lo que $g(0, t) < \infty$ y así $(0, t) \in D_g$ para todo $t \in I$.
2. Si $t = t_0$ entonces $g(\theta, t_0) = E_r[e_k^{\beta w} e_k^{ut_0}]$, que por la igualdad (4) se escribe

$$g(\theta, t_0) = E_p[e_k^{\theta w}] E_r[e_k^{ut_0}]. \quad (5)$$

Como $w \in L^{\phi_k}(p \cdot \mu)$, existe una constante positiva $\bar{\theta} > 0$ tal que $E_p[\phi_k(\bar{\theta} w)] < \infty$. Por el teorema (5.6) se sigue que $E_p[e_k^{\bar{\theta} w}] < \infty$, si $\theta \in [-\bar{\theta}, \bar{\theta}]$ entonces $E_p[e_k^{\theta w}] < \infty$.

Como $E_r[e_k^{ut_0}] < \infty$ por ser $u \in \mathcal{C}_{r,k}$ y $t_0 \in I$ entonces, en (5) resulta $g(\theta, t_0) < \infty$ siempre que $\theta \in [-\bar{\theta}, \bar{\theta}]$.

Sea $\theta \in [-\bar{\theta}, \bar{\theta}]$ y α de tal forma que $\alpha \in [0, \bar{\theta}]$ por lo que $0 \leq \alpha \leq \bar{\theta}$. Sea $b \in I$ de tal forma que $t_0 < t_1 < b < \lambda$, la función restricción de g

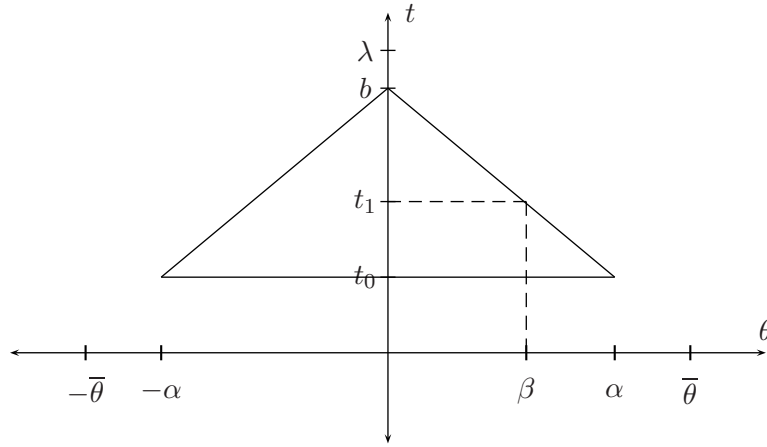
$$g(0, \cdot): \quad I \rightarrow \mathbb{R}$$

$$t \mapsto g(0, t) = E_r[e_k^{ut}].$$

Es finita para todo $t \in I$ en particular para el intervalo $[t_0, b]$. La otra función restricción

$$\begin{aligned} g(\cdot, t_0) : \mathbb{R} &\rightarrow \mathbb{R} \\ \theta &\mapsto g(\theta, t_0) = E_r[e_k^{\theta w} e_k^{ut_0}], \end{aligned}$$

es finita para θ en el intervalo $[-\bar{\theta}, \bar{\theta}]$, para cualquier subintervalo también es finita, en particular para $[-\alpha, \alpha]$. Por tanto, la función g es finita en el interior y la frontera del rectángulo $[-\alpha, \alpha] \times [t_0, b]$, así como en la región triangular cuyos vértices son los puntos (α, t_0) , $(0, b)$ y $(-\alpha, t_0)$ como se ilustra a continuación



Sea β la preimagen de t_1 respecto de la recta determinada por los puntos (α, t_0) y $(0, b)$ cuya ecuación cartesiana es $t = m\theta + b$ donde $m = \frac{t_0 - b}{\alpha} < \alpha$. Para t_1 se escribe $t_1 = \left(\frac{t_0 - b}{\alpha}\right)\beta + b$, para obtener la proporción

$$\frac{t_1 - b}{t_0 - b} = \frac{\beta}{\alpha}, \quad (6)$$

donde $0 < \frac{\beta}{\alpha} < 1$ ya que $t_0 < t_1$. De la convexidad de g sobre el segmento de línea comprendido entre los puntos (α, t_0) y $(0, b)$, existe $0 \leq \eta \leq 1$ tal que

$$g(\beta, t_1) \leq \eta g(\alpha, t_0) + (1 - \eta)g(0, b).$$

Si se hace $\eta = \frac{\beta}{\alpha}$ entonces por la igualdad (6), $1 - \eta = 1 - \frac{\beta}{\alpha} = 1 - \frac{t_1 - b}{t_0 - b} = \frac{t_0 - t_1}{t_0 - b}$. Con base en las igualdades halladas en (3) y (4) se tiene

$$\begin{aligned} g(\beta, t_1) &\leq \left(\frac{t_1 - b}{t_0 - b}\right) g(\alpha, t_0) + \left(\frac{t_0 - t_1}{t_0 - b}\right) g(0, b) \\ E_z[e_k^{\beta w}] E_r[e_k^{ut_1}] &\leq \left(\frac{t_1 - b}{t_0 - b}\right) E_p[e_k^{\alpha w}] E_r[e_k^{ut_0}] + \left(\frac{t_0 - t_1}{t_0 - b}\right) E_r[e_k^{ub}] \end{aligned}$$

$$E_z[e_k^{\beta w}] \leq \left(\frac{t_1 - b}{t_0 - b} \right) \frac{E_p[e_k^{\alpha w}] E_r[e_k^{ut_0}]}{E_r[e_k^{ut_1}]} + \left(\frac{t_0 - t_1}{t_0 - b} \right) \frac{E_r[e_k^{ub}]}{E_r[e_k^{ut_1}]} .$$

Como $b \in I$ entonces $E_r[e_k^{ub}] < \infty$. Además $E_p[e_k^{\alpha w}] < \infty$ puesto que $\alpha \leq \bar{\theta}$ y $E_p[e_k^{\bar{\theta} w}] < \infty$. En la desigualdad anterior resulta $E_z[e_k^{\beta u}] < \infty$, por lo que $E_z[\phi_k(\beta u)] < \infty$. Esto permite garantizar la inclusión conjuntista $L^{\phi_k}(p \cdot \mu) \subset L^{\phi_k}(z \cdot \mu)$ (de igual forma se demuestra la otra inclusión). Se concluye que $L^{\phi_k}(p \cdot \mu) = L^{\phi_k}(z \cdot \mu)$ siempre que las densidades p y z están conectadas por un modelo k -exponencial. \square

Definición 5.6. Sean $p \in \mathfrak{M}_\mu$ y $u \in \mathcal{O}$, se define el funcional acumulante k -deformado como el mapeo

$$\begin{aligned} K_{p,k}(\cdot) : \mathcal{O} &\rightarrow [0, \infty] \\ u &\mapsto K_{p,k}(u) = \ln_k(E_p[e_k^u]) . \end{aligned}$$

De acuerdo con la definición 5.3 se sigue que $K_{k,p}(u) = \ln_k(M_{p,k}(u))$, lo cual es cierto ya que $u \in \mathcal{O} \subset \text{dom} M_{p,k}$. En el siguiente teorema se enuncian las propiedades del funcional acumulante k -deformado.

Teorema 5.8. El funcional acumulante k -deformado $K_{p,k}$ satisface

1. $K_{p,k}(0) = 1$, para $u \neq 0$ centrada $K_{p,k}$ es estrictamente positivo, continuo, creciente y cóncavo.
2. Para cada $u \in \mathcal{O}$ la función $z = \frac{e_k^u}{E_k[e_k^u]} p$ es una densidad de probabilidad en \mathfrak{M}_μ . El funcional $K_{p,k}$ es infinitamente Fréchet diferenciable y la n -ésima derivada en $u \in \mathcal{O}$ en la dirección $(v_1, \dots, v_n) \in \mathcal{O}^n$ de $K_{p,k}$ está dada por

$$D^n K_{p,k}(u) \cdot (v_1, \dots, v_n) = [M_{p,k}(u)]^{1-k} a_n(k) E_z[v_1, \dots, v_n] .$$

3. $K_{p,k}$ es analítica en \mathcal{O} .

Demostración

- i. Si $u = 0$ entonces $K_p(0) = \ln_k(M_{p,k}(0)) = \ln_k 1 = 0$. Si $u \neq 0$ por el teorema 5.5 se tiene $M_{p,k}(u) > 1$, ya que $\ln_k(\cdot)$ es una función creciente para todo $0 < |k| < 1$ entonces $\ln_k(M_{p,k}(u)) > \ln_k 1$, lo cual permite concluir que $K_{p,k}(u) > 0$ esto para $u \neq 0$ y $0 < |k| < 1$.

- ii. De acuerdo con la expresión dada, $z = \frac{e_k^u}{K_{p,k}(u)}p$, como $K_{p,k}(u) = \ln_k(E_p[e_k^u])$ entonces por las propiedades de las funciones deformadas se sigue que $e_k^{K_{p,k}(u)} = E_p[e_k^u]$, así $z = \frac{e_k^u}{E_p[e_k^u]}p$ que por el teorema 5.4 se sigue que $z \in \mathfrak{M}_\mu$. Derivando $K_{p,k}$ y haciendo uso del teorema 5.5 resulta

$$\begin{aligned} D^n K_{p,k}(u) \cdot (v_1, v_2, \dots, v_n) &= \frac{1}{[M_{p,k}(u)]^k} D^n M_{p,k}(u) \cdot (v_1, v_2, \dots, v_n) \\ &= [M_{p,k}(u)]^{-k} a_n(k) E_p[v_1 \dots v_n e_k^u]. \end{aligned}$$

Como $z = \frac{e_k^u}{M_{p,k}(u)}p$ entonces $e_k^u p = M_{p,k}(u)z$ en la expresión anterior se tiene

$$\begin{aligned} D^n K_{p,k}(u) \cdot (v_1, v_2, \dots, v_n) &= [M_{p,k}(u)]^{-k} a_n(k) \int_{\Omega} v_1 \dots v_n e_k^u p d\mu \\ &= [M_{p,k}(u)]^{-k} a_n(k) \int_{\Omega} v_1 \dots v_n M_{p,k}(u) z d\mu \\ &= [M_{p,k}(u)]^{1-k} a_n(k) E_z[v_1 \dots v_n], \end{aligned}$$

el cual era el objetivo de la demostración.

- iii. En el teorema 5.5 se demostró que $M_{p,k}$ es analítico en $B_{\phi_k}(0, 1)$ y $\ln_k(\cdot)$ también es analítica por lo que $K_{p,k}(u) = \ln_k(M_{p,k}(u))$ también es analítica (composición de funciones analíticas). \square

5.4. Variedad de información k -deformada

En el capítulo 4 se indicó que $E_p \left[\left(\frac{q}{p} \right)^k \right]$ es finito para dos densidades arbitrarias p y q en \mathfrak{M}_μ , pero se asume que $E_p \left[\left(\frac{p}{q} \right)^k \right]$ es finita. Situación análoga se supondrá en esta sección.

Definición 5.7. Sean $p \in \mathfrak{M}_\mu$, $0 < k < 1$ y $u \in \mathcal{V}_{p,k}$, se define el mapeo

$$\begin{aligned} e_{p,k}(\cdot) : \mathcal{V}_{p,k} &\rightarrow \mathfrak{M}_\mu \\ u &\mapsto q = e_{p,k}(u) \end{aligned}$$

dado por $q = e_{p,k}(u) = e_k^{u \ominus K_{p,k}(u)} p$.

Nótese que dicho mapeo está bien definido ya que en el teorema (5.4) se concluyó que $q = e_k^{u \ominus \ln_k(E_p[e_k^u])} p$ es una densidad de probabilidad donde $\ln_k(E_p[e_k^u])$ es el acumulante

k -deformado.

Sean $p \in \mathfrak{M}_\mu$ y $0 < k < 1$, para el mapeo $e_{p,k}$ se denota $\mathcal{U}_{p,k}$ como el rango de dicho mapeo, lo cual se escribe como $e_{p,k} : \mathcal{V}_{p,k} \rightarrow \mathcal{U}_{p,k}$ y así $e_{p,k}(\mathcal{V}_{p,k}) = \mathcal{U}_{p,k} \subseteq \mathfrak{M}_\mu$. En este caso, $p \in \mathcal{U}_{p,k}$ haciendo $u = 0$ casi en todas partes. En la definición 5.7 se sigue que $e_{p,k}$ es un modelo k -exponencial no paramétrico ya que $K_{p,k}(u)$ es una función real.

Lema 5.3. *Sea $0 < k < 1$. La función $g(x) = \frac{x}{\sqrt{1+k^2x^2}}$ es invertible y acotada para todo valor de x .*

Demostración

Debido que $|kx| = k|x| < \sqrt{1+k^2x^2}$ para todo $x \in \mathbb{R}$ y $0 < k < 1$ entonces se sigue que $\left| \frac{x}{\sqrt{1+k^2x^2}} \right| = \frac{|x|}{\sqrt{1+k^2x^2}} < \frac{1}{k}$, con lo cual se obtiene que $|g(x)| < \frac{1}{k}$ y por tanto para todo x la función $g(x)$ está acotada, donde $y = \frac{1}{k}$ y $y = -\frac{1}{k}$ actúan como asíntotas horizontales, siendo $\frac{1}{k} > 1$; se tiene además que la función es continua para cualquier valor de x por ser la composición de funciones continuas.

Para determinar que g es inyectiva se supondrá que $g(x_1) = g(x_2)$ para concluir que $x_1 = x_2$, por medio de la definición de la función g se tiene la igualdad

$$\frac{x_1}{\sqrt{1+k^2x_1^2}} = \frac{x_2}{\sqrt{1+k^2x_2^2}}, \quad (1)$$

La cual se puede escribir como $x_1\sqrt{1+k^2x_2^2} = x_2\sqrt{1+k^2x_1^2}$, se eleva al cuadrado y distribuye logrando la igualdad $x_1^2 + k^2x_1^2x_2^2 = x_2^2 + k^2x_2^2x_1^2$, así $x_1^2 = x_2^2$; se afirma que $x_1 = x_2$, ya que si $x_1 = -x_2$ entonces al reemplazar en (1) los denominadores se cancelan y se tiene $-x_2 = x_2$ cuya única solución es $x_2 = 0$, por lo que $x_1 = x_2$ en términos generales y así $g(x)$ es una función uno a uno y por tanto invertible para cada valor de x . Para hallar la inversa se hace $g(x) = y$ y se despeja la x para tener

$$x = \frac{y}{\sqrt{1-k^2y^2}}.$$

Y por tanto la inversa de g es $g^{-1}(x) = \frac{x}{\sqrt{1-k^2x^2}}$ siempre que $k^2x^2 < 1$, esta última desigualdad implica que $-\frac{1}{k} < x < \frac{1}{k}$. \square

Ya que la función g del lema 5.3 es invertible entonces la ecuación $g(x) = -a$ tiene solución única siempre que $-\frac{1}{k} < a < \frac{1}{k}$ y la solución es

$$x = g^{-1}(-a) = \frac{-a}{\sqrt{1-k^2a^2}}. \quad (5.12)$$

El propósito ahora es hallar la inversa del mapeo $q = e_{p,k}(u) = e_k^{u \ominus^k K_{p,k}(u)} p$ definido en (5.7). Como $\frac{q}{p} = e_k^{u \ominus^k K_{p,k}(u)}$ entonces se tiene la expresión

$$\ln_k \left(\frac{q}{p} \right) = u \ominus^k K_{p,k}(u) ,$$

que es equivalente a

$$u = \ln_k \left(\frac{q}{p} \right) \oplus^k K_{p,k}(u) . \quad (5.13)$$

Puesto que u tiene esperanza cero respecto de la medida $p \cdot \mu$ entonces

$$E_p \left[\ln_k \left(\frac{q}{p} \right) \oplus^k K_{p,k}(u) \right] = 0 .$$

Con base en la definición de la suma k -deformada se logra la expresión

$$E_p \left[\ln_k \left(\frac{q}{p} \right) \sqrt{1 + k^2 K_{p,k}^2(u)} + K_{p,k}(u) \sqrt{1 + k^2 \ln_k^2 \left(\frac{q}{p} \right)} \right] = 0 .$$

Debido a que la integral es un operador lineal y $K_{p,k}(u)$ y $\sqrt{1 + k^2 K_{p,k}^2(u)}$ son constantes respecto de la medida $p \cdot \mu$ entonces la anterior igualdad se escribe como

$$\sqrt{1 + k^2 K_{p,k}^2(u)} E_p \left[\ln_k \left(\frac{q}{p} \right) \right] + K_{p,k}(u) E_p \left[\sqrt{1 + k^2 \ln_k^2 \left(\frac{q}{p} \right)} \right] = 0 ,$$

equivalente a

$$\frac{K_{p,k}(u)}{\sqrt{1 + k^2 K_{p,k}^2(u)}} = - \frac{E_p \left[\ln_k \left(\frac{q}{p} \right) \right]}{E_p \left[\sqrt{1 + k^2 \ln_k^2 \left(\frac{q}{p} \right)} \right]} . \quad (5.14)$$

Consideremos la función $Z_{p,k}$, la cual depende de la densidad de probabilidad q como

$$Z_{p,k}(q) = \frac{E_p \left[\ln_k \left(\frac{q}{p} \right) \right]}{E_p \left[\sqrt{1 + k^2 \ln_k^2 \left(\frac{q}{p} \right)} \right]} . \quad (5.15)$$

En el caso en que $k = 0$ la función $Z_{p,0}$ se escribe como $Z_{p,0}(q) = E_p \left[\ln \left(\frac{q}{p} \right) \right]$. En la igualdad (5.14) al reemplazar por la función $Z_{p,k}$ se tiene la expresión

$$\frac{K_{p,k}(u)}{\sqrt{1 + k^2 K_{p,k}^2(u)}} = -Z_{p,k}(q) ,$$

que en términos de la función g definida en el lema 5.3, se tiene el problema $g(K_{p,k}(u)) = -Z_{p,k}(q)$ cuya solución se presentó en la ecuación 5.12 para tener

$$K_{p,k}(u) = -\frac{Z_{p,k}(q)}{\sqrt{1 - k^2 Z_{p,k}^2(q)}}. \quad (5.16)$$

Se debe verificar entonces que $|Z_{p,k}(q)| < \frac{1}{k}$ para garantizar la existencia de la solución del problema. Como se verifica la desigualdad

$$k \ln_k \left(\frac{q}{p} \right) < \sqrt{1 + k^2 \ln_k^2 \left(\frac{q}{p} \right)},$$

de la monotonía de la integral resulta $kE_p \left[\ln_k \left(\frac{q}{p} \right) \right] < E_p \left[\sqrt{1 + k^2 \ln_k^2 \left(\frac{q}{p} \right)} \right]$ que permite concluir que $Z_{p,k}(q) < \frac{1}{k}$ y así la expresión en (5.16) está bien definida. Sustituyendo el valor del funcional acumulante (5.16) en la expresión (5.13) se sigue que una variable aleatoria centrada u se escribe a través de la densidad q como

$$u = \ln_k \left(\frac{q}{p} \right) \oplus \frac{-Z_{p,k}(q)}{\sqrt{1 - k^2 Z_{p,k}^2(q)}}.$$

Por las propiedades de la suma k -deformada se puede escribir la anterior igualdad como

$$u = \ln_k \left(\frac{q}{p} \right) \ominus \frac{Z_{p,k}(q)}{\sqrt{1 - k^2 Z_{p,k}^2(q)}} \quad (5.17)$$

El resultado anterior, donde se logró despejar la variable aleatoria u en términos de la densidad q motiva la siguiente definición.

Definición 5.8. Sea $p \in \mathfrak{M}_\mu$, $0 < k < 1$ y $q \in \mathfrak{M}_\mu$, se define el mapeo

$$\begin{aligned} s_{p,k}(\cdot) : \mathfrak{M}_\mu &\rightarrow B_{p,k} \\ q &\mapsto u = s_{p,k}(q) \end{aligned}$$

como $u = s_{p,k}(q) = \ln_k \left(\frac{q}{p} \right) \ominus \frac{Z_{p,k}(q)}{\sqrt{1 - k^2 Z_{p,k}^2(q)}}.$

En el caso en que $k = 0$ se dijo que $Z_{p,0}(q) = E_p \left[\ln \left(\frac{q}{p} \right) \right]$ y por tanto el mapeo $s_{p,k}$ en este valor equivale a $s_{p,0}(q) = \ln \left(\frac{q}{p} \right) - E_p \left[\ln \left(\frac{q}{p} \right) \right]$ que coincide con los mapeos que generan las cartas para \mathfrak{M}_μ en la variedad de Pistone y Sempì usual. A continuación se encontrarán

expresiones equivalentes para $s_{p,k}$.

Por la definición de resta k -deformada, el mapeo $s_{p,k}$ se escribe como

$$\begin{aligned} s_{p,k}(q) &= \ln_k \left(\frac{q}{p} \right) \sqrt{1 + k^2 \left(\frac{Z_{p,k}(q)}{\sqrt{1 - k^2 Z_{p,k}^2(q)}} \right)^2} - \frac{Z_{p,k}(q)}{\sqrt{1 - k^2 Z_{p,k}^2(q)}} \sqrt{1 + k^2 \ln_k^2 \left(\frac{q}{p} \right)} \\ &= \frac{1}{\sqrt{1 - k^2 Z_{p,k}^2(q)}} \ln_k \left(\frac{q}{p} \right) - \frac{Z_{p,k}(q)}{\sqrt{1 - k^2 Z_{p,k}^2(q)}} \sqrt{1 + k^2 \ln_k^2 \left(\frac{q}{p} \right)} \\ &= \frac{1}{\sqrt{1 - k^2 Z_{p,k}^2(q)}} \left[\ln_k \left(\frac{q}{p} \right) - Z_{p,k}(q) \sqrt{1 + k^2 \ln_k^2 \left(\frac{q}{p} \right)} \right] \end{aligned}$$

Se reemplaza la expresión de $Z_{p,k}$ definida en (5.15) para tener en la anterior igualdad que

$$s_{p,k}(q) = \frac{1}{\sqrt{1 - k^2 Z_{p,k}^2(q)}} \left[\ln_k \left(\frac{q}{p} \right) - \frac{\sqrt{1 + k^2 \ln_k^2 \left(\frac{q}{p} \right)}}{E_p \left[\sqrt{1 + k^2 \ln_k^2 \left(\frac{q}{p} \right)} \right]} E_p \left[\ln_k \left(\frac{q}{p} \right) \right] \right] \quad (5.18)$$

En la igualdad (5.18) se tiene que $E_p[s_{p,k}(q)] = 0$, lo cual indica que la elección de que $u = s_{p,k}(q)$ está bien definida, es decir, dada una densidad de probabilidad q se obtiene una variable aleatoria u centrada.

Los procedimientos siguientes están encaminados a hallar otras expresiones para la función $Z_{p,k}$. Con base en la elección de $Z_{p,k}$ en (5.15) se tienen las siguientes igualdades, donde se asumirá la notación $E_p^2[\cdot]$ para indicar que el valor esperado se eleva al cuadrado ($E_p[\cdot]$)².

$$1 - k^2 Z_{p,k}^2(q) = 1 - \frac{k^2 E_p^2 \left[\ln_k \left(\frac{q}{p} \right) \right]}{E_p^2 \left[\sqrt{1 + k^2 \ln_k^2 \left(\frac{q}{p} \right)} \right]} = \frac{E_p^2 \left[\sqrt{1 + k^2 \ln_k^2 \left(\frac{q}{p} \right)} \right] - k^2 E_p^2 \left[\ln_k \left(\frac{q}{p} \right) \right]}{E_p^2 \left[\sqrt{1 + k^2 \ln_k^2 \left(\frac{q}{p} \right)} \right]}$$

Por diferencia de cuadrados y linealidad de la integral de Lebesgue se obtiene la expresión

$$1 - k^2 Z_{p,k}^2(q) = \frac{E_p \left[\sqrt{1 + k^2 \ln_k^2 \left(\frac{q}{p} \right)} - k \ln_k \left(\frac{q}{p} \right) \right] E_p \left[\sqrt{1 + k^2 \ln_k^2 \left(\frac{q}{p} \right)} + k \ln_k \left(\frac{q}{p} \right) \right]}{E_p^2 \left[\sqrt{1 + k^2 \ln_k^2 \left(\frac{q}{p} \right)} \right]} \quad (5.19)$$

Por la definición de la exponencial k -deformada aplicada a $\ln_k \left(\frac{q}{p} \right)$ se tiene

$$e_k^{\ln_k \left(\frac{q}{p} \right)} = \left(k \ln_k \left(\frac{q}{p} \right) + \sqrt{1 + k^2 \ln_k^2 \left(\frac{q}{p} \right)} \right)^{\frac{1}{k}}.$$

Esta igualdad es equivalente a tener

$$\left(\frac{q}{p} \right)^k = k \ln_k \left(\frac{q}{p} \right) + \sqrt{1 + k^2 \ln_k^2 \left(\frac{q}{p} \right)}. \quad (5.20)$$

Por situación análoga y sabiendo que $-\ln_k \left(\frac{q}{p} \right) = \ln_k \left(\frac{p}{q} \right)$ se tiene

$$e_k^{\ln_k \left(\frac{p}{q} \right)} = e_k^{-\ln_k \left(\frac{q}{p} \right)} = \left(-k \ln_k \left(\frac{q}{p} \right) + \sqrt{1 + k^2 \ln_k^2 \left(\frac{q}{p} \right)} \right)^{\frac{1}{k}}$$

y obtener así una expresión para $\left(\frac{p}{q} \right)^k$ como

$$\left(\frac{p}{q} \right)^k = -k \ln_k \left(\frac{q}{p} \right) + \sqrt{1 + k^2 \ln_k^2 \left(\frac{q}{p} \right)}. \quad (5.21)$$

Al reemplazar las igualdades (5.20) y (5.21) en la expresión (5.19) se tiene que

$$1 - k^2 Z_{p,k}^2(q) = \frac{E_p \left[\left(\frac{q}{p} \right)^k \right] E_p \left[\left(\frac{p}{q} \right)^k \right]}{E_p^2 \left[\sqrt{1 + k^2 \ln_k^2 \left(\frac{q}{p} \right)} \right]},$$

para tener así la expresión

$$\sqrt{1 - k^2 Z_{p,k}^2(q)} = \frac{\sqrt{E_p \left[\left(\frac{q}{p} \right)^k \right] E_p \left[\left(\frac{p}{q} \right)^k \right]}}{E_p \left[\sqrt{1 + k^2 \ln_k^2 \left(\frac{q}{p} \right)} \right]}. \quad (5.22)$$

Por la definición de la función $Z_{p,k}$ en (5.15) y la igualdad (5.22) se tiene entonces que

$$K_{p,k}(u) = -\frac{Z_{p,k}(q)}{\sqrt{1 - k^2 Z_{p,k}^2(q)}} = -\frac{E_p \left[\ln_k \left(\frac{q}{p} \right) \right]}{\sqrt{E_p \left[\left(\frac{q}{p} \right)^k \right] E_p \left[\left(\frac{p}{q} \right)^k \right]}}. \quad (5.23)$$

Consideremos ahora la función $h_{p,k}(q)$ definida como

$$h_{p,k}(q) = \frac{1}{\sqrt{E_p \left[\left(\frac{q}{p} \right)^k \right] E_p \left[\left(\frac{p}{q} \right)^k \right]}}, \quad (5.24)$$

donde $h_{p,k}(q)$ es un número real de tal forma que si $k = 0$ entonces $h_{p,0}(q) = 1$ y en el caso en que $q = p$ casi en todas partes de Ω entonces $h_{p,k}(p) = 1$ independiente del valor de k . La igualdad (5.23) se escribe como

$$K_{p,k}(u) = -\frac{Z_{p,k}(q)}{\sqrt{1 - k^2 Z_{p,k}^2(q)}} = -h_{p,k}(q) E_p \left[\ln_k \left(\frac{q}{p} \right) \right] \quad (5.25)$$

Con la expresión (5.25) es posible expresar el mapeo $s_{p,k}$ definido en (5.8) como

$$s_{p,k}(q) = \ln_k \left(\frac{q}{p} \right) \ominus^k h_{p,k}(q) E_p \left[\ln_k \left(\frac{q}{p} \right) \right]. \quad (5.26)$$

Sumando ahora las expresiones obtenidas en (5.20) y (5.21) se tiene que

$$\sqrt{1 + k^2 \ln_k^2 \left(\frac{q}{p} \right)} = \frac{1}{2} \left[\left(\frac{q}{p} \right)^k + \left(\frac{p}{q} \right)^k \right], \quad (5.27)$$

por lo que la igualdad 5.22 se escribe como

$$\sqrt{1 - k^2 Z_{p,k}^2(q)} = \frac{\sqrt{E_p \left[\left(\frac{q}{p} \right)^k \right] E_p \left[\left(\frac{p}{q} \right)^k \right]}}{\frac{1}{2} \left(E_p \left[\left(\frac{q}{p} \right)^k \right] + E_p \left[\left(\frac{p}{q} \right)^k \right] \right)},$$

debido que $\sqrt{1 - k^2 Z_{p,k}^2(q)} \leq 1$ para todo $0 < k < 1$ entonces en la anterior igualdad se concluye la propiedad

$$\sqrt{E_p \left[\left(\frac{q}{p} \right)^k \right] E_p \left[\left(\frac{p}{q} \right)^k \right]} \leq \frac{1}{2} \left(E_p \left[\left(\frac{q}{p} \right)^k \right] + E_p \left[\left(\frac{p}{q} \right)^k \right] \right); \quad (5.28)$$

es decir, la media geométrica entre $E_p \left[\left(\frac{q}{p} \right)^k \right]$ y $E_p \left[\left(\frac{p}{q} \right)^k \right]$ no excede su media aritmética.

Teorema 5.9. Sean $p \in \mathfrak{M}_\mu$ y $0 < k < 1$ entonces el mapeo $e_{p,k}$ es inyectivo.

Demostración

Sean u_1 y u_2 variables aleatorias en $B_{p,k}$ tales que $e_{p,k}(u_1) = e_{p,k}(u_2)$ se demuestra que $u_1 = u_2$. Sea $q = e_{p,k}(u_1)$ por lo que $q = e_{p,k}(u_2)$, por la definición del mapeo $e_{p,k}(\cdot)$ se tienen las igualdades

$$q = e_k^{u_1 \oplus K_{p,k}(u_1)} p \quad \wedge \quad q = e_k^{u_2 \oplus K_{p,k}(u_2)} p ,$$

expresiones equivalentes a

$$u_1 = \ln_k \left(\frac{q}{p} \right) \oplus^k K_{p,k}(u_1) \quad \wedge \quad u_2 = \ln_k \left(\frac{q}{p} \right) \oplus^k K_{p,k}(u_2) . \quad (1)$$

Por un proceso análogo que condujo a la expresión (5.16), siendo u_1 y u_2 variables aleatorias centradas, resulta

$$K_{p,k}(u_1) = -\frac{Z_{p,k}(q)}{\sqrt{1 - k^2 Z_{p,k}^2(q)}} \quad \wedge \quad K_{p,k}(u_2) = -\frac{Z_{p,k}(q)}{\sqrt{1 - k^2 Z_{p,k}^2(q)}} ,$$

por lo tanto, $K_{p,k}(u_1) = K_{p,k}(u_2)$ y al sumarle $\ln_k \left(\frac{q}{p} \right)$ con respecto a la k -deformación se tiene

$$\ln_k \left(\frac{q}{p} \right) \oplus^k K_{p,k}(u_1) = \ln_k \left(\frac{q}{p} \right) \oplus^k K_{p,k}(u_2) ,$$

que por las expresiones en (1) se concluye que $u_1 = u_2$ y por lo tanto el mapeo $e_{p,k}(\cdot)$ es inyectivo. \square

Como el mapeo $e_{p,k}$ es inyectivo de acuerdo con lo demostrado en el teorema (5.9) y además es sobreyectivo sobre el conjunto $\mathcal{U}_{p,k}$ entonces posee una inversa, la cual se demostrará que es efectivamente $s_{p,k}$ definido en (5.8), para lo cual se debe demostrar que $s_{p,k}(e_{p,k}(u)) = u$ y $e_{p,k}(s_{p,k}(q)) = q$ siendo $u \in B_{p,k}$ y $q \in \mathcal{U}_{p,k} \subseteq \mathfrak{M}_\mu$. Para llegar a estas igualdades es necesario hacer algunas deducciones previas.

Como $e_{p,k}(u) = e_k^{u \oplus K_{p,k}(u)} p$, se obtiene

$$\ln_k \left(\frac{e_{p,k}(u)}{p} \right) = u \oplus^k K_{p,k}(u) . \quad (5.29)$$

Ya que $u \in B_{p,k}$ entonces $E_p[u] = 0$ y por las propiedades de la suma k -deformada se tiene $u = (u \oplus^k K_{p,k}(u)) \oplus^k K_{p,k}(u)$, que al utilizar la igualdad (5.29) se escribe $u = \ln_k \left(\frac{e_{p,k}(u)}{p} \right) \oplus^k K_{p,k}(u)$ y así

$$E_p \left[\ln_k \left(\frac{e_{p,k}(u)}{p} \right) \oplus^k K_{p,k}(u) \right] = 0 .$$

Por medio de la definición de la suma k -deformada la expresión anterior equivale a

$$E_p \left[\ln_k \left(\frac{e_{p,k}(u)}{p} \right) \sqrt{1 + k^2 K_{p,k}^2(u)} + K_{p,k}(u) \sqrt{1 + k^2 \ln_k^2 \left(\frac{e_{p,k}(u)}{p} \right)} \right] = 0 ,$$

de donde

$$\frac{K_{p,k}(u)}{\sqrt{1 + k^2 K_{p,k}^2(u)}} = - \frac{E_p \left[\ln_k \left(\frac{e_{p,k}(u)}{p} \right) \right]}{E_p \left[\sqrt{1 + k^2 \ln_k^2 \left(\frac{e_{p,k}(u)}{p} \right)} \right]} .$$

En términos de la función g del lema (5.3) y la escogencia de la función $Z_{p,k}$ se tiene el problema $g(K_{p,k}(u)) = -Z_{p,k}(e_{p,k}(u))$ cuya solución es $K_{p,k}(u) = -g^{-1}(Z_{p,k}(e_{p,k}(u)))$ dada por

$$K_{p,k}(u) = \frac{-Z_{p,k}(e_{p,k}(u))}{\sqrt{1 - k^2 Z_{p,k}^2(e_{p,k}(u))}} . \quad (5.30)$$

Por otro lado, con base en la definición del mapeo $s_{p,k}$ en (5.8) se escribe

$$\ln_k \left(\frac{q}{p} \right) = s_{p,k}(q) \oplus^k \frac{Z_{p,k}(q)}{\sqrt{1 - k^2 Z_{p,k}^2(q)}} ,$$

a través de la definición de la exponencial k -deformada y sus propiedades se tiene

$$\frac{q}{p} = e_k^{s_{p,k}(q)} e_k \left(\frac{Z_{p,k}(q)}{\sqrt{1 - k^2 Z_{p,k}^2(q)}} \right) ;$$

como $E_p \left[\frac{q}{p} \right] = E_q[1] = 1$ entonces la anterior igualdad es equivalente a

$$e_k \left(\frac{Z_{p,k}(q)}{\sqrt{1 - k^2 Z_{p,k}^2(q)}} \right) E_p \left[e_k^{s_{p,k}(q)} \right] = 1 . \quad (5.31)$$

Se demuestra ahora que efectivamente $s_{p,k}$ es el mapeo inverso del mapeo $e_{p,k}$.

Teorema 5.10. Sean $p \in \mathfrak{M}_\mu$ y $0 < k < 1$. El mapeo $s_{p,k}$ es el inverso del mapeo $e_{p,k}$.

Demostración

Para que $s_{p,k}$ sea la inversa de $e_{p,k}$ se debe demostrar que si $u \in B_{p,k}$ entonces $(s_{p,k} \circ e_{p,k})(u) = u$, esto es, $s_{p,k}(e_{p,k}(u)) = u$; por la definición del mapeo $s_{p,k}$ resulta

$$s_{p,k}(e_{p,k}(u)) = \ln_k \left(\frac{e_{p,k}(u)}{p} \right) \ominus \frac{Z_{p,k}(e_{p,k}(u))}{\sqrt{1 - k^2 Z_{p,k}^2(e_{p,k}(u))}}.$$

Por las expresiones halladas en (5.29) y (5.30) la igualdad anterior se resume a

$$\begin{aligned} s_{p,k}(e_{p,k}(u)) &= (u \ominus K_{p,k}(u)) \ominus (-K_{p,k}(u)) \\ &= u \oplus ((-K_{p,k}(u) \oplus K_{p,k}(u))) = u \oplus 0 = u, \end{aligned}$$

esto es independiente de la variable aleatoria u . Consideremos ahora una densidad $q \in \mathcal{U}_{p,k}$ se debe demostrar que $(e_{p,k} \circ s_{p,k})(q) = q$, esto es, $e_{p,k}(s_{p,k}(q)) = q$; por la definición del mapeo $e_{p,k}$ resulta

$$e_{p,k}(s_{p,k}(q)) = e_k^{s_{p,k}(q)} \ominus K_{p,k}(s_{p,k}(q)) p;$$

por las propiedades de la función k -exponencial se escribe

$$e_{p,k}(s_{p,k}(q)) = \frac{e_k^{s_{p,k}(q)}}{e_k^{K_{p,k}(s_{p,k}(q))}} p. \quad (1)$$

Por la definición del funcional acumulante k -deformado en (5.6) resulta que $e_k^{K_{p,k}(s_{p,k}(q))} = E_p \left[e_k^{s_{p,k}(q)} \right]$ al sustituir esta expresión en (1) y definir el mapeo $s_{p,k}$ se logra

$$e_{p,k}(s_{p,k}(q)) = \frac{e_k \left[\ln_k \left(\frac{q}{p} \right) \ominus \frac{Z_{p,k}(q)}{\sqrt{1 - k^2 Z_{p,k}^2(q)}} \right]}{E_p \left[e_k^{s_{p,k}(q)} \right]} p = \frac{e_k \left(\frac{q}{p} \right)}{\left(\frac{Z_{p,k}(q)}{\sqrt{1 - k^2 Z_{p,k}^2(q)}} \right) E_p \left[e_k^{s_{p,k}(q)} \right]} p.$$

Por la igualdad obtenida en (5.31) y al ser la exponencial y la logarítmica k -deformadas inversas se llega a que $e_{p,k}(s_{p,k}(q)) = \frac{q}{p} p = q$; lo cual se logra independiente de la densidad $q \in \mathcal{U}_{p,k}$. Se concluye que los mapeos $e_{p,k}$ y $s_{p,k}$ son inversos uno del otro. \square

Con base en el teorema (5.10) se sigue que la familia

$$\mathcal{F} = \{(\mathcal{U}_{p,k}, s_{p,k}) : p \in \mathfrak{M}_\mu, 0 < k < 1\}$$

es un atlas para el espacio \mathfrak{M}_μ , ya que: Si $0 < k < 1$ y p_1, p_2 son densidades de probabilidad en el espacio lineal convexo \mathfrak{M}_μ , tales que $\mathcal{U}_{p_1,k} \cap \mathcal{U}_{p_2,k} \neq \emptyset$. Los mapeos transición entre los

conjuntos $\mathcal{V}_{p_1,k}$ y $\mathcal{V}_{p_2,k}$ están representados por las composiciones $s_{p_2,k} \circ e_{p_1,k}$ y su inversa $s_{p_1,k} \circ e_{p_2,k}$.

Sea $u \in B_{p_1,k}$ tal que $e_{p_1,k}(u) = q$ donde $q \in \mathcal{U}_{p_1,k}$ y define la variable $\bar{u} \in B_{p_2,k}$ tal que $s_{p_2,k}(q) = \bar{u}$ de donde $(s_{p_2,k} \circ e_{p_1,k})(u) = \bar{u}$. Para escribir el mapeo transición se hará uso de la expresión (5.26), donde

$$\bar{u} = (s_{p_2,k} \circ e_{p_1,k})(u) = \ln_k \left(\frac{e_{p_1,k}(u)}{p_2} \right) \stackrel{k}{\ominus} h_{p_2,k}(e_{p_1,k}(u)) E_{p_2} \left[\ln_k \left(\frac{e_{p_1,k}(u)}{p_2} \right) \right]. \quad (5.32)$$

Por la definición del mapeo $e_{p_1,k}$ y las propiedades del logaritmo deformado se tiene

$$\begin{aligned} \ln_k \left(\frac{e_{p_1,k}(u)}{p_2} \right) &= \ln_k \left(\frac{e_k^{u \stackrel{k}{\ominus} K_{p_1,k}(u)} p_1}{p_2} \right) = \ln_k \left(e_k^{u \stackrel{k}{\ominus} K_{p_1,k}(u)} \right) \stackrel{k}{\oplus} \ln_k \left(\frac{p_1}{p_2} \right) \\ &= \left(u \stackrel{k}{\ominus} K_{p_1,k}(u) \right) \stackrel{k}{\oplus} \ln_k \left(\frac{p_1}{p_2} \right) = u \stackrel{k}{\oplus} \left(-K_{p_1,k}(u) \stackrel{k}{\oplus} \ln_k \left(\frac{p_1}{p_2} \right) \right). \end{aligned}$$

El término $-K_{p_1,k}(u) \stackrel{k}{\oplus} \ln_k \left(\frac{p_1}{p_2} \right)$ es un número real para cada variable aleatoria u , entonces se puede denotar como A ; así, el mapeo transición se escribir como

$$(s_{p_2,k} \circ e_{p_1,k})(u) = \left(u \stackrel{k}{\oplus} A \right) \stackrel{k}{\ominus} h_{p_2,k}(e_{p_1,k}(u)) E_{p_2} \left[u \stackrel{k}{\oplus} A \right]$$

Por medio de la definición de la suma k -deformada, el término $u \stackrel{k}{\oplus} A$, se escribe como $u \stackrel{k}{\oplus} A = u\sqrt{1+k^2A^2} + A\sqrt{1+k^2u^2}$. Sean $B = \sqrt{1+k^2A^2}$ que es un número real, $f(u) = u$ y $g(u) = \sqrt{1+k^2u^2}$, donde f y g son funciones que dependen de la variable aleatoria centrada u , es por ello que $u \stackrel{k}{\oplus} A = Bf(u) + Ag(u)$. Las funciones $f(u)$ y $g(u)$ son analíticas, el mapeo transición se escribe como

$$(s_{p_2,k} \circ e_{p_1,k})(u) = (Bf(u) + Ag(u)) \stackrel{k}{\ominus} h_{p_2,k}(e_{p_1,k}(u)) E_{p_2} [Bf(u) + Ag(u)]. \quad (5.33)$$

Se concluye que la pareja $(\mathfrak{M}_\mu, \mathcal{F})$ es una variedad topológica modelada por los espacios de Banach $L^{\phi_k}(p \cdot \mu)$.

Capítulo 6

Problemas Abiertos

1. Demostrar la diferenciabilidad de los mapeos transición $s_{p_2,k} \circ e_{p_1,k}$ obtenidos en (5.33) entre los espacios \mathcal{V}_{p_1} y \mathcal{V}_{p_2} , para lograr obtener una variedad diferenciable sobre el espacio de densidades de probabilidad estrictamente positivas.
2. En la variedad de información estándar, se construye la estructura de variedad para el espacio \mathfrak{M}_μ (densidades estrictamente positivas) a través del espacio de Banach $L^{\phi_1}(p \cdot \mu)$, donde $\phi_1(x) = \cosh x - 1$. Alberto Cena en su tesis doctoral [5], hace la construcción de la variedad mixta, es decir, una extensión de la variedad construida para \mathfrak{M}_μ al espacio \mathcal{P} que está conformada por las densidades de probabilidad de cualquier signo. Tal construcción se hace a partir de la función de Young $\phi_3(x) = (1+|x|) \ln(1+|x|) - |x|$ que forma un par conjugado con $\phi_2(x) = e^{|x|} - |x| - 1$ y ésta, a su vez, es equivalente a la función ϕ_1 . De otro lado, G. Pistone en su artículo *K-exponential models from the geometrical viewpoint*, propone que la función $\phi_k(x) = \cosh_k(x) - 1$ es equivalente a la función $|x|^{1/k}$, que genera el espacio de Lebesgue $L^{1/k}(p \cdot \mu)$ y cuya función complementaria es $|x|^{1-1/k}$ que genera al espacio $L^{1-1/k}(p \cdot \mu)$. Se propone así la construcción de la variedad mixta k -deformada para \mathcal{P} a través de estos espacios de Lebesgue.
3. Respecto al capítulo 5, mediante algún tipo de convergencia en redes, caracterizar la topología inducida por el atlas sobre \mathfrak{M}_μ .
4. Desarrollar un estudio del espacio tangente, según la variedad construida en el capítulo 5, en la misma vía que se ha hecho con la variedad de información de Pistone y Sempi [23].

Capítulo 7

Conclusiones

En el artículo titulado “*k-exponential models from the geometrical viewpoint*”, G. Pistone deja abierta la posibilidad de construir una nueva estructura que permita interpretaciones geométricas de aspectos relacionados con la información en el contexto del formalismo de la mecánica estadística según G. Kaniadakis. Precisamente propone una la construcción de una variedad que sea una k -deformación para la variedad de información estadística, considerando modelos que conectan a densidades positivas p y q de la forma $q = e_k^{u \overset{k}{\ominus} \Psi(u)} p$, donde: $\overset{k}{\ominus}$ es la k -diferencia según G. Kaniadakis y $\Psi(u)$ es alguna función de una variable aleatoria u .

En este trabajo se ha dado respuesta afirmativa al planteamiento de G. Pistone. Precisamente se establece que la función Ψ que hace posible la construcción de la variedad es definida por $\Psi(\cdot) = K_{p,k}(\cdot)$ acorde con la definición 5.6. De tal forma los modelos propuestos son del tipo

$$q = e_k^{u \overset{k}{\ominus} K_{p,k}(u)} p.$$

Propiamente, la variedad se construye como sigue. La colección de pares $\{(\mathcal{U}_{p,k}, s_{p,k})\}_{p \in \mathfrak{M}_\mu}$ es una atlas modelado sobre \mathfrak{M}_μ , donde los conjuntos $\mathcal{U}_{p,k}$ son abiertos en el espacio de Banach $B_{p,k} = \{L^{\phi_k}(p \cdot \mu) : E_p[u] = 0\}$ y $s_{p,k}$ es dada por la definición 5.8 y los mapeos de transición son como los indica la expresión (5.32); para así tener una variedad topológica sobre \mathfrak{M}_μ modelada por los espacios de Banach $L^{\phi_k}(p \cdot \mu)$.

Bibliografía

- [1] Amari, S. I.: *Differential-Geometric methods in Statistics*. Berlin, Springer-Verlag, 1985 (Lectures Notes in statistics, vol. 28)
- [2] Amari, S. I. and Nagaoka: *Methods of Information Geometry*. American Mathematics Society Providence, R.I., translated from Japanese by D. Harada, 1993
- [3] Blanco Castañeda, Liliana: *Probabilidad*. Universidad Nacional de Colombia, Bogotá, 2004.
- [4] Borja, Ruben: *Aspectos matemáticos en termoestadística generalizada*. Tesis de maestría, Universidad EAFIT. Medellín (2010).
- [5] Cena, Alberto: *Geometric Structures on the non-parametric statistical manifold*. Tesi di Dottorato. Dottorato in Matematica, Università di Milano. Eguchi, S. (2005)
- [6] Cena, Alberto and Pistone, Giovanni: *Exponential statistical manifold*. Annals of the Institute of Statistical Mathematics. ISSN 0020-3157 (Print) 1572-9052 (Online) 59, Pag 27-56 (2006).
- [7] Deossa, Dora: *Sobre funciones exponenciales y logarítmicas deformadas según Kaniadakis*. Tesis de maestría, Universidad EAFIT. Medellín (2011).
- [8] Gibilisco, P and Isola T: *Some open problems in information geometry*. 2006.
- [9] Graselli, M. R.: *Dual connections in nonparametric classical information geometry*. The Institute of Statistical Mathematics, Tokio, 2008.
- [10] Gregori Puerta, Pablo: *Espacios de medida vectoriales*. Tesis Doctoral, Universitat de Valencia, Departamento de análisis matematico, 2001.
- [11] Jeffreys, H.: *An invariant from for the prior probability in estimacion problems*. Proceedings of Royal Society A, Vol. 186, (1946), pag 453-461.
- [12] Jencová, Anna: *A construction of a nonparametric quantum information manifold*. Journal Of Functional Analysis, 239, 1-20, 2006.

- [13] Kadets, M. I. and Kadets, V. M.: *Series in Banach spaces Conditional and unconditional convergence*, Birkhauser Verlag, Basel, , Traslated for the Russian by Andrei Iacob. (1997).
- [14] G. Kaniadakis: *Statistical mechanics in the context of special relativity*. The American Physical Society, Physical Review E 66, 056125 1 (2002).
- [15] G. Kaniadakis: *Statistical mechanics in the context of special relativity II*. Print, 2008.
- [16] Kulback, S and Leibler, R.A. *On Information and Sufficiency*. Annals of Mathematics and Statistics. Vol 22. pag 79-86. (1951).
- [17] Lang, Serge.: *Differential Manifolds*. Springer Verlag. New York, 1985.
- [18] Loaiza, Gabriel y Quiceno, Héctor: *A q-exponential statistical Banach Manifold*. Journal of Mathematical Analysis and Applications, Vol. 398, (2013), pág 466-476.
- [19] Naudts, Jan: *Deformed exponentials and logarithms in generalized thermostatics*. Physica A. Statistical Mechanics and its Applications, Volumen 316, Issues 1-4, (2002) pp 323-334.
- [20] Naudts, Jan.: *Estimators, escort probabilities, and ϕ -exponential families in statical physics*. Journal of Inequalities in Pure and Applied Mathematics., Vol. 5., Article 102, (2004). pp., 1-15.
- [21] Pistone, Giovanni and Rogantin, Maria: *The exponential statistical manifold: mean parameters, orthogonality and space transformations*. Bernoulli. N° 4, pag 721-760 (1999)
- [22] Pistone, Giovanni and Sempi, Carlo: *An infinite-dimensional geometric structure on the space of all the probability measures equivalent to a given one*. The Annals of Statistics., Vol. 23., No. 5., (1995). pp.2, 1543–1561.
- [23] Pistone, Giovanni: *K-exponential models from the geometrical viewpoint*. The European Physical Journal B. Springer Berlin. ISSN 1434-6028 (Print) 1434-6036 (Online) 70, Pag 29-37 (2009)
- [24] Quiceno Echavarría, Héctor: *Variedad de información estadística q-exponencial*. Tesis de maestría, Universidad Eafit. Medellín (2010).
- [25] Rao, C. R. : *Information and accuracy attainable in the estimation of statistical parameters*. Bull. Calcutta Math. Soc. 37 (1945) 81-91
- [26] Rao, M. M. and Ren, Z.D.: *Application of Orlicz spaces*. Marcel Dekker Inc. 2002.
- [27] Rao, M. M. and Ren, Z.D.: *Theory of Orlicz Spaces*. Marcel Dekker Inc. New York, N.Y. 1991.

-
- [28] Royden, Halsey: *Real Analysis*. Segunda edición. The Macmillan Company, 1968.
- [29] Streater, R. F. : *Quantum Orlicz spaces in information geometry*. Open Systems & Information Dynamics. Vol 11. N0. 4 (2002), 359-375.
- [30] Zhang, Jun and Hästö, Peter. *Statistical manifold as an affine space: A functional equation approach*. Journal of Mathematical Psychology 50, 2006.
- [31] Young. W.H. *On classes of summable functions and their Fourier series*. Proc. Roy. Soc. 87 (1912), 225-229.