

SOBRE ESTRUCTURAS GEOMÉTRICAS PARA MODELOS q -EXPONENCIALES

Marinela Andrea Restrepo Zuleta
Leonardo Fabio Carvajal Fernández

ESCUELA DE CIENCIAS Y HUMANIDADES

DEPARTAMENTO DE CIENCIAS BÁSICAS

MAESTRÍA EN MATEMÁTICAS APLICADAS

MEDELLÍN

Mayo 2014

SOBRE ESTRUCTURAS GEOMÉTRICAS PARA MODELOS q -EXPONENCIALES

**Trabajo de investigación presentado como requisito para optar al título de
Magíster en Matemáticas Aplicadas**

**Marinela Andrea Restrepo Zuleta
Leonardo Fabio Carvajal Fernández**

Director

Gabriel Ignacio Loaiza Ossa

Doctor en Ciencias Matemáticas

Escuela de Ciencias y Humanidades

Departamento de Ciencias Básicas

Maestría en Matemáticas Aplicadas

Medellín

Mayo 2014

Dedicatoria

Este trabajo de grado está dedicado a la memoria de **Gerardo Arango**, gran mentor.

Agradecimientos

En este espacio queremos expresar nuestra gratitud a las personas que con su amistad, apoyo y compañía han contribuido a este logro, que nos ha dejado experiencias de incalculable valor en lo cognoscitivo, afectivo, moral y social, permitiéndonos progresar de manera integral como personas y profesionales en un mundo caracterizado por la globalización.

En nuestros agradecimientos, nos dirigimos a Dios, pues, todo lo hace posible, y nos concedió el privilegio de tener como director al Dr. Gabriel Ignacio Loaiza, a quien le debemos muchísimo, particularmente su disposición para orientarnos. A nuestros familiares y amigos, quienes constantemente con su confianza nos motivaron para alcanzar la meta, sacrificando mucho de sí por nuestro bienestar.

También nuestro merecido reconocimiento para nuestros profesores, quienes han dejado huellas de gran estima en cada uno de nosotros, bajo la coordinación de Francisco Zuluaga Díaz y Jairo Villegas Gutiérrez. Así mismo, a los docentes de la Unicórdoba, Hugo Aduén Muskus y Abraham Arenas Tawil, a Hugo Arbeláez, docente de la Universidad Nacional de Medellín, por su incondicional apoyo.

Finalmente queremos manifestar nuestra gratitud a Telvia Rosa Castilla, Mónica Arteaga Sierra, José Guerra Bonet y Juan Pablo Ballen, nuestro equipo de estudio.

Dios los bendiga a todos.

Estimada por la sociedad.

Avanza exitosamente.

Fénix de mi esperanza.

Icono de acreditación.

Te contemplo de gratitud.

Resumen

Las estructuras geométricas para las variedades de información que describen modelos exponenciales, tanto en el caso paramétrico como no paramétrico, han sido estudiadas desde el siglo pasado. Respecto a las variedades que describen modelos q -exponenciales, acordes con el índice de entropía en la mecánica estadística de Tsallis, el desarrollo apenas comienza y se discutirá en la presente memoria.

Índice general

Introducción	1
1. Preliminares	5
1.1. Espacios de funciones de Orlicz	5
1.2. Analiticidad	8
1.3. Variedades de Banach	9
1.4. Funciones q -exponenciales y q -logarítmicas	10
2. Variedades estándar de información estadística	25
2.1. Variedad estándar de Amari (caso paramétrico)	25
2.2. Variedad de Pistone y Sempi(caso no paramétrico)	26
3. Variedad de información q-exponencial: caso paramétrico	31
3.1. Familia q -exponencial o q -Gibbs	32
3.2. Estructura Riemanniana y q -divergencia	39
3.3. Estructura dualmente plana de la variedad	41
4. Variedad de información q-exponencial: caso no paramétrico	49
4.1. Funcional acumulante	50
4.2. Definición de la variedad	55
4.3. Geometría inducida por q -Divergencias	62
Bibliografía	73

Introducción

La matemática como ciencia que establece patrones y permite formular nuevas conjeturas tiende a generalizar resultados clásicos para poder describir mediante algunos modelos ciertos fenómenos no estándares. En ese sentido, con el paso del tiempo, el concepto de variedad como una generalización del concepto métrico, diferencial y topológico del espacio euclidiano ha dado origen a un gran desarrollo matemático que ha permitido integrar diferentes áreas de esta disciplina. El primer aporte lo realiza Gauss mediante un estudio intrínseco de las superficies, seguido por Riemann, que aporta una estructura métrica para las variedades [17, 18]. En ese recorrido, la asociación de los conceptos de espacios de Orlicz [6, 7], geometría riemannianna, estadística y la teoría de la información (teoría de las variedades de información estadística) han contribuido en el avance de esta línea de investigación.

Así como la ecuación fundamental de Newton constituye el pilar de la dinámica clásica, la distribución de Gibbs lo es para la física estadística, su cumbre, pues, permite derivar importantes relaciones que son objeto de muchos tratados. En este sentido, la distribución de Gibbs se encuentra ligada a la estructura geométrica diferencial de los modelos estadísticos, en relación con su estructura riemanniana, que ha sido objeto de estudio en estos dos últimos siglos. Estas herramientas geométricas se han aplicado en el estudio de la teoría de información generando nuevos desarrollos matemáticos en la teoría de información geométrica. Además, G. Pistone [35] presenta una discusión sobre el uso del modelo k -exponencial de Kaniadakis en la construcción de una variedad estadística modelada sobre los espacios de Lebesgue de variables aleatorias reales. También se consideran algunos aspectos algebraicos de los modelos exponenciales deformados.

El estudio de estructuras geométricas de modelos estadísticos, especialmente las estructuras riemannianas, se han desarrollado de dos formas: el finito y el infinito dimensional, que corresponden a los modelos estadísticos paramétricos y no paramétricos respectivamente. En esas estructuras, todo punto de la variedad es una función de densidad de probabilidad en algún espacio medible. Para el caso paramétrico, con los trabajos de Rao [11] y Jeffreys [12], la información de Fisher es dada como una métrica para un modelo estadístico paramétrico, mediante las conexiones de Levi-Civita. Efron [13], introdujo el concepto de curvatura usando implícitamente una conexión exponencial, que fué desarrollada por Dawid [14]. El estudio del caso finitodimensional culmina con Amari [20], quien define una familia uniparamétrica de α -conexiones, que particularizan a la conexión exponencial cuando α tiende a uno. Respecto al caso de dimensión infinita, el trabajo fué

desarrollado por G. Pistone y C. Sempì (1995), construyendo una variedad de Banach para la familia exponencial con espacios de Orlicz como espacios coordinados y definida sobre las funciones de densidad de probabilidad no parametrizadas [26]. Gibilisco y Pistone [19] definieron la conexión exponencial como una conexión natural inducida por el uso de espacios de Orlicz y probaron que las conexiones exponencial y mixtas tienen una relación de dualidad, justo como en el caso paramétrico.

Por otra parte algunos modelos, llamados modelos complejos, no encajan en las representaciones anteriormente mencionadas y se hace necesario definir nuevas familias de distribuciones para contener dichos modelos. Una de estas familias es la q -exponencial, basada en una deformación de la exponencial usando la función q -exponencial que se ha utilizado en varias aplicaciones de la mecánica estadística no extensiva según Tsallis [15, 16], donde el número real q es llamado índice de entropía.

Con el fin de dar una estructura geométrica para estos modelos, Amari y Ohara [1] estudiaron la geometría de la familia q -exponencial en el entorno de dimensión finita y encontraron que esta familia tiene una estructura geométrica doblemente plana. En 2013, Loaiza y Quiceno [2] han construido para esta familia, una variedad estadística no paramétrica inspirados en espacios de funciones esencialmente acotadas, de manera que cada modelo paramétrico q -exponencial se identifica con el espacio tangente y los mapas de coordenadas son definidos de forma natural en términos de entropías relativas en el contexto de Tsallis y en [3] se caracteriza la estructura riemanniana. La variedad construida recupera la variedad estándar de Pistone y Sempì (salvo inmersiones continuas en el espacio de modelado) cuando q tiende a uno.

Los resultados que conciernen a la familia q -exponencial son recientes, aparecen en [1, 2, 3] y no aparecen aún en otros textos. La presente tesis aborda la discusión sobre la estructura geométrica para los modelos q -exponenciales en los casos paramétrico y no paramétrico, exponiendo dichos resultados con detalles en la ampliación de las pruebas y características geométricas. Más que una compilación de resultados sobre este tema específico, la expectativa es tener un material que contribuya a un estudio más eficiente para quienes aborden investigaciones con modelos q -exponenciales y, particularmente, apoyar las investigaciones del grupo de Análisis Funcional y Aplicaciones de la Universidad Eafit.

El orden a seguir será: El primer capítulo corresponde a preliminares matemáticos. El segundo capítulo describe las variedades estándar de información estadística. En el tercer y cuarto capítulos se discute sobre las estructuras geométricas de la familia q -exponencial en el caso paramétrico y en el caso no paramétrico, respectivamente. Por último se dan las conclusiones y algunos problemas abiertos.

Objetivos

Objetivo general

Presentar una discusión sobre la estructura riemanniana de las variedades de información estadísticas asociadas a modelos q -exponenciales, tanto en el caso paramétrico como no paramétrico.

Objetivos específicos

1. Presentar la construcción de variedades de dimensión tanto finita como infinita, en términos de la mecánica estadística de C. Tsallis.
2. Discutir las características de las geometrías riemannianas en cada variedad q -exponencial.
3. Mostrar el desarrollo sobre las estructuras geométricas para variedades de información estadística q -exponencial.

Capítulo 1

Preliminares

1.1. Espacios de funciones de Orlicz

El conocimiento sobre varios aspectos de la teoría de espacios de Orlicz es fundamental para entender mejor el trabajo de Pistone y Sempi, sobre la variedad q -exponencial para el caso no paramétrico; de hecho, los espacios coordinados para la variedad son espacios de funciones de Orlicz. A pesar de que dichos espacios son una generalización natural de los espacios clásicos L^p , no son muy conocidos en la comunidad matemática y es por ello que a continuación presentamos algunos conceptos, propiedades y resultados importantes sobre teoría de espacios de Orlicz, tomados de los textos de Rao y Ren [6, 7]. Para tener un contexto claro, iniciamos con aspectos de la teoría de la medida y probabilidad que pueden ser consultados en [8, 9].

Definición 1.1.1: Una σ -álgebra en Ω es una colección \mathcal{A} de subconjuntos sobre Ω que satisface las siguientes propiedades:

- a) $\Omega \in \mathcal{A}$.
- b) Si $A \in \mathcal{A}$ entonces $A^c \in \mathcal{A}$.
- c) Si $A_1, A_2, \dots, A_n, \dots \in \mathcal{A}$ entonces $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$.

\mathcal{A} es un álgebra, si la tercera condición se cumple sólo para uniones finitas.

Definición 1.1.2: El par (Ω, \mathcal{A}) será llamado *espacio medible*, donde Ω es un conjunto y \mathcal{A} es una σ -álgebra sobre Ω y *conjuntos medibles* a los elementos de \mathcal{A} .

Definición 1.1.3: Dada una familia de eventos \mathcal{A} de un espacio muestral Ω , es decir, una σ -álgebra de subconjuntos de Ω . Asignaremos a cada evento $A \in \Omega$ un número real $P(A)$, llamado la probabilidad de A , con las siguientes propiedades:

1. $P(A) \geq 0$ para todo $A \in \Omega$.
2. $P(\Omega) = 1$.
3. Si $A_n \in \mathcal{A}$ para $n = 1, 2, \dots$ son eventos disjuntos dos a dos, entonces $P(\bigcup_{n=1}^{\infty} A_n) =$

$$\sum_{n=1}^{\infty} P(A_n).$$

Una terna (Ω, \mathcal{A}, P) , donde Ω es un espacio muestral, \mathcal{A} una familia de eventos y P una probabilidad es llamada espacio de probabilidad.

Definición 1.1.4: Dado $(\Omega, \mathcal{A}, \mu)$ un espacio de probabilidad y p una función de densidad de probabilidad. Denotamos por $p \cdot \mu$, la medida de probabilidad con densidad p , dada por

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

para cada $A \in \mathcal{A}$.

Es conocido el siguiente resultado: fp es μ medible, si y sólo si, f es $p \cdot \mu$ medible, en tal caso

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

Definición 1.1.5: Una función $\Phi : \mathfrak{R} \rightarrow \mathfrak{R}^+$ es una función de Young si:

- i. Φ es convexa.
- ii. $\Phi(0) = 0$.
- iii. $\Phi(x) = \Phi(-x)$.
- iv. $\lim_{x \rightarrow \infty} \Phi(x) = \infty$.

Algunas funciones de Young son:

- i. $\Phi_1(x) = \cosh(x) - 1$.
- ii. $\Phi_2(x) = |x|^p$, con $p \geq 1$.

Dos funciones de Young son equivalentes, en el siguiente sentido: $\Phi = \Psi$ si existen dos constantes, $0 < c_1 < c_2 < \infty$ y un $x_0 \geq 0$ tales que

$$\Phi(c_1(x)) \leq \Psi(x) \leq \Phi(c_2(x)) \text{ para } x \geq x_0.$$

Una función de Young Φ se dice que cumple la condición Δ_2 , si existen $s_0 \geq 0$ y $k > 0$ tal que $\Phi(2s) \leq k\Phi(s)$ para $s_0 \leq s < \infty$.

Definición 1.1.6: Sea Φ una función de Young. La clase de Orlicz \mathcal{L}^Φ consiste de todas las funciones $u \in L^1(p \cdot \mu)$ para las cuales el funcional

$$\int_{\Omega} \Phi(u) p d\mu$$

es finito.

De la definición se puede notar que $\mathcal{L}^\Phi(p \cdot \mu) \subset L^1(p \cdot \mu)$.

Teorema 1.1.1: Sea Φ una función de Young entonces $\mathcal{L}^\Phi(p \cdot \mu)$ es un conjunto convexo.

Prueba: Sean u y v en $\mathcal{L}^\Phi(p \cdot \mu)$. Por lo tanto, $\int_\Omega \Phi(u)pd\mu < \infty$ y $\int_\Omega \Phi(v)pd\mu < \infty$. Como Φ es convexa, para $0 \leq \lambda \leq 1$, $\Phi((1-\lambda)u + \lambda v) \leq (1-\lambda)\Phi(u) + \lambda\Phi(v)$. De donde,

$$\begin{aligned} \int_\Omega \Phi((1-\lambda)u + \lambda v)pd\mu &\leq \int_\Omega [(1-\lambda)\Phi(u) + \lambda\Phi(v)]pd\mu \\ &\leq (1-\lambda) \int_\Omega \Phi(u)pd\mu + \lambda \int_\Omega \Phi(v)pd\mu < \infty. \end{aligned}$$

En consecuencia, $(1-\lambda)u + \lambda v \in \mathcal{L}^\Phi(p \cdot \mu)$, lo que prueba la convexidad de este conjunto. \square

Definición 1.1.7: Sea Φ una función de Young. El espacios de Orlicz $L^\Phi(p \cdot \mu)$ asociado a Φ , está dado por

$$L^\Phi(p \cdot \mu) = \{u \in L^1(p \cdot \mu) : \int_\Omega \Phi(\alpha u)pd\mu < \infty \text{ para alguna } \alpha > 0\}$$

El siguiente resultado muestra que el espacio de Orlicz constituye un espacio lineal convexo.

Teorema 1.1.2: Si Φ una función de Young, $L^\Phi(p \cdot \mu)$ es un espacio lineal convexo.

Definición 1.1.8: Sea $(\Omega, \mathcal{A}, \mu)$ un espacio de probabilidad. Denotamos por \mathfrak{M}_μ el conjunto

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

Sobre este conjunto G. Pistone y C. Sempì construyen una variedad de información modelada sobre espacios de Banach para los cuales cada $p \in \mathfrak{M}_\mu$ constituye un subespacio cerrado del correspondiente espacio de Orlicz $L^{\Phi_1}(p \cdot \mu)$.

Teorema 1.1.3: \mathfrak{M}_μ es un conjunto convexo.

Prueba: Sean $p_1, p_2 \in \mathfrak{M}_\mu$. Luego, por definición, se tiene que $\int_\Omega p_1d\mu = 1$ y $\int_\Omega p_2d\mu = 1$, con $p_1 > 0, p_2 > 0$ en casi toda parte.

Si $0 \leq \lambda \leq 1$, $\lambda \in \mathfrak{R}$, tenemos que

$$\int_\Omega [\lambda p_1 + (1-\lambda)p_2]d\mu = \lambda \int_\Omega p_1d\mu + (1-\lambda) \int_\Omega p_2d\mu = \lambda + (1-\lambda) = 1$$

es decir, $\lambda p_1 + (1-\lambda)p_2 \in \mathfrak{M}_\mu$. En consecuencia, \mathfrak{M}_μ es un conjunto convexo. \square

Es claro que \mathfrak{M}_μ no constituye un espacio lineal, pués, no es cerrado bajo la suma.

Definición 1.1.9: Sea $u \in L^\Phi(p \cdot \mu)$ una variable aleatoria. El mapeo $\hat{u}_p(\cdot) : \mathfrak{R} \rightarrow \bar{\mathfrak{R}}$ dado por

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

es conocido como la función generadora de momentos de la variable u .

Asociada a la función generadora de momentos tenemos la clase de Cramer C_p , donde

$$C_p = \{u \in L^\Phi(p \cdot \mu) : \hat{u}_p(t) < \infty, t \in I, \text{ con } I \text{ en una vecindad de cero}\}.$$

Teorema 1.1.4: Sea $p \in \mathfrak{M}_\mu$. La clase de Cramer C_p es un espacio lineal convexo.

Definición 1.1.10: Dos densidades de probabilidad $p, z \in \mathfrak{M}_\mu$ están conectadas por un modelo q -exponencial unidimensional si existen $r \in \mathfrak{M}_\mu$, $u \in L^\infty(p \cdot \mu)$, una función real de variable real ψ y $\delta > 0$ tal que para todo $t \in (-\delta, \delta)$, la función f definida por

$$f(t) = e_q^{tu \ominus_q \psi(t)} r,$$

satisface la existencia de $t_0, t_1 \in (-\delta, \delta)$ para la cual $p = f(t_0)$ y $z = f(t_1)$. La función f es llamada modelo q -exponencial unidimensional.

Lema 1.1.1: Sean $p, z \in \mathfrak{M}_\mu$ densidades de probabilidad relacionadas por un modelo q -exponencial unidimensional. Entonces $L^\infty(p \cdot \mu) = L^\infty(z \cdot \mu)$ y la transformación identidad, $id : L^\infty(p \cdot \mu) \rightarrow L^\infty(z \cdot \mu)$ es una isometría. En consecuencia, se cumple $\frac{-1}{1-q} < \ln_q(\frac{p}{z})$.

1.2. Analiticidad

Requerimos de la propiedad de analiticidad, para demostrar posteriormente que los mapeos son diferenciables.

Denotamos por $\mathcal{L}^n(E, F)$ los mapeos multilineales continuos de E^n a F , donde E y F son espacios de Banach.

Un operador multilineal λ es acotado si existe $M > 0$ tal que

$$\|\lambda(u_1, u_2, \dots, u_n)\|_F \leq M \|u_1\|_E \dots \|u_n\|_E, \text{ para todo } u_1, u_2, \dots, u_n \in E^n.$$

Un operador se dice simétrico si el intercambio del orden no altera el resultado de la aplicación.

Definición 1.2.1 Sean E y F espacios de Banach. Un polinomio n -homogéneo continuo es una función $\hat{\lambda} : E \rightarrow F$, si existe un operador lineal, continuo y simétrico λ tal que

$$\hat{\lambda}(x) = \lambda x^n, \text{ para todo } x \in E.$$

Definición 1.2.2: Sean E y F espacios de Banach. Un mapeo analítico es una función $f : U \rightarrow F$, con $U \subset E$ abierto, para la cual existe una serie de potencias convergente $\sum_0^\infty \hat{\lambda}_n$ con radio de convergencia positivo tal que, para todo $x_0 \in U$ y todo x en una vecindad de x_0

$$f(x) = \sum_0^\infty \hat{\lambda}_n(x - x_0).$$

Una función analítica es infinitamente diferenciable sobre la vecindad de x_0 y la k -ésima derivada en series de potencias viene dada por

$$D^k f(x) = \sum_0^\infty \frac{(k+n)!}{n!} \lambda_{k+n}(x - x_0)^n.$$

Teorema 1.2.1: Sean E , F y G espacios de Banach tales que $U \subset E$ y $V \subset F$ abiertos. Si $f : U \rightarrow F$ y $g : V \rightarrow G$ son funciones analíticas tales que $f(U) \subset V$ entonces la composición de funciones $g \circ f$ es también analítica.

1.3. Variedades de Banach

Definición 1.3.1: Una variedad topológica X modelada por un espacio de Banach E es un espacio topológico X dotado de una colección de abiertos U que lo recubren y una colección de mapas (U, ϑ) llamados cartas que consisten de un homeomorfismo $\vartheta : U \rightarrow \vartheta(U) \subset E$, con $\vartheta(U)$ abierto en E , con la condición de compatibilidad: si (V, ϱ) es cualquier otra carta de X , entonces la función de transición $\varrho \circ \vartheta^{-1} : \vartheta(U \cap V) \subset E \rightarrow E$ es continua. Un atlas es una colección de cartas compatibles que cubre todo X .

Se dice que una colección de cartas recubren a X , si para todo punto de X siempre hay un entorno de coordenadas de una carta en el que está contenido.

Definición 1.3.2: Sea $(*)$ alguna de las siguientes categorías:

1. Diferenciable
2. C^k
3. C^∞
4. C^w o analítica.

Diremos que la variedad X es de clase $(*)$ si las funciones de transición son $(*)$. Si Y es otra variedad modelada por un espacio de Banach F , diremos que una función $f : X \rightarrow Y$ entre variedades es $(*)$ si para todo par de cartas $(U, \vartheta), (U', \xi)$ de X y Y respectivamente, la función $\xi \circ f \circ \vartheta^{-1} : E \rightarrow F$ es $(*)$ en el abierto $f(U) \subset E$.

Definición 1.3.3: Un vector tangente a X en p es definido como la clase de equivalencia de sus representantes dados por cartas que cubren p .

Definición 1.3.4: El espacio tangente T_p en el punto p de una variedad X es, en términos generales, un espacio vectorial obtenido por linealización local de X alrededor de p .

Definición 1.3.5: El fibrado tangente es un espacio formado por todos los planos tangentes a la variedad en todos los puntos.

1.4. Funciones q -exponenciales y q -logarítmicas

El carácter imponente de la distribución de Boltzmann-Gibbs en la física estadística es reconocido a nivel general, sin embargo, mediante la implementación de funciones exponenciales y logarítmicas q -deformadas obtenemos modelos más generales. Estas funciones constituyen el fundamento de la mecánica estadística no-extensiva. Un gran mentor de este trascendental aporte es C. Tsallis, quien definió un funcional de entropía empleando un parámetro real q , originando nuevas estructuras y desarrollos matemáticos, de los cuales exponemos algunos en esta sección tomados de [16].

Para $0 < q < 1$, las funciones señaladas se definen respectivamente por

$$e_q^x = (1 + (1 - q)x)^{1/(1-q)}, \text{ si } \frac{-1}{1-q} \leq x \text{ y } \ln_q(x) = \frac{x^{1-q} - 1}{1-q}, \text{ si } x > 0.$$

Es fácil verificar que una función es la inversa de la otra:

$$\ln_q(e_q^x) = e_q^{\ln_q x} = x.$$

Las funciones q -exponencial y q -logaritmo surgen como soluciones de ecuaciones diferenciales de variables separables. Precisamente, al considerar la ecuación $y' = x^{-q}$, $y(1) = 0$, obtenemos como solución la función q -logaritmo, $\ln_q(x) = \frac{x^{1-q} - 1}{1-q}$, si $x > 0$.

Para verlo, resolvamos la ecuación $y' = x^{-q}$ sujeta a la condición $y(1) = 0$.

$$\frac{dy}{dx} = x^{-q} \text{ entonces } dy = x^{-q} dx$$

integrando a ambos lados tenemos,

$$\int dy = \int x^{-q} dx$$

$$y = \frac{x^{1-q}}{1-q} + c$$

Al aplicar la condición $y(1) = 0$, se tiene que $c = 0$ y la solución de la ecuación diferencial es:

$$y = \frac{x^{1-q} - 1}{1-q} = \ln_q(x).$$

Consideremos la ecuación $y' = y^q$ con la condición de contorno $y(0) = 1$.

Resolvamos la ecuación $y' = y^q$ con $y(0) = 1$.

$$\frac{dy}{dx} = y^q \quad \text{entonces} \quad y^{-q} dy = dx$$

integrando a ambos lados tenemos,

$$\int y^{-q} dy = \int dx$$

$$\frac{y^{1-q}}{1-q} = x + k$$

$$y^{1-q} = (1-q)x + (1-q)k$$

$$y^{1-q} = (1-q)x + c, \quad \text{con } c = (1-q)k$$

Luego, $y = [c + (1-q)x]^{\frac{1}{1-q}}$.

Puesto que $y(0) = 1$, se tiene que $c = 1$. Así, $y(x) = [1 + (1-q)x]^{1/(1-q)} = e_q^x$.

De otra parte, si hacemos $f(q) = \frac{x^{1-q}-1}{1-q}$ y consideramos el límite cuando $q \rightarrow 1$, por L'Hopital, obtenemos

$$\lim_{q \rightarrow 1} f(q) = \lim_{q \rightarrow 1} \frac{-x^{1-q} \log(x)}{-1} = \log(x) = \log_1(x).$$

Se puede notar que q -logaritmo es una función creciente, pues,

$$\frac{d}{dx} \log_q(x) = \frac{1}{x^q} > 0.$$

por tanto, tiene inversa.

Llamamos q -exponencial, la inversa de la función q -logaritmo, y está definida por

$$\exp_q(x) = \{1 + (1 - q)x\}^{\frac{1}{1-q}}, \quad x > \frac{-1}{1-q}, \quad q > 0.$$

con $q \neq 1$. El caso límite $q \rightarrow 1$ se reduce a $\exp_1(x) = \exp(x)$. En efecto, sea $g(q) = \{1 + (1 - q)x\}^{\frac{1}{1-q}}$, tenemos que

$$y = \lim_{q \rightarrow 1} \{1 + (1 - q)x\}^{\frac{1}{1-q}}.$$

Tomando logaritmo en ambos lados, y por la regla de L'Hopital, se nota que

$$\ln(y) = \lim_{q \rightarrow 1} \frac{1}{1-q} \ln\{1 + (1 - q)x\} = \lim_{q \rightarrow 1} \frac{x}{1+(1-q)x} = x.$$

$$\ln(y) = x$$

$$y = \exp(x)$$

Así,

$$\exp_1(x) = \exp(x).$$

q -Álgebra.

Asociadas a las funciones exponenciales y logarítmicas q -deformadas definimos las siguientes operaciones para números reales x e y por

1. La q -suma $\oplus_q : \mathbb{R}^2 \rightarrow \mathbb{R}$ está dada por

$$x \oplus_q y = x + y + (1 - q)xy.$$

2. La q -diferencia.

$$x \ominus_q y := \frac{x-y}{1+(1-q)y}, \quad \text{para } y \neq (q-1)^{-1}.$$

3. El q -producto $\otimes_q : \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow \mathbb{R}$, es dado por

$$x \otimes_q y = [x^{1-q} + y^{1-q} - 1]_+^{\frac{1}{1-q}}, \quad x > 0 \text{ y } y > 0.$$

Proposición 1.4.1: Las operaciones (1) y (3) son asociativas, modulativas y conmutativas, siendo 0 y 1 los módulos respectivos.

Prueba:

I. $x \oplus_q y = x + y + (1 - q)xy$.

Asociativa:

$$\begin{aligned}
(x \oplus_q y) \oplus_q z &= (x \oplus_q y) + z + (1 - q)(x \oplus_q y)z \\
&= (x + y + (1 - q)xy) + z + (1 - q)(x + y + (1 - q)xy)z \\
&= x + y + z + (1 - q)xy + (1 - q)xz + (1 - q)yz + (1 - q)^2xyz \\
&= x + y + z + (1 - q)yz + (1 - q)xy + (1 - q)xz + (1 - q)^2xyz \\
&= x + (y + z + (1 - q)yz) + (1 - q)x(y + z + (1 - q)yz) \\
&= x + (y \oplus_q z) + (1 - q)x(y \oplus_q z) \\
&= x \oplus_q (y \oplus_q z).
\end{aligned}$$

Modulativa: El 0 es el modulo de la operación \oplus_q

$$x \oplus_q 0 = x + 0 + (1 - q)x0 = x.$$

$$0 \oplus_q x = 0 + x + (1 - q)0x = x.$$

Conmutativa:

$$x \oplus_q y = x + y + (1 - q)xy = y + x + (1 - q)yx = y \oplus_q x.$$

II. $x \otimes_q y = [x^{1-q} + y^{1-q} - 1]_+^{\frac{1}{1-q}}$, $x > 0$ y $y > 0$.

Asociativa:

$$\begin{aligned}
(x \otimes_q y) \otimes_q z &= [(x \otimes_q y)^{1-q} + z^{1-q} - 1]_+^{\frac{1}{1-q}} \\
&= [(x^{1-q} + y^{1-q} - 1) + z^{1-q} - 1]_+^{\frac{1}{1-q}} \\
&= [(x^{1-q} - 1) + (y^{1-q} + z^{1-q} - 1)]_+^{\frac{1}{1-q}} \\
&= [(x^{1-q} - 1) + (y \otimes_q z)^{1-q}]_+^{\frac{1}{1-q}} \\
&= [x^{1-q} + (y \otimes_q z)^{1-q} - 1]_+^{\frac{1}{1-q}} \\
&= x \otimes_q (y \otimes_q z).
\end{aligned}$$

Modulativa: El 1 es el modulo de la operación \otimes_q

$$x \otimes_q 1 = [x^{1-q} + 1 - 1]_+^{\frac{1}{1-q}} = x.$$

$$1 \otimes_q x = [1 + x^{1-q} - 1]_+^{\frac{1}{1-q}} = x.$$

Conmutativa:

$$x \otimes_q y = [x^{1-q} + y^{1-q} - 1]_+^{\frac{1}{1-q}} = [y^{1-q} + x^{1-q} - 1]_+^{\frac{1}{1-q}} = y \otimes_q x.$$

□

Proposición 1.4.2: Sea $x \neq \frac{1}{q-1}$, entonces existe un único inverso b , para \oplus_q ; denotado por $b = \ominus_q x$ y dado por $b = -\frac{x}{1+(1-q)x}$.

Entonces

$$x \oplus_q (\ominus_q x) = 0.$$

Prueba:

Existencia.

Sea $x \neq \frac{1}{q-1}$, luego

$$\begin{aligned} x \oplus_q b &= x + b + (1 - q)xb \\ &= x - \frac{x}{1 + (1 - q)x} - \frac{(1 - q)x^2}{1 + (1 - q)x} \\ &= \frac{x + (1 - q)x^2 - x - (1 - q)x^2}{1 + (1 - q)x} = 0. \end{aligned}$$

Unicidad.

Supongamos que existen b_1 y b_2 tales que $x \oplus_q b_1 = 0$ y $x \oplus_q b_2 = 0$. Veamos que $b_1 = b_2$. En efecto,

$$b_1 = b_1 \oplus_q 0 = b_1 \oplus_q (x \oplus_q b_2) = (b_1 \oplus_q x) \oplus_q b_2 = (x \oplus_q b_1) \oplus_q b_2 = 0 \oplus_q b_2 = b_2. \square$$

Proposición 1.4.3: Para $x > 0$ existe una inversa única para \otimes_q , denotada \oslash_q y dada por $b = (2 - x^{1-q})_+^{\frac{1}{1-q}}$.

En efecto, sea $x > 0$, luego para $b = (2 - x^{1-q})_+^{\frac{1}{1-q}}$, se tiene que

$$\begin{aligned} x \otimes_q b &= [x^{1-q} + b^{1-q} - 1]_+^{\frac{1}{1-q}} \\ &= [x^{1-q} + (2 - x^{1-q})_+^{\frac{1}{1-q}} - 1]_+^{\frac{1}{1-q}} \\ &= [x^{1-q} + 2 - x^{1-q} - 1]_+^{\frac{1}{1-q}} \\ &= 1. \end{aligned}$$

Con la proposición anterior, se define una q -cociente para reales x, y , positivos, de tal manera que $x^{1-q} \leq y^{1-q} + 1$, dada por

$$x \oslash_q y = x \otimes_q (\oslash_q y) = (x^{1-q} + (2 - y^{1-q})_+ - 1)_+^{\frac{1}{1-q}} = (x^{1-q} - y^{1-q} + 1)_+^{\frac{1}{1-q}}$$

Note que para $0 < x \leq 2$ se sigue que $1 \oslash_q (1 \oslash_q x) = x$; y si $q < 1$, la expresión $1 \oslash_q 0$ es convergente.

$$x \oslash_q y = 1 \oslash_q (y \oslash_q x) \text{ si } x^{1-q} \leq y^{1-q} + 1$$

$$x \oslash_q (y \oslash_q z) = (x \oslash_q y) \otimes_q z = (x \otimes_q z) \oslash_q y$$

si $z^{1-q} - 1 \leq y^{1-q} \leq x^{1-q} + 1$.

Proposición 1.4.4:

1. Para $q < 0$, $x \in [0, \infty)$, $exp_q(x)$ es positiva, continua, creciente, cóncava y de tal manera que

$$\lim_{x \rightarrow \infty} exp_q(x) = \infty.$$

2. Para $0 < q < 1$, $x \in [0, \infty)$, $exp_q(x)$ es positiva, continua, creciente, convexa y de tal manera que

$$\lim_{x \rightarrow \infty} exp_q(x) = \infty.$$

3. Para $1 < q$, $x \in [0, \frac{1}{q-1})$, $exp_q(x)$ es positiva, continua, creciente, convexa y de tal manera que

$$\lim_{x \rightarrow (\frac{1}{q-1})^-} exp_q(x) = \infty.$$

Prueba:

1. Por definición de la función $exp_q(x)$ se tiene que $\frac{-1}{1-q} \leq x$, así que $-1 \leq (1-q)x$. Luego, $0 \leq 1 + (1-q)x$, de donde $0 < exp_q(x)$. Así, $exp_q(x)$ es positiva.

De otra parte, para $q < 0$ se tiene que $0 < \frac{1}{1-q} < 1$ y $-1 < \frac{-1}{1-q}$. También es claro que $f(x) = 1 + (1-q)x \geq 0$ es continua en $a \in [0, +\infty)$. Así mismo, $g(x) = x^{\frac{1}{1-q}}$ es continua en $f(a) \geq 0$. Luego, por el teorema de la composición de dos funciones, se tiene que $exp_q(x) = g(f(x))$ es continua en a .

Veamos que $exp_q(x)$ es creciente. En efecto, sean x_1, x_2 en el dominio de $exp_q(x)$ tales que $x_1 < x_2$. Luego,

$$1 + (1-q)x_1 < 1 + (1-q)x_2 \text{ implica } [1 + (1-q)x_1]^{\frac{1}{1-q}} < [1 + (1-q)x_2]^{\frac{1}{1-q}},$$

es decir, $exp_q(x_1) < exp_q(x_2)$. Así, $exp_q(x)$ es creciente.

Otra forma era notar que $[exp_q(x)]' = (exp_q(x))^q > 0$, pues, $exp_q(x)$ es positiva.

Es fácil observar que $exp_q(x)$ es cóncava, pues, $exp_q(x) \leq 1+x$. En particular, la desigualdad se obtiene de la de Bernoulli, $(1+x)^n \geq 1+nx$, tomando $n = (1-q) \in \mathbb{Z}^+$.

Veamos que $\lim_{x \rightarrow \infty} exp_q(x) = \infty$. Necesitamos mostrar que, dado $A > 0$, existe $k > 0$ tal que para toda $x > k$, se tiene que $exp_q(x) > A$. En efecto, dado $A > 0$, por la propiedad Arquimediana, existe $k > 0$ tal que $k > \frac{2A^{1-q}}{1-q}$. Luego, para $x > k$ se tiene que

$$x > \frac{2A^{1-q}}{1-q} \quad \text{y} \quad x \geq \frac{-1}{1-q} > \frac{-2}{1-q},$$

de donde, $2x > \frac{-2}{1-q} + \frac{2A^{1-q}}{1-q}$, entonces $x > \frac{-1+A^{1-q}}{1-q}$. Luego, $1 + (1-q)x > A^{1-q}$. Así, $\exp_q(x) > A$. Esto prueba el límite indicado.

Similarmente se prueban las partes 2 y 3. \square

Proposición 1.4.5: Para la función $\ln_q(x)$

1. Para $q < 0$, $x \in [0, \infty)$, $\ln_q(x)$ es continua, creciente, convexa y de tal manera que

$$\lim_{x \rightarrow \infty} \ln_q(x) = \infty.$$

2. Para $0 < q < 1$, $x \in [0, \infty)$, $\ln_q(x)$ es continua, creciente, cóncava y de manera que

$$\lim_{x \rightarrow \infty} \ln_q(x) = \infty.$$

3. Para $1 < q$, $x \in (0, \infty)$, $\ln_q(x)$ es creciente, continua, cóncava y de manera que

$$\lim_{x \rightarrow \infty} \ln_q(x) = \frac{1}{q-1}.$$

Prueba:

1. Para $x \geq 0$ y $q < 0$, tenemos que $1 - q > 0$. Además, la función $g(x) = \frac{x-1}{1-q}$ es lineal, por tanto, es continua en todos los reales. Así mismo, $f(x) = x^{1-q}$ es continua en $[0, \infty)$. Luego, por el teorema de la composición de dos funciones, se tiene que $\ln_q(x) = g(f(x))$ es continua en $[0, \infty)$.

Veamos que $\ln_q(x)$ es creciente. En efecto, sean x_1, x_2 en el dominio de $\ln_q(x)$ tales que $x_1 < x_2$. Luego,

$$x_1^{1-q} - 1 < x_2^{1-q} - 1 \text{ implica } \frac{x_1^{1-q}-1}{1-q} < \frac{x_2^{1-q}-1}{1-q}$$

es decir, $\ln_q(x_1) < \ln_q(x_2)$. Así, $\ln_q(x)$ es creciente.

Otra forma consiste en observar que $[\ln_q(x)]' = \frac{1}{x^q} > 0$.

Es fácil ver que $\ln_q(x)$ es convexa, pues, $x - 1 \leq \ln_q(x)$ para $q < 0$ y $x \geq 0$.

Veamos que $\lim_{x \rightarrow \infty} \ln_q(x) = \infty$. Necesitamos mostrar que dado $A > 0$, existe $k > 0$ tal que para toda $x > k$, se tiene que $\ln_q(x) > A$. En efecto, dado $A > 0$, por la propiedad Arquimediana, existe $k > 0$ tal que $k > [A(1-q) + 1]^{\frac{1}{1-q}}$. Luego, para $x > k$ se tiene que $x > [A(1-q) + 1]^{\frac{1}{1-q}}$ entonces $x^{1-q} > A(1-q) + 1$, de donde $\frac{x^{1-q}-1}{1-q} > A$. Así, $\ln_q(x) > A$.

Esto prueba el límite indicado.

Similarmente se prueban las partes 2 y 3.

□

Las funciones exponenciales y logarítmicas q -deformadas satisfacen las siguientes propiedades:

Proposición 1.4.6:

1. $e_q^{x \oplus_q y} = e_q^x e_q^y$ y $e_q^{x \ominus_q y} = \frac{e_q^{(x)}}{e_q^{(y)}}$.
2. $\ln_q(xy) = \ln_q(x) \oplus_q \ln_q(y)$ y $\ln_q(\frac{x}{y}) = \ln_q(x) \ominus_q \ln_q(y)$.
3. $d[\ln_q(u)] = \frac{1}{u^q} du$ y $d[e_q^u] = (e_q^u)^q du$.

Prueba:

1. $e_q^{x \oplus_q y} = e_q^x e_q^y$

$$\begin{aligned}
 e_q^{x \oplus_q y} &= e_q^{x+y+(1-q)xy} \\
 &= (1 + (1-q)(x+y+(1-q)xy))^{\frac{1}{1-q}} \\
 &= (1 + (x+y)(1-q) + (1-q)^2 xy)^{\frac{1}{1-q}} \\
 &= ((1 + (1-q)x)(1 + (1-q)y))^{\frac{1}{1-q}} \\
 &= (1 + (1-q)x)^{\frac{1}{1-q}} (1 + (1-q)y)^{\frac{1}{1-q}} = e_q^x e_q^y.
 \end{aligned}$$

$$e_q^{x \ominus_q y} = \frac{e_q^{(x)}}{e_q^{(y)}}$$

$$\begin{aligned}
 e_q^{x \ominus_q y} &= [1 + (1-q)(x \ominus_q y)]^{\frac{1}{1-q}} \\
 &= [1 + (1-q) \frac{x-y}{1 + (1-q)y}]^{\frac{1}{1-q}} \\
 &= \left[\frac{1 + (1-q)x}{1 + (1-q)y} \right]^{\frac{1}{1-q}} \\
 &= \frac{[1 + (1-q)x]^{\frac{1}{1-q}}}{[1 + (1-q)y]^{\frac{1}{1-q}}} = \frac{e_q^{(x)}}{e_q^{(y)}}.
 \end{aligned}$$

$$2. \ln_q(xy) = \ln_q(x) \oplus_q \ln_q(y)$$

$$\begin{aligned} \ln_q(xy) &= \frac{(xy)^{\frac{1}{1-q}} - 1}{1-q} \\ &= \frac{(x^{1-q} - 1) + (y^{1-q} - 1) + (xy)^{1-q} - x^{1-q} - y^{1-q} + 1}{1-q} \\ &= \frac{(x^{1-q} - 1) + (y^{1-q} - 1) + (x^{1-q} - 1)(y^{1-q} - 1)}{1-q} \\ &= \frac{x^{1-q} - 1}{1-q} + \frac{y^{1-q} - 1}{1-q} + (1-q) \left(\frac{x^{1-q} - 1}{1-q} \right) \left(\frac{y^{1-q} - 1}{1-q} \right) \\ &= \frac{x^{1-q} - 1}{1-q} + \frac{y^{1-q} - 1}{1-q} + (1-q)\ln_q(x)\ln_q(y) \\ &= \ln_q(x) \oplus_q \ln_q(y). \end{aligned}$$

$$\ln_q\left(\frac{x}{y}\right) = \ln_q(x) \ominus_q \ln_q(y)$$

$$\begin{aligned} \ln_q\left(\frac{x}{y}\right) &= \frac{\left(\frac{x}{y}\right)^{1-q} - 1}{1-q} \\ &= \frac{\left(\frac{x^{1-q}}{y^{1-q}}\right) - 1}{1-q} \\ &= \frac{x^{1-q} - y^{1-q}}{(1-q)y^{1-q}} \\ &= \frac{x^{1-q} - 1 - y^{1-q} + 1}{1-q} \\ &= \frac{1 + (y^{1-q} - 1)}{1 + (1-q)\frac{(y^{1-q}-1)}{(1-q)}} \\ &= \frac{\frac{x^{1-q}-1}{1-q} - \frac{y^{1-q}-1}{1-q}}{1 + (1-q)\frac{(y^{1-q}-1)}{(1-q)}} \\ &= \frac{\ln_q(x) - \ln_q(y)}{1 + (1-q)\ln_q(y)} \\ &= \ln_q(x) \ominus_q \ln_q(y). \end{aligned}$$

$$3. d[\ln_q(u)] = \frac{1}{u^q} du$$

$$\begin{aligned} d[\ln_q(u)] &= d\left[\frac{u^{1-q} - 1}{1-q}\right] \\ &= \frac{(1-q)u^{-q}}{(1-q)} du \\ &= \frac{1}{u^q} du. \end{aligned}$$

$$d[e_q^u] = (e_q^u)^q du.$$

$$\begin{aligned} d[e_q^u] &= d \left[(1 + (1 - q)u)^{\frac{1}{1-q}} \right] \\ &= \frac{1}{1 - q} (1 + (1 - q)u)^{\frac{q}{1-q}} (1 - q) du \\ &= \left((1 + (1 - q)u)^{\frac{1}{1-q}} \right)^q = (e_q^u)^q du \end{aligned}$$

□

Proposición 1.4.7:

1. **Producto** $exp_q(x)exp_q(y) = exp_q(x + y + (1 - q)xy)$. (3)
2. **Cociente** $\frac{exp_q(x)}{exp_q(y)} = exp_q\left(\frac{x-y}{1+(1-q)y}\right)$. (4)
3. **Ley de potencia** $(exp_q x)^n = exp_{1-\frac{1-q}{n}}(nx)$. (5)
4. **Inversa** $[exp_q(x)]^{-1} = exp_q\left(\frac{-x}{1+(1-q)x}\right) = exp_{2-q}(-x)$. (6)
5. **Derivada** $\frac{d}{dx}[exp_q(x)] = (exp_q(x))^q = exp_{2-\frac{1}{q}}(qx)$. (7)
6. **Integral** $\int exp_q(nx)dx = \frac{1}{(2-q)n}(exp_q(nx))^{2-q}$. (8)
7. **Producto** $ln_q(xy) = ln_q(x) + ln_q(y) - (1 - q)ln_q(x)ln_q(y)$. (9)
8. **Cociente** $ln_q\left(\frac{x}{y}\right) = \frac{ln_q(x) - ln_q(y)}{1 + (1 - q)ln_q(y)}$. (10)
9. **Ley de potencia** $ln_q(x^n) = \frac{n}{1-q}ln_{1-n}(x^{1-q})$. (11)
10. **Inversa** $ln_q(x^{-1}) = \frac{ln_q(x)}{1 + (1 - q)ln_q(x)} = \frac{-1}{x^{1-q}}ln_q(x)$. (12)
11. **Derivada** $\frac{d}{dx}[ln_q(x)] = \frac{1}{x^q}$. (13)
12. **Integral** $\int ln_q(x)dx = \frac{x(ln_q(x) - 1)}{2 - q}$. (14)

Prueba:

$$1. exp_q(x)exp_q(y) = exp_q(x + y + (1 - q)xy).$$

$$\begin{aligned} exp_q(x)exp_q(y) &= (1 + (1 - q)x)^{\frac{1}{1-q}} (1 + (1 - q)y)^{\frac{1}{1-q}} \\ &= ((1 + (1 - q)x)(1 + (1 - q)y))^{\frac{1}{1-q}} \\ &= (1 + (x + y)(1 - q) + (1 - q)^2 xy)^{\frac{1}{1-q}} \\ &= (1 + (1 - q)(x + y + (1 - q)xy))^{\frac{1}{1-q}} \\ &= exp_q(x + y + (1 - q)xy) \end{aligned}$$

$$2. \frac{\exp_q(x)}{\exp_q(y)} = \exp_q\left(\frac{x-y}{1+(1-q)y}\right).$$

$$\begin{aligned} \frac{\exp_q(x)}{\exp_q(y)} &= \frac{[1 + (1-q)x]^{\frac{1}{1-q}}}{[1 + (1-q)y]^{\frac{1}{1-q}}} \\ &= \left[\frac{1 + (1-q)x}{1 + (1-q)y} \right]^{\frac{1}{1-q}} \\ &= \left[\frac{(1 + (1-q)y) + ((1-q)x - (1-q)y)}{1 + (1-q)y} \right]^{\frac{1}{1-q}} \\ &= \left[1 + (1-q) \frac{x-y}{1 + (1-q)y} \right]^{\frac{1}{1-q}} \\ &= \exp_q\left(\frac{x-y}{1 + (1-q)y}\right). \end{aligned}$$

$$3. (\exp_q x)^n = \exp_{1-\frac{1-q}{n}}(nx).$$

$$\begin{aligned} (\exp_q x)^n &= ([1 + (1-q)x]^{\frac{1}{1-q}})^n \\ &= [1 + (1-q)x]^{\frac{n}{1-q}} \\ &= \left[1 + \left(1 - \frac{1-q}{n}\right)nx \right]^{\frac{1}{1-\frac{1-q}{n}}} \\ &= \exp_{1-\frac{1-q}{n}}(nx). \end{aligned}$$

$$4. [\exp_q(x)]^{-1} = \exp_q\left(\frac{-x}{1+(1-q)x}\right) = \exp_{2-q}(-x).$$

$$\begin{aligned} [\exp_q(x)]^{-1} &= [1 + (1-q)x]^{-\frac{1}{1-q}} \\ &= \left[\frac{1}{1 + (1-q)x} \right]^{\frac{1}{1-q}} \\ &= \left[\frac{1 + (1-q)x - (1-q)x}{1 + (1-q)x} \right]^{\frac{1}{1-q}} \\ &= \left[1 + (1-q) \left(\frac{-x}{1 + (1-q)x} \right) \right]^{\frac{1}{1-q}} \\ &= \exp_q\left(\frac{-x}{1 + (1-q)x}\right). \end{aligned}$$

$$\begin{aligned} [\exp_q(x)]^{-1} &= [(1 + (1-q)x)^{\frac{1}{1-q}}]^{-1} \\ &= [(1 + (1-q)x)^{\frac{1}{1-q}}]^{-\frac{1-q}{1-q}} \\ &= [1 + (1-q)x]^{-\frac{1}{1-q}} \\ &= [1 + (1-q)x]^{\frac{1}{1-(2-q)}} \\ &= [1 + (1-(2-q))(-x)]^{\frac{1}{1-(2-q)}} \\ &= \exp_{2-q}(-x). \end{aligned}$$

5. $d[e_q^u] = (e_q^u)^q du$.

$$\begin{aligned} d[e_q^u] &= d\left[(1 + (1 - q)u)^{\frac{1}{1-q}}\right] \\ &= \frac{1}{1 - q}(1 + (1 - q)u)^{\frac{q}{1-q}}(1 - q)du \\ &= \left((1 + (1 - q)u)^{\frac{1}{1-q}}\right)^q = (e_q^u)^q du \end{aligned}$$

6. $\int exp_q(nx)dx = \frac{1}{(2-q)n}(exp_q(nx))^{2-q}$. Haremos la sustitución $u = 1 + (1 - q)nx$.

$$\begin{aligned} \int exp_q(nx)dx &= \int (1 + (1 - q)nx)^{\frac{1}{1-q}} dx \\ &= \frac{1}{(1 - q)n} \int (1 + (1 - q)nx)^{\frac{1}{1-q}} (1 - q)ndx \\ &= \frac{1}{(1 - q)n} \int u^{\frac{1}{1-q}} du \\ &= \frac{1}{(1 - q)n} \frac{u^{\frac{2-q}{1-q}}}{\frac{2-q}{1-q}} \\ &= \frac{1}{(2 - q)n} [(1 + (1 - q)nx)^{\frac{1}{1-q}}]^{2-q} \\ &= \frac{1}{(2 - q)n} (exp_q(nx))^{2-q}. \end{aligned}$$

7. $ln_q(xy) = ln_q(x) + ln_q(y) - (1 - q)ln_q(x)ln_q(y)$.

$$\begin{aligned} ln_q(xy) &= \frac{x^{1-q}y^{1-q} - 1}{1 - q} \\ &= \frac{x^{1-q} - 1 + y^{1-q} - 1 + x^{1-q}y^{1-q} - x^{1-q} - y^{1-q} + 1}{1 - q} \\ &= \frac{(x^{1-q} - 1) + (y^{1-q} - 1) + (x^{1-q} - 1)(y^{1-q} - 1)}{1 - q} \\ &= \frac{x^{1-q} - 1}{q - 1} + \frac{y^{1-q} - 1}{q - 1} + (1 - q) \frac{(x^{1-q} - 1)(y^{1-q} - 1)}{(1 - q)(1 - q)} \\ &= ln_q(x) + ln_q(y) - (1 - q)ln_q(x)ln_q(y). \end{aligned}$$

$$8. \ln_q\left(\frac{x}{y}\right) = \frac{\ln_q(x) - \ln_q(y)}{1 + (1-q)\ln_q(y)}.$$

$$\begin{aligned} \ln_q\left(\frac{x}{y}\right) &= \frac{\frac{x^{1-q}}{y^{1-q}} - 1}{1 - q} \\ &= \frac{x^{1-q} - y^{1-q}}{(1-q)y^{1-q}} \\ &= \frac{(x^{1-q} - y^{1-q})(1-q)}{(1-q)^2} \\ &= \frac{(1-q)y^{1-q}}{(1-q)^2} \\ &= \frac{(x^{1-q} - y^{1-q})(1-q)}{y^{1-q}} \\ &= \frac{(x^{1-q} - 1)(1-q) - (y^{1-q} - 1)(1-q)}{(1-q)^2} \\ &= \frac{1 + (1-q)\frac{y^{1-q} - 1}{1-q}}{1 + (1-q)\frac{y^{1-q} - 1}{1-q}} \\ &= \frac{\frac{x^{1-q} - 1}{1-q} - \frac{y^{1-q} - 1}{1-q}}{1 + (1-q)\frac{y^{1-q} - 1}{1-q}} \\ &= \frac{\ln_q(x) - \ln_q(y)}{1 + (1-q)\ln_q(y)} \end{aligned}$$

$$9. \ln_q(x^n) = \frac{n}{1-q} \ln_{1-n}(x^{1-q}).$$

$$\begin{aligned} \ln_q(x^n) &= \frac{(x^n)^{1-q} - 1}{1 - q} \\ &= \frac{(x^{1-q})^n - 1}{1 - q} \\ &= \frac{n}{1 - q} \frac{(x^{1-q})^n - 1}{n} \\ &= \frac{n}{1 - q} \frac{(x^{1-q})^{1-(1-n)} - 1}{1 - (1-n)} \\ &= \frac{n}{1 - q} \ln_{1-n}(x^{1-q}). \end{aligned}$$

$$10. \ln_q(x^{-1}) = \frac{\ln_q(x)}{1+(1-q)\ln_q(x)} = \frac{-1}{x^{1-q}} \ln_q(x).$$

$$\begin{aligned} -\frac{\ln_q(x)}{1+(1-q)\ln_q(x)} &= -\frac{\frac{x^{1-q}-1}{1-q}}{1+(1-q)\frac{x^{1-q}-1}{1-q}} \\ &= -\frac{\frac{x^{1-q}-1}{1-q}}{x^{1-q}} \\ &= -\frac{x^{1-q}-1}{(1-q)x^{1-q}} \\ &= -\frac{1}{x^{1-q}} \frac{x^{1-q}-1}{1-q} \\ &= -\frac{1}{x^{1-q}} \ln_q(x) \end{aligned}$$

$$11. d[\ln_q(u)] = \frac{1}{u^q} du$$

$$\begin{aligned} d[\ln_q(u)] &= d\left[\frac{u^{1-q}-1}{1-q}\right] \\ &= \frac{(1-q)u^{-q}}{(1-q)} du \\ &= \frac{1}{u^q} du. \end{aligned}$$

$$12. \int \ln_q(x) dx = \frac{x(\ln_q(x)-1)}{2-q}.$$

$$\begin{aligned} \int \ln_q(x) dx &= \int \frac{x^{1-q}-1}{1-q} dx \\ &= \frac{1}{(1-q)} \int (x^{1-q}-1) dx \\ &= \frac{1}{(1-q)} \left[\frac{x^{2-q}}{2-q} - x \right] \\ &= \frac{1}{(1-q)(2-q)} [x(x^{1-q}-2+q)] \\ &= \frac{1}{(2-q)} x \left[\frac{x^{1-q}-1-1+q}{1-q} \right] \\ &= \frac{1}{(2-q)} x \left[\frac{x^{1-q}-1}{1-q} - 1 \right] \\ &= \frac{1}{(2-q)} x [\ln_q(x) - 1] \\ &= \frac{x(\ln_q(x)-1)}{2-q}. \end{aligned}$$

□

Capítulo 2

Variedades estándar de información estadística

En este capítulo se comentan las variedades de información estadísticas (estándares) construidas por Amari y Pistone y Sempi. No se presentan los desarrollos matemáticos ya que la mayoría de ellos se pueden deducir como límite, cuando q tiende a uno, de las respectivas generalizaciones que se mostrarán en los capítulos siguientes.

2.1. Variedad estándar de Amari (caso paramétrico)

A diferencia de Pistone y Sempi que construyen una variedad no paramétrica infinito dimensional sobre el conjunto de todas las densidades de probabilidad estrictamente positivas \mathfrak{M}_μ , Amari define su variedad paramétrica sobre la familia exponencial, es decir, cuando la función de densidad puede ser escrita en la forma

$$p(x, \theta) = \exp\{\theta^i x_i - \psi(\theta)\} \quad (2.1)$$

con respecto a alguna medida adecuada $P(x)$, eligiendo una parametrización apropiada $\theta = (\theta^i)$ y adecuada variable aleatoria $x = (x_i)$.

Considerando la familia de distribuciones de probabilidad $S = \{p(x, \theta)\}$ de un modelo estadístico, podemos definir un mapeo

$$\phi : S \rightarrow \mathbb{R}^n, \text{ por } \phi[p(x, \theta)] = \theta.$$

Usando este mapeo como función de coordenada, el vector θ es considerado como coordenada de la distribución $p(x, \theta)$, dando lugar a una estructura diferenciable en S . Finalmente, tenemos que S es una variedad diferenciable.

Si consideramos otro parámetro vectorial ξ que este conectado a θ por un difeomorfismo, observaremos que ξ define otro sistema de coordenadas en S , que no afecta sus propiedades

geométricas intrínsecas, pues, son independientes de éstos [20]. A fin de estructurar las propiedades intrínsecas de un modelo estadístico, Amari introduce las α -conexiones, que coinciden con la conexión exponencial cuando α tiende a uno.

Una función divergente es aquella que cumple dos condiciones:

1. $D[x : y] \geq 0$, la igualdad se da cuando $x = y$;
2. $D[x : y]$ es diferenciable y la Hessiana con respecto a x en $y = x$ es definida positiva.

$D[x : y]$ no es necesariamente simétrica con respecto a x e y . Por lo tanto, no es una distancia. La desigualdad triangular no se sostiene [25].

Dada una función divergencia en una variedad, podemos derivar una métrica de Riemann y un dual par de conexiones afines, que son los componentes esenciales de la geometría de la información. En el caso de una familia de distribuciones de probabilidad, su estructura geométrica-información se da a partir de la propiedad de invariancia. Consiste de la métrica de información de Fisher y las α -conexiones.

De otra parte, la variedad de distribuciones de probabilidad discretas, que es el conjunto de todas las distribuciones de probabilidad en un conjunto finito, tiene una estructura de Riemann dualmente plana. Estas propiedades están relacionadas con la función de divergencia subyacente. Definimos una divergencia invariante en términos de monotonía información, lo que nos lleva a la clase de f -divergencias. Necesitamos considerar una función de divergencia que le de estructura afín dual plana. Esto está dado por la divergencia Bregman en términos de una función convexa. La divergencia invariante y plana en la variedad de distribuciones de probabilidad es la divergencia de Kullback-Leibler, y esto es único, pero generalmente es la clase de α -divergencias en la variedad de medidas positivas. Una transformación no lineal de una función de divergencia o una función convexa provoca un cambio conformacional de la estructura geométrica dual. En este sentido, podemos hablar de la doble geometría derivada de la entropía de Tsallis o Rényi. La familia q -exponencial y la q -estructura afín, es una generalización de la actual geometría de información invariante [25, 24].

2.2. Variedad de Pistone y Sempi(caso no paramétrico)

El tema principal de la geometría diferencial moderna ha sido caracterizar las propiedades globales de las variedades, y la teoría se ha desarrollado hacia este fin. La geometría de información comenzó como el estudio geométrico de estimación estadística. Esto implicó ver el conjunto de distribuciones de probabilidad que constituyen un modelo estadístico como una variedad, y el análisis de la relación entre la estructura geométrica de esta variedad y la estimación estadística utilizando este modelo. A continuación se describe la variedad de información estadística para el caso no paramétrico, ver [26, 34, 4].

Pistone y Sempi construyen una variedad estadística de información (modelada sobre espacios de funciones de Banach) sobre el conjunto de todas las densidades de probabilidad estrictamente positivas \mathfrak{M}_μ , permitiendo interpretar la entropía clásica relativa a dos funciones de densidad de probabilidad conectadas por un modelo exponencial. Esto es, dado el espacio de probabilidad $(\Omega, \mathcal{B}, \mu)$ tenemos que

$$\mathfrak{M}_\mu = \{p : \Omega \rightarrow \mathbb{R}, \mu - \text{medible}, p > 0 \mu.a.s, E_p = 1\}$$

donde, para cada $p \in \mathfrak{M}_\mu$, E_p es un operador lineal llamado el valor esperado respecto a la medida con densidad p , cuyo valor viene dado por la expresión

$$E_p(u) = \int_{\Omega} u p d\mu$$

dada una variable aleatoria $u : \Omega \rightarrow \mathbb{R}$.

Dada la función de Young, $\Phi_1(x) = \cosh(x) - 1$, construimos la correspondiente clase de Cramer

$$B_p = \{u \in L^{\Phi_1}(p \cdot \mu) : E_p(u) = 0\}$$

que constituye un subespacio cerrado de $L^1(p \cdot \mu)$.

B_p constituye un subespacio convexo y cerrado de $L^{\Phi_1}(p \cdot \mu)$. En consecuencia, B_p es de Banach.

Es importante notar que probabilísticamente, $L^1(p \cdot \mu)$ coincide con $\{u : E_p(e^{tu}) < \infty, \forall t \in \mathbb{R}\}$.

Sea

$$\mathcal{V}_p = \{u \in B_p : \|u\|_p \leq 1\}.$$

Definamos la función inyectiva $\xi_p : \mathcal{V}_p \rightarrow \mathfrak{M}_\mu$ de tal forma que $\xi_p(u) := e^{u - \varpi_p(u)} \cdot p$, donde $\varpi_p(u) = \ln E_p(u) = \ln \int u p d\mu$.

Si denotamos por \mathcal{U}_p el rango de la función ξ_p , tenemos que $\xi_p : \mathcal{V}_p \rightarrow \mathcal{U}_p$ es una función sobreyectiva. En consecuencia, es biyectiva, y su respectiva inversa S_p viene dada por

$$S_p : \mathcal{U}_p \rightarrow \mathcal{V}_p$$

donde, si $g \in \mathcal{U}_p \subset \mathfrak{M}_\mu$,

$$S_p(g) = \ln \frac{g}{p} - E_p[\ln \frac{g}{p}]$$

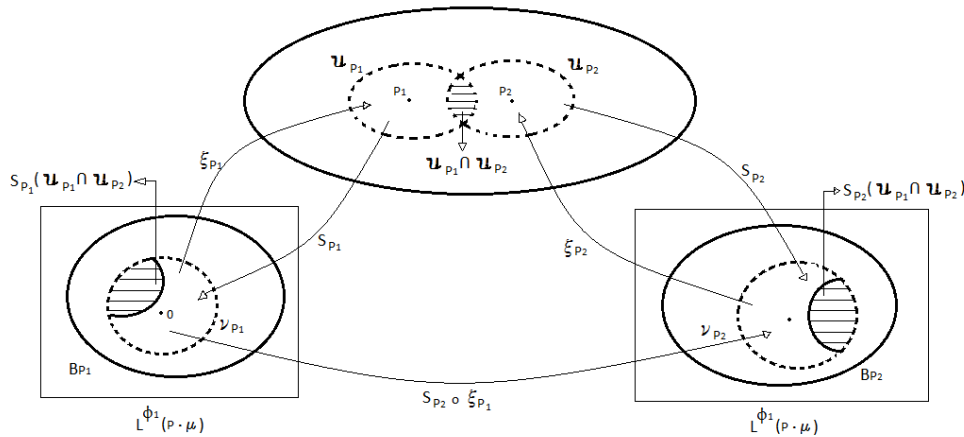
Las funciones $S_p, p \in \mathfrak{M}_\mu$, serán los mapeos de coordenadas de la variedad. El par (\mathcal{U}_p, S_p) es una carta. Esto da origen a las siguientes funciones

$$S_{p_2} \circ \xi_{p_1} : S_{p_1}(\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2}) \rightarrow S_{p_2}(\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2})$$

donde $(S_{p_2} \circ \xi_{p_1})(u) = u + \ln \frac{p_1}{p_2} - E_{p_2}[u + \ln \frac{p_1}{p_2}]$ con $p_1, p_2 \in \mathfrak{M}_\mu$.

Se puede notar que los mapeos de transición difieren de la función identidad en una constante dada por $\ln \frac{p_1}{p_2} - E_{p_2}[u + \ln \frac{p_1}{p_2}]$, por tanto son continuos y de clase C^∞ .

Los cálculos algebraicas se realizan en el espacio de las μ -clases de funciones medibles y la esperanza está bien definida, siempre y cuando $\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2} \neq \emptyset$ debido a que esto implica $u + \ln \frac{p_1}{p_2} \in \mathcal{V}_p$.



Según el teorema 3.6 [26], la colección de pares $\{(\mathcal{U}_p, S_p) : p \in \mathfrak{M}_\mu\}$ es un atlas de clase C^∞ sobre \mathfrak{M}_μ e induce una topología equivalente a la topología de la convergencia exponencial en las sucesiones de \mathfrak{M}_μ . Además, esta colección define una variedad infinito dimensional.

Esta estructura de variedad nos muestra que la noción geométrica de vector tangente corresponde a la noción estadística de la probabilidad score.

La variedad construida por Pistone y Sempi, donde las divergencias son la base de la construcción, ha contribuido en el estudio de los estados físicos de sistemas clásicos y cuánticos.

Pistone y Sempi (1995) introducen una variedad estadística exponencial no paramétrica consistente de todas las densidades que son absolutamente continuas con respecto a una dada, señalando una fórmula explícita para una carta de dicha variedad.

Entre todas las curvas que conectan dos puntos Q y K , la que da la distancia mínima se

denomina geodésica Riemann conectando Q y K .

Loaiza-Quiceno construyen una variedad de Banach que permite pormenorizar la entropía relativa según Tsallis, con la finalidad que sirva como modelo geométrico para caracterizar sistemas no extensivos. Adicionalmente, se obtiene la variedad de información de Pistone y Sempi bajo la condición $q \rightarrow 1$, constituyéndose en una generalización de la variedad señalada.

En [1], para $q \rightarrow 1$, se obtiene la matrix de información de Fisher.

La métrica de información de Fisher es una particular métrica de Riemann [32], que se puede definir en una variedad estadística múltiple suave, es decir, una variedad suave cuyos puntos son medidas de probabilidad definidas en un espacio de probabilidad común. La métrica es interesante en varios aspectos. En primer lugar, se puede interpretar como la forma infinitesimal de la entropía relativa (es decir, la divergencia de Kullback-Leibler); específicamente, es el Hessian de la divergencia. Alternativamente, se puede entender como la métrica inducida por el espacio métrico euclidiano plano, después de cambios apropiados de la variable. Cuando se extiende al espacio de Hilbert proyectivo complejo, se convierte en la métrica Fubini-Study, y cuando se escribe en términos de estados mixtos, es la métrica quantum Bures.

Dada una variedad estadística, con coordenadas dadas por $\theta = (\theta_1, \theta_2, \dots, \theta_n)$, denotamos por $p(x, \theta)$ la distribución de probabilidad. Aquí, x es un valor específico elaborado a partir de una colección de variables aleatorias (discretas o continuas) X .

La métrica de información de Fisher entonces toma la forma

$$g_{j,k}(\theta) = \int_X \frac{\partial \log p(x, \theta)}{\partial \theta_j} \frac{\partial \log p(x, \theta)}{\partial \theta_k} p(x, \theta) dx.$$

La integral se realiza sobre todos los valores x de todas las variables aleatorias X .

Capítulo 3

Variedad de información q -exponencial: caso paramétrico

La mecánica estadística actúa como nexo de diferentes ramas de la física, a este fin busca caracterizar un sistema a nivel macroscópico a partir de las propiedades microscópicas. Boltzmann y Gibbs difieren al explicar el comportamiento irreversible de los sistemas macroscópicos: Gibbs considera la ergodicidad como requisito para la irreversibilidad, pero no necesariamente un número alto de grados de libertad del sistema; en cambio, Boltzmann señala la necesidad de un número alto de grados de libertad pero que el sistema sea ergódico no es necesario. A éste respecto Lombardi-Labarca puntualizan que existen condiciones físicas definidas bajo las cuales los resultados de ambos enfoques se aproximan lo suficiente para ser admitidos desde la práctica de la física; y la divergencia radica, primordialmente, en la utilización de diferentes conceptos de equilibrio [32]. Sin embargo, se han hecho importantes avances en el estudio de microestados. Particularmente, el enfoque de Gibbs ha contribuido en esta línea de estudio. Nos interesaremos en la familia q -exponencial como una generalización de las funciones de potencia que son útiles para estudiar varios fenómenos complejos o físicos no estándares, obteniendo una nueva estructura matemática con geometría dualmente plana derivada de la transformación de Legendre y la geometría conformal es útil para entenderla [1].

La q -Gaussiana tiene aplicaciones de la mecánica estadística, la geología, la anatomía, la astronomía, la economía y las finanzas, entre otras. Un ejemplo representativo de la situación planteada lo constituye la q -entropía de Tsallis, quien obtiene y proporciona nuevas posibilidades para construir la física estadística generalizada que recupera no sólo las tradicionales mecánica estadística de Boltzmann-Gibbs, sino que permite describir los comportamientos de la ley de potencia de forma sistemática. Con base en la geometría de la información, obtenemos una familia de distribuciones q -exponenciales con estructura ubicua y dualmente plana, mediante la incursión de una métrica Riemanniana y un par de conexiones duales afines, en la que son válidos los teoremas q -pitagóricos y entropía q -máxima.

3.1. Familia q -exponencial o q -Gibbs

Dado que algunos sistemas complejos o físicos no estándares no encajan en la conocida distribución de Gibbs, sino en una ley de potencia, se hace una generalización de la actual geometría de información invariante mediante la distribución q -Gibbs o familia q -exponencial, utilizando el índice q de la entropía Tsalli, lo que permite elucidar varios fenómenos físicos de este tipo.

La familia q -exponencial

La siguiente es una q -generalización de la familia exponencial (2.1), que resulta conveniente para algunos modelos físicos, y que se llamará familia q -exponencial.

$$f_{\beta}(x) = c(x) \exp_q(-\alpha(\beta) - \beta H(x)) \quad \text{y} \quad f_{\beta}(x) = 0 \text{ cuando la exponencial diverge, (3.1)}$$

donde la función $H(x)$ es el Hamiltoniano del sistema, el parametro β es usualmente la temperatura inversa, $\alpha(\beta)$ es la normalización y la función $c(x)$ es una distribución a priori que no debe depender del parámetro β . Esta clase de modelos también se puede derivar usando el principio de máxima entropía optimizando el funcional de entropía de Tsallis. Si un modelo es de la forma anterior, entonces se dice que pertenecen a la familia q -exponencial; la cual recupera a la familia exponencial estándar (2.1) cuando $q \rightarrow 1$; que en este caso queda descrita de la forma

$$f_{\beta}(x) = c(x) \exp(-\alpha(\beta) - \beta H(x)),$$

y se conoce como la distribución de Boltzmann-Gibbs.

Una reformulación de (3.1) es por lo tanto que, o bien $f(x) = 0$ o

$$\ln_q\left(\frac{f_{\beta}(x)}{c(x)}\right) = -\alpha(\beta) - \beta H(x)$$

Los modelos que pertenecen a tal familia comparten un número de propiedades interesantes. En particular, todos ellos encajan en el formalismo termodinámico. Como consecuencia, la densidad de probabilidad $f_{\beta}(x)$ puede ser considerada para ser la distribución equilibrio del modelo en el valor del parámetro dado.

A continuación mediante ejemplos mostraremos cómo el formalismo generalizado se ve cuando lo restringimos al contexto de la física estadística no extensiva.

La distribución q -Gaussiana ($q < 3$)

Las aplicaciones de la distribución normal o gaussiana en diversas áreas del conocimiento es algo reconocido por todos. La generalización a la familia q -exponencial, particularmente, es relevante para la descripción estadística de pequeños sistemas aislados. Jan Naudts

hace una generalización de la distribución de Gauss, y como aplicación señala que la distribución del momento de una sola partícula es una q -gaussiana [28].

La q -Gaussiana es una distribución de probabilidad dada por

$$f(x) = \frac{1}{c_q \sigma} \exp_q\left(-\frac{x^2}{\sigma^2}\right), \quad (3.2)$$

con

$$c_q = \int_{-\infty}^{\infty} \exp_q(-x^2) dx = \sqrt{\frac{\pi}{q-1}} \frac{\Gamma(-\frac{1}{2} + \frac{1}{q-1})}{\Gamma(\frac{1}{q-1})} \text{ si } 1 < q < 3$$

$$c_q = \sqrt{\frac{\pi}{q-1}} \frac{\Gamma(1 + \frac{1}{q-1})}{\Gamma(\frac{3}{2} + \frac{1}{q-1})} \text{ si } q < 1.$$

Esta distribución puede ser derivada de la maximización de la entropía de Tsallis bajo restricciones adecuadas. El dominio de \mathbf{x} en el caso q -Gausiano depende de q si $0 < q < 1$. En todo caso (3.2) es un ejemplo de una distribución Tsallis y puede ser llevada a la forma (3.1).

Tomemos, por ejemplo $q = \frac{1}{2}$. Entonces q -Gaussiana se convierte en

$$f(x) = \frac{15\sqrt{2}}{32\sigma} \left[1 - \frac{x^2}{\sigma^2}\right]_+.$$

Tenga en cuenta que esta distribución se anula fuera del intervalo $[-\sigma, \sigma]$. Entonces, el caso $q \rightarrow 1$, reproduce la distribución de Gauss convencional. Para $q = 2$ se obtiene

$$f(x) = \frac{1}{\pi} \frac{\sigma}{x^2 + \sigma^2}.$$

Esta se conoce como la distribución de Cauchy, la cual no tiene momentos finitos de orden mayor o igual a uno; sólo existen momentos absolutos fraccionarios. Su importancia en la física es el resultado de ser la solución a la ecuación diferencial que describe la resonancia forzada. En su forma estándar, es la distribución de probabilidad de máxima entropía para una variable aleatoria X para la que $E[\ln(1 + X^2)] = \ln(4)$ [29]. También se conoce, especialmente entre los físicos, como la distribución de Lorentz, la distribución de Cauchy-Lorentz, o la distribución de Breit-Wigner. En el rango $1 \leq q < 3$ la q -Gaussiana es estrictamente positiva en toda la línea. Para $q < 1$ es cero fuera de un intervalo. Para $q < 3$ la distribución no puede ser normalizada, porque

$$f(x) \sim \frac{1}{|x|^{\frac{2}{q-1}}} \text{ cuando } |x| \rightarrow \infty.$$

Distribuciones Kappa $1 < q < \frac{5}{3}$

El uso de distribuciones Kappa viene teniendo mayor importancia para la física del plasma espacial y la astrofísica. Una expresión empírica para estas distribuciones se introdujo

en 1968, y desde entonces, se emplean en numerosos estudios del viento solar. Recientes observaciones indican que iones en la heliosfera exterior están bien descritos por distribuciones Kappa; análisis teóricos de los iones y átomos neutros energéticos ya han comenzado a depender en gran medida de estas distribuciones [30].

La distribución Kappa viene dada por la expresión

$$f(v) = \frac{1}{A(k)v_0^3} \frac{v^2}{\left(1 + \frac{1}{k-a} \frac{v^2}{v_0^2}\right)^{1+k}}. \quad (3.3)$$

Esta distribución se obtiene mediante una deformación de la distribución de Maxwell, y ha sido utilizada en la física de espacio durante más de cuatro décadas. Ella presenta una ley de potencia como decaimiento $f(v) \simeq v^{-2k}$ para v grande. Los valores finitos de k corresponden a estados estacionarios fuera de equilibrio.

Una forma funcional empírica para describir la distribución de energía a lo largo de todo el espectro, tanto en el núcleo de Maxwell de baja energía como en la alta energía de la cola de la ley de potencia, fue propuesto por primera vez por *Vasyliūnas* [1968].

El trabajo de *Vasyliūnas* se relacionó con un estudio de electrones de baja energía de la magnetosfera de la Tierra. Desde entonces, esta distribución empírica se ha utilizado para describir los iones en diversas magnetosferas.

En contraste con el uso de la mecánica estadística de Tsallis, intentar derivar teóricamente una distribución Kappa de la mecánica estadística estándar Boltzmann-Gibbs es altamente problemático. Este enfoque ha sido intentado por varios autores, sin embargo, no es de extrañar que no sea posible desarrollar una distribución kappa sólida en el marco de la mecánica estadística de Boltzmann-Gibbs, ya que las estadísticas de Boltzmann-Gibbs no incluyen los sistemas en los estados estacionarios fuera del equilibrio.

La expresión (3.3) se puede escribir en la forma q -exponencial con $q = 1 + \frac{1}{1+k}$ y

$$f(v) = \frac{v^2}{A(k)v_0^3} \exp_q\left(-\frac{1}{2-q-(q-1)a} \frac{v^2}{v_0^2}\right). \quad (3.4)$$

Sin embargo, para poder ser de la forma (3.1), el prefactor de (3.4) debe no depender del parámetro v_0 . Introduce una constante arbitraria $c > 0$ con las dimensiones de una velocidad. La que puede escribirse

$$f(v) = \frac{4\pi v^2}{c^3} \exp_q\left(-\ln_{2-q}\left[4\pi A(k)\left(\frac{v_0}{c}\right)^3\right] - \left[4\pi A(k)\left(\frac{v_0}{c}\right)^3\right]^{q-1} h(q, v)\right).$$

donde $h(q, v) = \frac{1}{2-q-(q-1)a} \frac{v^2}{v_0^2}$.

En el caso $q \rightarrow 1$, se obtiene la distribución de Maxwell, que es ampliamente conocida como la base de la teoría cinética de los gases.

Velocidad del oscilador armónico ($q = 3$)

La distribución de las velocidades v de un oscilador armónico clásico está dada por

$$f(v) = \frac{1}{\pi} \frac{1}{\sqrt{v_0^2 - v^2}}.$$

Diverge cuando $|v|$ se acerca a su valor máximo v_0 y se desvanece para $|v| > v_0$. Esta distribución puede escribirse en la forma (3.1) de una familia q -exponencial con $q = 3$. Para ello, sea $x = v$ y

$$c(x) = \frac{\sqrt{2}}{\pi|v|},$$

$$\beta = \frac{1}{2}mv_0^2,$$

$$H(x) = \frac{1}{\frac{1}{2}mv^2} = \frac{1}{K},$$

$$\alpha(\beta) = -\frac{3}{2}.$$

Distribución binomial

Según [31] se tiene lo siguiente. Sea X una variable aleatoria n -dimensional con $n \geq 1$ y $0 < q < 1$ un parámetro. Representamos la función de probabilidad **mass** (pmf) de X en la forma de la familia exponencial con parámetro q .

$$\begin{aligned} f(x | q) &= \binom{n}{x} q^x (1-q)^{n-x} I_{\{x \in \{0,1,2,\dots,n\}\}} \\ &= \binom{n}{x} \left(\frac{q}{1-q}\right)^x (1-q)^n I_{\{x \in \{0,1,2,\dots,n\}\}} \\ &= \binom{n}{x} \exp\left[x \ln \frac{q}{1-q} + n \ln(1-q)\right] I_{\{x \in \{0,1,2,\dots,n\}\}} \\ &= \binom{n}{x} I_{\{x \in \{0,1,2,\dots,n\}\}} \exp\left[n \ln(1-q) + \ln \frac{q}{1-q} x\right] \\ &= c(x) \exp[-\alpha(\beta) - \beta H(x)]. \end{aligned}$$

Tomando $c(x) = \binom{n}{x} I_{\{x \in \{0,1,2,\dots,n\}\}}$, $\alpha(\beta) = -n \ln(1-q)$, $\beta = -\ln \frac{q}{1-q}$ y $H(x) = x$, tenemos representada la pmf $f(x | q)$ en la forma de la familia exponencial con parámetro q , siempre y cuando $q \in (0, 1)$. Para $q = 0$ o $q = 1$, la distribución se convierte en una distribución de un punto. En consecuencia, la familia de las distribuciones $\{f(x | q), 0 < q < 1\}$ forma una familia exponencial con parámetro q , pero si alguno de los valores límite $q = 0$ o $q = 1$, está incluido, la familia no está en la familia exponencial.

Distribución gamma

Supongamos que X tiene la densidad Gamma $\frac{\exp(-\frac{x}{\lambda})x^{q-1}}{\lambda^q \Gamma(q)} I_{x>0}$. Como tal, tiene dos parámetros λ y q . Si asumimos que λ se conoce, entonces podemos escribir la densidad en forma de la familia exponencial con parámetro q ,

$$f(x|q) = \exp[q \log x - q(\log \lambda) - \log \Gamma(q)] \exp\left(-\frac{x}{\lambda}\right) I_{x>0},$$

que tiene la forma de una densidad en la familia exponencial si tomamos $\beta = -q$, $H(x) = \log x$, $\alpha(\beta) = q(\log \lambda) + \log \Gamma(q)$, y $c(x) = \exp(-\frac{x}{\lambda}) I_{x>0}$. Γ es la función gamma.

Una inusual distribución Gamma

Supongamos que tenemos una densidad Gamma en la que se conoce la media, por ejemplo, $E(X) = 1$. Esto significa que $q\lambda = 1 \Rightarrow \lambda = \frac{1}{q}$. Parametrizando la densidad con q , tenemos

$$\begin{aligned} f(x|q) &= \exp[-qx + q\log(x)] \frac{q^q}{\Gamma(q)} \frac{1}{x} I_{x>0} \\ &= \exp[q(\log(x) - x) - (\log \Gamma(q) - q\log q)] \frac{1}{x} I_{x>0} \\ &= \frac{1}{x} I_{x>0} \exp[q(\log(x) - x) - (\log \Gamma(q) - q\log q)]. \end{aligned}$$

que también es de la forma de la familia exponencial con un parámetro, tomando $\beta = -q$, $H(x) = \log(x) - x$, $\alpha(\beta) = \log \Gamma(q) - q\log q$, y $c(x) = \frac{1}{x} I_{x>0}$.

A continuación presentamos la forma estándar de una familia exponencial de distribuciones. En efecto, sea $\mathbf{x} = (x_1, \dots, x_n)$ un conjunto de variables aleatorias y $\theta = (\theta^1, \dots, \theta^n)$ parámetros canónicos para describir el sistema latente. Entonces

$$p(\mathbf{x}, \theta) = \exp\{\sum \theta^i x_i - \psi(\theta)\} \quad (2.1)$$

con respecto a una medida adecuada $\mu(\mathbf{x})$. $\psi(\theta)$ es llamada la energía libre, que es la función generadora de acumulante.

Una generalización de las funciones de potencia de la distribución de Gibbs es

$$p(\mathbf{x}, \theta) = \exp_q\{\theta \cdot \mathbf{x} - \psi_q(\theta)\} \quad (3.5)$$

$$\log_q\{p(\mathbf{x}, \theta)\} = \theta \cdot \mathbf{x} - \psi_q(\theta)$$

donde $\theta \cdot \mathbf{x} = \sum \theta^i x_i$. Esta es la distribución q -Gibbs o la familia q -exponencial, la cual denotamos por S , donde el dominio de \mathbf{x} se restringe tal que $p(\mathbf{x}, \theta) > 0$ se cumple. La función $\psi_q(\theta)$ llamada la energía q -libre o la función q -potencial, es determinada por la condición de normalización:

$$\int \exp_q\{\theta \cdot \mathbf{x} - \psi_q(\theta)\} d\mathbf{x} = 1$$

donde reemplazamos $d\mu(\mathbf{x})$ por $d\mathbf{x}$ por brevedad. La distribución q -Gaussiana es dada por

$$p(x, \mu, \sigma) = \exp_q\left\{-\frac{(x-\mu)^2}{2\sigma^2} - \psi(\mu, \sigma)\right\}$$

y constituye una generalización de la de Gauss de la misma manera que la entropía de Tsallis es una generalización de la entropía estándar de Boltzmann-Gibbs o la entropía de Shannon. La distribución normal es un caso particular de ésta cuando $q \rightarrow 1$.

De otra parte, usamos la familia S_n de distribuciones discretas sobre $(n + 1)$ elementos $X = \{x_0, x_1, \dots, x_n\}$. La variable aleatoria x toma valores sobre X .

Supóngase que $p_i = \text{Prob}\{x = x_i\}$ y denotemos la distribución de probabilidad por el vector $p = (p_0, p_1, \dots, p_n)$ donde

$$\sum_{i=1}^n p_i = 1$$

La probabilidad de x también puede ser escrita como

$$p(x) = \sum_{i=0}^n p_i \delta_i(x) \tag{3.6}$$

donde

$$\delta_i(x) = \begin{cases} 1 & \text{si } x = x_i \\ 0 & \text{en otro caso.} \end{cases}$$

Teorema 3.1.1: La familia S_n de distribuciones de probabilidad discretas tiene la estructura de una familia q -exponencial para cualquier q .

Prueba: Para la prueba emplearemos el hecho que, para cualquier función $f(u)$, tenemos

$$f\{\sum_{i=1}^n p_i \delta_i(x)\} = \sum_{i=0}^n f(p_i) \delta_i(x).$$

Además,

$$\delta_0(x) = 1 - \sum_{i=1}^n \delta_i(x).$$

Tomando \log_q a ambos lados de (3.6), y aplicando la linealidad señalada, tenemos que

$$\begin{aligned}
 \log_q p(x) &= \sum_{i=0}^n \frac{p_i^{1-q} - 1}{1-q} \delta_i(x) \\
 &= \frac{1}{1-q} \sum_{i=0}^n (p_i^{1-q} - 1) \delta_i(x) \\
 &= \frac{1}{1-q} \left[\sum_{i=1}^n (p_i^{1-q} - 1) \delta_i(x) + (p_0^{1-q} - 1) \delta_0(x) \right] \\
 &= \frac{1}{1-q} \left[\sum_{i=1}^n (p_i^{1-q} - 1) \delta_i(x) + (p_0^{1-q} - 1) \left(1 - \sum_{i=1}^n \delta_i(x) \right) \right] \\
 &= \frac{1}{1-q} \left[\sum_{i=1}^n p_i^{1-q} \delta_i(x) - \sum_{i=1}^n \delta_i(x) + (p_0^{1-q} - 1) - \sum_{i=1}^n (p_0^{1-q} - 1) \delta_i(x) \right] \\
 &= \frac{1}{1-q} \left[\sum_{i=1}^n p_i^{1-q} \delta_i(x) + (p_0^{1-q} - 1) - \sum_{i=1}^n p_0^{1-q} \delta_i(x) \right] \\
 &= \frac{1}{1-q} \left[\sum_{i=1}^n (p_i^{1-q} - p_0^{1-q}) \delta_i(x) + p_0^{1-q} - 1 \right].
 \end{aligned}$$

Así,

$$\begin{aligned}
 \log_q p(x) &= \frac{1}{1-q} \left[\sum_{i=1}^n (p_i^{1-q} - p_0^{1-q}) \delta_i(x) + p_0^{1-q} - 1 \right] \\
 &= \left[\sum_{i=1}^n \frac{1}{1-q} (p_i^{1-q} - p_0^{1-q}) \delta_i(x) + \frac{p_0^{1-q} - 1}{1-q} \right].
 \end{aligned}$$

Haciendo,

$$\theta^i = \frac{1}{1-q} (p_i^{1-q} - p_0^{1-q}), \quad i = 1, \dots, n, \quad (\text{parámetros } q\text{-canónicos})$$

$$x_i = \delta_i(x) \quad (\text{variables aleatorias})$$

$$\psi(\theta) = -\log_q p_0 \quad (\text{función } q\text{-potencial})$$

tenemos que, S_n es una familia q -exponencial (3.5) para cualquier q . \square

Se puede observar que el potencial $\psi(\theta)$ y el parámetro canónico θ dependen de q . Además, la variedad tiene estructura lineal con respecto a cualquier θ_q .

3.2. Estructura Riemanniana y q -divergencia

A continuación consideraremos la estructura q -geometrica de S . La q -log-verosimilitud es una forma lineal definida por

$$l_q(\mathbf{x}, \theta) = \log_q p(\mathbf{x}, \theta) = \sum_{i=1}^n \theta^i x_i - \psi(\theta) \quad (3.7)$$

Diferenciando con respecto a θ^i , con la anotación abreviada $\partial_i = \frac{\partial}{\partial \theta^i}$, tenemos

$$\partial_i l_q(\mathbf{x}, \theta) = x_i - \partial_i \psi(\theta) \quad (3.8)$$

$$\partial_i \partial_j l_q(\mathbf{x}, \theta) = -\partial_i \partial_j \psi(\theta) \quad (3.9)$$

El siguiente teorema señala la convexidad de la energía q -libre.

Teorema 3.2.1: La energía q -libre o q -potential $\psi_q(\theta)$ es una función convexa de θ_q .

Prueba: Omitimos el sufijo q por simplicidad. Tenemos

$$\partial_i p(\mathbf{x}, \theta) = p(\mathbf{x}, \theta)^q (x_i - \partial_i \psi) \quad (3.10)$$

$$\partial_i \partial_j p(\mathbf{x}, \theta) = qp(\mathbf{x}, \theta)^{2q-1} (x_i - \partial_i \psi)(x_j - \partial_j \psi) - p(\mathbf{x}, \theta)^q \partial_i \partial_j \psi \quad (3.11)$$

Las siguientes identidades son válidas:

$$\int \partial_i p(\mathbf{x}, \theta) d\mathbf{x} = \partial_i \int p(\mathbf{x}, \theta) d\mathbf{x} = 0 \quad (3.12)$$

$$\int \partial_i \partial_j p(\mathbf{x}, \theta) d\mathbf{x} = \partial_i \partial_j \int p(\mathbf{x}, \theta) d\mathbf{x} = 0 \quad (3.13)$$

Aquí, definimos un funcional importante

$$h_q(\theta) = h_q[p(\mathbf{x}, \theta)] = \int p(\mathbf{x}, \theta)^q d\mathbf{x} \quad (3.14)$$

en particular para S_n discreto,

$$h_q(\mathbf{p}) = \sum_{i=0}^n p_i^q \quad (3.15)$$

para $0 < q < 1$. Esta función juega un papel importante en lo siguiente. De (3.10) y (3.12), tenemos

$$\int p(\mathbf{x}, \theta)^q (x_i - \partial_i \psi) d\mathbf{x} = 0.$$

$$\int \partial_i \psi p(\mathbf{x}, \theta)^q d\mathbf{x} = \int x_i p(\mathbf{x}, \theta)^q d\mathbf{x}.$$

$$\partial_i \psi h_q(\theta) d\mathbf{x} = \int x_i p(\mathbf{x}, \theta)^q d\mathbf{x}.$$

$$\partial_i \psi(\theta) = \frac{1}{h_q(\theta)} \int x_i p(\mathbf{x}, \theta)^q d\mathbf{x} \quad (3.16)$$

De (3.11) y (3.13), se obtiene

$$\int qp(\mathbf{x}, \theta)^{2q-1} (x_i - \partial_i \psi)(x_j - \partial_j \psi) - p(\mathbf{x}, \theta)^q \partial_i \partial_j \psi d\mathbf{x} = 0$$

$$\partial_i \partial_j \psi \int p(\mathbf{x}, \theta)^q d\mathbf{x} = \int qp(\mathbf{x}, \theta)^{2q-1} (x_i - \partial_i \psi)(x_j - \partial_j \psi) d\mathbf{x}$$

$$\partial_i \partial_j \psi h_q(\theta) = q \int p(\mathbf{x}, \theta)^{2q-1} (x_i - \partial_i \psi)(x_j - \partial_j \psi) d\mathbf{x}$$

$$\partial_i \partial_j \psi(\theta) = \frac{q}{h_q(\theta)} \int (x_i - \partial_i \psi)(x_j - \partial_j \psi) p(\mathbf{x}, \theta)^{2q-1} d\mathbf{x} \quad (3.17)$$

Esto muestra que $\partial_i \partial_j \psi(\theta)$ está positivo-definido, y por lo tanto ψ es convexo. \square

Pasamos ahora a la estructura riemanniana. La matriz de información q -Fisher viene dada por la expresión $[g_{ij}^{(q)}(\theta)]$, donde $g_{ij}^{(q)} = \partial_i \partial_j \psi(\theta)$ es una matriz definida positiva. Como caso particular, cuando $q = 1$, se obtiene la matriz de información de Fisher estándar

$$g_{ij}^{(1)}(\theta) = g_{ij}^{(F)}(\theta) = E[\partial_i \log p(\mathbf{x}, \theta) \partial_j \log p(\mathbf{x}, \theta)] \quad (3.18)$$

La matriz definida positiva $g_{ij}^{(q)}(\theta)$ define una métrica Riemanniana sobre S_n , dándole estructura q -Riemanniana.

La transformación conformal de divergencia que induce la métrica Riemanniana viene dada por

$$\tilde{g}_{ij}(\theta) = \sigma(\theta) g_{ij}(\theta) \quad (3.19)$$

donde $\sigma(\theta)$ es una función positiva y $g_{ij}(\theta)$ es un tensor métrico.

Teorema 3.2.2: La métrica de información q -Fisher es dada por una transformación conformal de la métrica de información de Fisher $g_{ij}^{(F)}$ como

$$g_{ij}^{(q)}(\theta) = \frac{q}{h_q(\theta)} g_{ij}^{(F)}(\theta) \quad (3.20)$$

□

S es una variedad Riemanniana dada por la matriz de información de q -Fisher.

La ortogonalidad, o más generalmente el ángulo, de dos vectores X y Y no cambia por una transformación conformal, aunque sus magnitudes cambien.

De otro lado, una divergencia se utiliza para medir la discrepancia entre dos puntos. Entre los diversos tipos de divergencia, una de las más conocidas es la divergencia de Bregman, que a menudo se emplea en el proceso de la optimización y de la señal. Esta es la única clase de divergencias que tienen estructura geométrica doblemente plana.

Mediante el empleo de una función convexa $\psi(\theta)$ es posible definir una divergencia de este tipo entre dos distribuciones de probabilidad $p(\mathbf{x}, \theta_1)$ y $p(\mathbf{x}, \theta_2)$:

$$D_q[p(\mathbf{x}, \theta_1) : p(\mathbf{x}, \theta_2)] = \psi(\theta_2) - \psi(\theta_1) - \nabla\psi(\theta_1) \cdot (\theta_2 - \theta_1) \quad (3.21)$$

donde el gradiente, es dado por $\nabla\partial/\partial\theta$, y satisface la condición de no negatividad

$$D_q[p(\mathbf{x}, \theta_1) : p(\mathbf{x}, \theta_2)] \geq 0 \quad (3.22)$$

dandose la igualdad sólo cuando $\theta_1 = \theta_2$. Esto origina una q -divergencia en la familia de distribuciones de probabilidad discretas S_n , que se describe en el siguiente teorema.

Teorema 3.2.3: Para dos distribuciones discretas $p(x) = \mathbf{p}$ y $r(x) = \mathbf{r}$, la q -divergencia está dada por

$$D_q[\mathbf{p} : \mathbf{r}] = \frac{1}{(1-q)h_q(\mathbf{p})} \left(1 - \sum_{i=0}^n p_i^q r_i^{1-q}\right). \quad (3.23)$$

□.

3.3. Estructura dualmente plana de la variedad

A continuación mostraremos la relación existente entre las coordenadas duales y su correspondiente potencial dual, para ello se requiere definir el concepto de q -esperanza de la variable aleatoria $f(\mathbf{x})$ y la transformada de Legendre. También se muestra la estructura

dualmente plana de la variedad, lo cual se aprovechará para mostrar una versión del teorema de Pitágoras y su correspondiente teorema de la proyección.

La q -esperanza de la variable aleatoria $f(\mathbf{x})$ viene dada por

$$E_{\hat{p}}[f(\mathbf{x})] = \frac{1}{h_q[p(\mathbf{x})]} \int p(\mathbf{x})^q f(\mathbf{x}) d\mathbf{x}. \quad (3.24)$$

donde \hat{p} es la distribución de probabilidad q -escort, definida por

$$\hat{p}_q(\mathbf{x}) = \frac{1}{h_q[p(\mathbf{x})]} p(\mathbf{x})^q. \quad (3.25)$$

Dada una función convexa $\psi(\theta)$, la transformación de Legendre es definida por

$$\eta = \nabla\psi(\theta) \quad (3.26)$$

donde $\nabla = (\partial/\partial\theta^i)$ es el gradiente. Puesto que la correspondencia entre θ y η es uno a uno, nosotros podemos considerar η como otro sistema de coordenada de S .

La función potencial dual es definida por

$$\varphi(\eta) = \max_{\theta} \{\theta \cdot \eta - \psi(\theta)\} \quad (3.27)$$

que es convexa con respecto a η . Las coordenadas originales se recuperan de la transformación inversa dada por

$$\theta = \nabla\varphi(\eta) \quad (3.28)$$

donde $\nabla = (\partial/\partial\eta_i)$, así que θ y η están en correspondencia dual.

Teorema 3.3.1: Las coordenadas duales η son dadas por

$$\eta = E_{\hat{p}}[\mathbf{x}] \quad (3.29)$$

y el potencial dual es dado por

$$\varphi(\eta) = \frac{1}{1-q} \left\{ \frac{1}{h_q(\mathbf{p})} - 1 \right\} \quad (3.30)$$

Prueba: La relación (3.29) es inmediata de (3.16). De la dualidad de Legendre, el potencial dual satisface

$$\varphi(\eta) + \psi(\theta) - \theta \cdot \eta = 0 \quad (3.31)$$

cuando θ y η corresponden el uno al otro por $\eta = \nabla\psi(\theta)$. Por lo tanto,

$$\varphi(\eta) = \sum_{i=1}^n \theta^i \eta_i - \psi(\theta) \quad (3.32)$$

$$= E_{\hat{p}}[\log_q p(\mathbf{x}, \theta)] \quad (3.33)$$

$$= \frac{1}{(1-q)h_q(\theta)} (1 - \int p^q p(\mathbf{x}, \theta) d\mathbf{x}) \quad (3.34)$$

$$= \frac{1}{(1-q)} \left(\frac{1}{h_q(\theta)} - 1 \right) \quad (3.35)$$

Ésta es una función convexa de η . \square

Llamamos el potencial q -dual

$$\varphi(\eta) = E[\log_q p(\mathbf{x}, \theta)] = \frac{1}{1-q} \left\{ \frac{1}{h_q} - 1 \right\} \quad (3.36)$$

La q -entropy negativa, porque es el Legendre-dual de la energía q -libre $\psi(\theta)$. Hay varias definiciones de q -entropy. La q -entropy de Tsallis [3] es definida originalmente por

$$H_{Tsallis} = \frac{1}{1-q} (h_q - 1) \quad (3.37)$$

mientras que la q -entropy de Rényi [36] es

$$H_{Renyi} = \frac{1}{1-q} \log h_q \quad (3.38)$$

Éstas están mutuamente relacionadas por funciones monotonas. Cuando $q \rightarrow 1$, todas ellas se reducen a la entropía de Shannon.

Nuestra definición de

$$H_q = \frac{1}{1-q} \left(1 - \frac{1}{h_q} \right) = \frac{H_{Tsallis}}{h_q} \quad (3.39)$$

también es monotonamente conectada con la anterior, pero es más natural desde el punto de vista de q -geometría. La entropía H_q es conocida como la q -entropía normalizada, que se estudió en [37], [38].

Hay dos sistemas de coordenadas dualmente acoplados θ y η en la familia q -exponencial S con dos funciones potenciales $\psi(\theta)$ y $\varphi(\eta)$ para cada q . Dos estructuras afines son introducidas por las dos funciones convexas ψ y φ . Consideremos S como una variedad afín donde θ es un sistema de coordenadas afín. Ellos representan cantidades intensivas de un sistema físico. Dualmente, introducimos una estructura afín dual a S , donde η es otro sistema de coordenadas afín. Ellos representan cantidades extensas. Nosotros podemos definir dos tipos de líneas rectas o geodésicas en S debido a las estructuras q -afín.

Para dos distribuciones $p(\mathbf{x}, \theta_1)$ y $p(\mathbf{x}, \theta_2)$ en S , una curva $p(\mathbf{x}, \theta(t))$ se dice que es una q -geodésica conecta, cuando

$$\theta(t) = t\theta_1 + (1-t)\theta_2 \quad (3.40)$$

donde t es el parámetro de la curva. Dualmente, en términos de coordenadas duales η , cuando

$$\eta(t) = t\eta_1 + (1-t)\eta_2 \quad (3.41)$$

también se cumple, se dice que la curva es un dual q -geodésico.

Más generalmente, la q -geodésica conectando dos distribuciones $p_1(\mathbf{x})$ y $p_2(\mathbf{x})$ es dada por

$$\log_q p(\mathbf{x}, t) = t \log_q p_1(\mathbf{x}) + (1 - t) \log_q p_2(\mathbf{x}) - c(t) \quad (3.42)$$

donde $c(t)$ es un término normalizando. Esto se puede reescribir como

$$p(\mathbf{x}, t)^{1-q} = t p_1(\mathbf{x})^{1-q} + (1 - t) p_2(\mathbf{x})^{1-q} - c(t) \quad (3.43)$$

Dualmente, la q -geodésica dual conectando $p_1(\mathbf{x})$ y $p_2(\mathbf{x})$ es dada usando las distribuciones escort como

$$\hat{p}(\mathbf{x}, t) = t \hat{p}_1(\mathbf{x}) + (1 - t) \hat{p}_2(\mathbf{x}) \quad (3.44)$$

Puesto que la variedad S tiene una estructura q -Riemanniana, la ortogonalidad de dos vectores tangentes es definida por la métrica Riemanniana. Volvemos a escribir la ortogonalidad de dos geodesicas en términos de las coordenadas afines. Consideremos dos desviaciones pequeñas $d_1 p(\mathbf{x})$ y $d_2 p(\mathbf{x})$ de $p(\mathbf{x})$, esto es, desde $p(\mathbf{x})$ a $p(\mathbf{x}) + d_1 p(\mathbf{x})$ y $p(\mathbf{x}) + d_2 p(\mathbf{x})$, que se consideran como dos vectores tangentes (infinitesimal) de S a $p(\mathbf{x})$.

Lema 3.3.1: El producto interno de dos desviaciones $d_1 p$ y $d_2 p$ es dado por

$$\langle d_1 p(\mathbf{x}), d_2 p(\mathbf{x}) \rangle = \int d_1 \hat{p}(\mathbf{x}) d_2 \log_q p(\mathbf{x}) d\mathbf{x} \quad (68)$$

Prueba: por simples cálculos, tenemos

$$\int d_1 \hat{p}(\mathbf{x}) d_2 \log_q p(\mathbf{x}) d\mathbf{x} = \frac{q}{h_q} \int \frac{\int d_1 p(\mathbf{x}) d_2 p(\mathbf{x})}{p(\mathbf{x})} d\mathbf{x}$$

del cual el lado derecho es el último factor de (3.20) . \square

Corolario 3.3.1: Dos curvas $\theta_1(t)$ y $\eta_2(t)$, intersectandose en $t = 0$, son ortogonales cuando $\langle \dot{\theta}_1(0), \dot{\eta}_2(0) \rangle = 0$.

Aquí, $\dot{\theta}_1(t)$ y $\dot{\eta}_2(t)$ denotan derivadas de $\theta_1(t)$ y $\eta_2(t)$ en t , respectivamente.

Las dos geodesicas y la ortogonalidad juegan un papel fundamental en S como se verá en lo siguiente.

Una variedad de Riemann dualmente plana admite una generalización del teorema de Pitágoras y su correspondiente teorema de la proyección.

A continuación introduciremos una estructura dual en S definida por la transformación de Legendre [24]. El vector tangente

$$z^* = \text{Grad} k(z)$$

para una función convexa $k(z)$, el cual está en correspondencia uno a uno con z . Esta es la

transformación de Legendre, y z^* puede ser considerado como otro sistema de coordenadas de S diferente de z .

Se puede calcular la función dual de k definida por

$$k^*(z^*) = mx_z \{z \cdot z^* - k(z)\}$$

que es una función convexa de z^* . Por lo tanto, podemos describir la geometría usando la función convexa dual k^* y las coordenadas dual de z^* . Claramente, z y z^* son duales, puesto que

$$z = \text{Grad}k^*(z^*).$$

Lema 3.3.2: Las dos divergencias D y D^* son mutuamente recíprocas, en el sentido de

$$D^*[y^* : z^*] = D[z : y].$$

La divergencia entre dos puntos z y y está dada por

$$D[z : y] = k(z) + k^*(y^*) - z \cdot y^*.$$

Prueba: sea $k^*(z^*) = mx_z \{z \cdot z^* - k(z)\}$

El lado derecho se maximiza cuando z y z^* corresponden uno al otro, es decir,

$$\frac{\partial}{\partial z} \{z \cdot z^* - k(z)\} = z^* - \text{Grad}k(z) = 0.$$

Por lo tanto, se obtiene la identidad

$$k(z) + k^*(z^*) - z \cdot z^* = 0.$$

Mediante el uso de esta relación para y , tenemos

$$D[z : y] = k(z) - k(y) - y^* \cdot (z - y) = k(z) + k^*(y^*) - z \cdot y^*.$$

Análogamente, tenemos para $k(z)$

$$D[z : y] = D^*[y^* : z^*].$$

De lo anterior se concluye que, es suficiente con considerar una sólo función de divergencia. \square

Teorema 3.3.2 q -Pitágoras: Para tres distribuciones $p_1(x)$, $p_2(x)$ y $p_3(x)$ en S , se cumple que

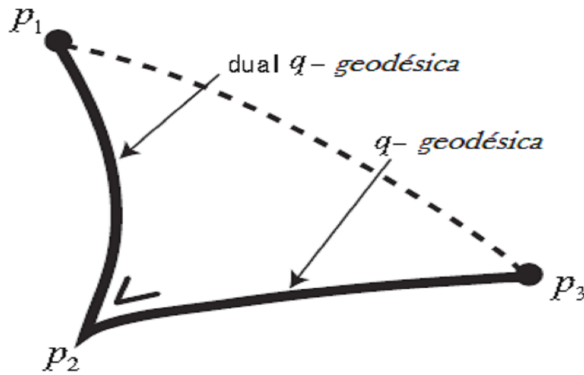
$$D_q[p_1 : p_2] + D_q[p_2 : p_3] = D_q[p_1 : p_3]$$

Prueba: por el lema 3.3.2. anterior se tiene que

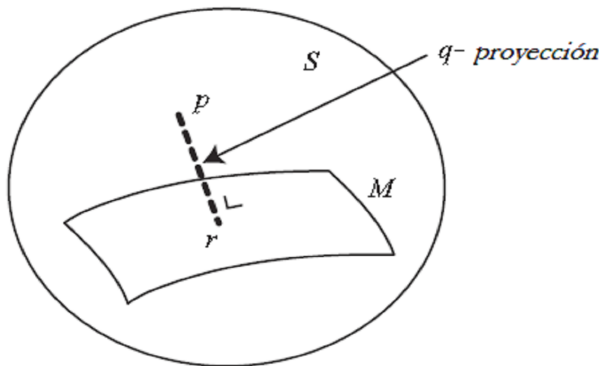
$$\begin{aligned} D[p_3 : p_2] + D[p_2 : p_1] &= k(z_{p_3}) + k^*(z_{p_2}^*) + k(z_{p_2}) + k^*(z_{p_1}^*) - z_{p_3} \cdot z_{p_2}^* - z_{p_2} \cdot z_{p_1}^* \\ &= k(z_{p_3}) + k^*(z_{p_1}^*) + z_{p_2} \cdot z_{p_2}^* - z_{p_3} \cdot z_{p_2}^* - z_{p_2} \cdot z_{p_1}^* \\ &= D[z_{p_3} : z_{p_1}^*] + (z_{p_2} - z_{p_3}) \cdot (z_{p_2}^* - z_{p_1}^*) \end{aligned}$$

El vector tangente de la geodésica que conecta a p_2 y p_3 es $z_{p_2} - z_{p_3}$ y el vector tangente de la geodésica dual que conecta a p_2 y p_1 es $z_{p_2}^* - z_{p_1}^*$ en el sistema dual de coordenadas. Por lo tanto, el segundo término del lado derecho de la anterior ecuación se anula, dado que las geodésicas prima y dual que conectan p_2 y p_3 , y p_2 y p_1 son ortogonales. \square

Cuando la geodésica dual conectando $p_1(x)$ y $p_2(x)$, es ortogonal a $p_2(x)$ hasta la conexión geodésica dada de $p_2(x)$ a $p_3(x)$, como lo muestra la figura.



Dada una distribución $p(x) \in S$ y una subvariedad $M \subset S$, una distribución de $r(x) \in M$ se dice que es el q -proyección (dual q -proyección) de $p(x)$ a M , cuando el q -geodesica (dual q -geodesica) que conecta $p(x)$ y $r(x)$ es ortogonal de M hasta $r(x)$. Ver figura.



Teorema 3.3.3 q -Proyección. Sea M una subvariedad de S . Dado $p(x) \in S$, el punto $r(x) \in M$ que minimiza $Dq[p(x) : r(x)]$ está dada por la doble q -proyección de $p(x)$ en M . El punto $r(x) \in M$ que minimiza $Dq[r(x) : p(x)]$ está dado por la q -proyección de $p(x)$ a M .

Mostramos que el bien conocido teorema de q -max-ent en el caso de q -entropía de Tsallis [1,4,9,11] es una consecuencia directa de los teoremas anteriores de q -Pitágoras y q -

proyección.

Teorema 3.3.4 q-Max-Ent. Las distribuciones de probabilidad que maximizan la H_{Tallis} q -entropía, $H_{Rényi}$ y H_q bajo q -lineales restricciones para m variables aleatorias $c_k(x)$ y distintos valores de a_k

$$E_{\hat{p}}[c_k(x)] = a_k, \quad k = 1, \dots, m$$

forman una familia q -exponencial

$$\log_q p(x, \theta) = \sum_{i=1}^m \theta^i c_i(x) - \psi(\theta) \quad (3.45)$$

La prueba es fácil de realizar usando algún método analítico convencional. Hasta aquí hemos dado una prueba geométrica.

Nos permitiremos considerar el subespacio $M^* \subset S$ cuyos miembros $p(x)$ satisface las m restricciones

$$E_{\hat{p}}[c_k(x)] = \int \hat{p}(x) c_k(x) dx = a_k \quad k = 1, \dots, m$$

Dado que las restricciones son lineales en las dobles coordenadas afines η o $\hat{p}(x)$, M^* es un subespacio lineal de S con respecto a la conexión doble afín. Sea $p_0(x, \theta_0)$ la distribución uniforme definida por $\theta_0 = 0$, lo que implica $p_0(x, \theta_0) = cte$ de (6). Sea $\hat{p}(x) \in M^*$ la q -proyección de $P_0(x)$ en M^* (figura siguiente). Entonces, la divergencia $D_q[p : p_0]$ de $p(x) \in M^*$ hasta $P_0(x)$ se descompone como

$$D_q[p : p_0] = D_q[p : \hat{p}] + D_q[\hat{p} : p_0]$$

Sea η_p la dual coordenada de $p(x)$. Puesto que la divergencia se escribe como

$$D_q[p : p_0] = \psi(\theta_0) + \varphi(\eta_p) - \theta_0 \cdot \eta_p$$

el reductor de $D_q[p : p_0]$ entre $p(x) \in M^*$ es $\hat{p}(x)$, que también es el maximizador de la entropía $-\varphi(\eta_p)$.

Las trayectorias de $\hat{p}(x)$ para distintos valores de k forman un subespacio plano ortogonal a M^* , lo que implica que formar una familia q -exponencial de la forma (6) (ver figura). Las direcciones tangenciales $d\hat{p}(x)$ de M^* satisfacen

$$\int d\hat{p}(x) c_k(x) dx = 0 \quad k = 1, \dots, m.$$

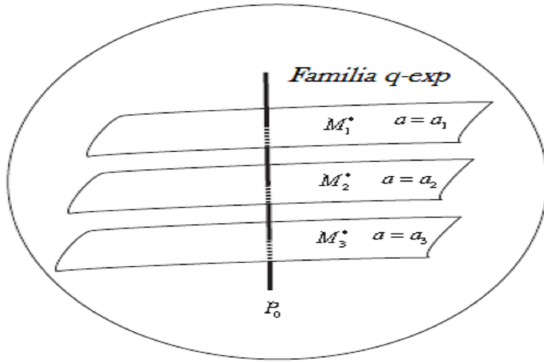
Por lo tanto, una familia q -exponencial de la forma

$$\log_q p(x, \xi) = \sum_{i=1}^m \xi_i d_i(x) - \psi(\xi)$$

es ortogonal a M^* , cuando

$$\int d\hat{p}(x) d \log_q p(x, \xi) dx = 0$$

Esto implica que $d_i(x) = c_i(x)$. Por lo tanto, tenemos la familia q -exponencial (3.45) que maximiza la q -entropía.



Conclusiones

Como consecuencia de las teorías de Tsallis se ha vuelto la mirada a las distribuciones de probabilidad sujetas a la ley de potencia. Sin embargo, se carece de una base geométrica, mientras que para las distribuciones comunes de probabilidad existe la geometría de la información. En esta sección hemos introducido un nuevo concepto de la q -geometría, con el objeto de suplir la necesidad antes señalada. La q -estructura se compone de un métrica riemanniana con un par de conexiones afines dualmente acoplados, que se encuentra en el marco de la geometría de información estándar. Sin embargo, la q -estructura es esencialmente diferente de la estándar derivada por el criterio de la invariancia de la variedad de distribuciones de probabilidad. Esta nueva mirada sobre la teoría relacionada con la q -entropía vista desde la transformación conformal nos conduce a definiciones unificadas de diversas cantidades tales como la q -entropía, q -divergencia, la función q -potencial y sus duales, así como nuevas interpretaciones de las cantidades conocidas. Esto constituye una base geométrica, y esperamos que el trabajo contribuya a proporcionar nuevos avances en este campo.

Capítulo 4

Variedad de información q -exponencial: caso no paramétrico

En este capítulo se expone la construcción de la variedad de información q -exponencial, así como la caracterización de la geometría que induce. Las contribuciones han sido establecidas en dos trabajos de G. Loaiza y H. R. Quiceno [2, 3], que introducen una variedad de Banach y una geometría riemanniana dependiendo de un parámetro real q inspirado en el índice de entropía según la mecánica estadística de Tsallis [15, 16] y que extienden como límite, cuando q tiende a uno, a la teoría formulada a partir de la variedad de información exponencial de Pistone y Sempi [26][34]; que se relaciona con la entropía de Shannon y la mecánica estadística clásica de Boltzman y Gibbs.

Concretamente, si μ es una medida de probabilidad dada y \mathfrak{M}_μ denota al conjunto de las densidades de probabilidad estrictamente positivas μ -equivalentes, se construye una variedad de Banach sobre \mathfrak{M}_μ para los modelos estadísticos q -exponenciales no paramétricos donde: $0 < q < 1$ es cualquier número real, cada modelo paramétrico q -exponencial se identifica con el espacio tangente de \mathfrak{M}_μ y los mapeos coordenados se definen naturalmente en términos de entropías relativas en el contexto de Tsallis (representadas como Φ -divergencias de Csiszár); tal como sucede con la construcción estándar de Pistone y Sempi respecto a la entropía relativa en el contexto de Shannon. La variedad construida por Pistone y Sempi, se recupera cuando q tiende a uno, salvo por inmersiones continuas en el espacio modelado.

Finalmente, mediante el funcional de q -divergancia se define una métrica que induce una geometría riemanniana que generaliza, por límite, a la geometría dada por la métrica de información de Fisher y conexiones Levi-Civita. Como resultado importante, se muestra la planeidad (curvatura cero) de la variedad, a partir de la geometría dada mediante el funcional de q -divergancia.

4.1. Funcional acumulante

Los mapeos de transición para la variedad de información q -exponencial, se definirán en términos del funcional acumulante K_p , respecto a cada $p \in \mathfrak{M}_\mu$, cuya definición y propiedades se presentan en esta sección, acorde a [2]. Ahora, para que la variedad de información q -exponencial sea una variedad de Banach, es necesario que sea diferenciable y en este caso, ello depende de la analiticidad del funcional acumulante K_p . Es por ello que en esta sección se presenta tanto la definición como las propiedades y la analiticidad de dicho funcional, allanando el camino para desarrollos presentados en secciones siguientes.

Antes de definir el funcional K_p , es necesario establecer algunos aspectos que conducen a definir bien dicho funcional y estudiar sus propiedades. Para ello se presenta a continuación algunos resultados de [2] realizados en el marco de la teoría de convergencia de series en espacios de Banach [33].

Se define una función analítica entre la bola unidad abierta $B_{p,\infty}(0,1)$ de $L^\infty(p \cdot \mu)$ y el espacio $L^\infty(p \cdot \mu)$ de la siguiente manera. Sea $0 < q < 1$, $p \in \mathfrak{M}_\mu$, $u \in B_{p,\infty}(0,1) \subset L^\infty(p \cdot \mu)$ y $n \in \mathbf{N}$. Se define el mapeo

$$\Lambda_n(u) : \prod_{i=1}^n L^\infty(p \cdot \mu) \rightarrow L^\infty(p \cdot \mu)$$

por

$$(v_1, \dots, v_n) \mapsto v_1 \dots v_n e_q^u. \quad (4.1)$$

Denotando por $\mathcal{L}_s^n(X, Y)$ al espacio de todos los mapeos n -multilineales continuos simétricos desde el espacio X en el espacio Y , que se usa en la siguiente proposición.

En repetidas ocasiones, usaremos el siguiente hecho. Si $\|u\|_{p,\infty} < 1$ entonces, $u < 1$ a.e sobre Ω y así

$$\|u\|_{p,\infty} < 1 \text{ implica que } \|e_q^u\|_{p,\infty} \leq e_q^{(1)} = (2 - q)^{\frac{1}{1-q}}. \quad (4.2)$$

Proposición 4.1.1: Para toda $u \in L^\infty(p \cdot \mu)$ con $\|u\|_{p,\infty} < 1$, se tiene que $\Lambda_n(u) \in \mathcal{L}_s^n(L^\infty(p \cdot \mu), L^\infty(p \cdot \mu))$.

Prueba: Si $\|u\|_{p,\infty} < 1$, es claro que $\Lambda_n(u)$ es un operador simétrico multilineal. Por (4.2), $\|e_q^u\|_{p,\infty} \leq (2 - q)^{\frac{1}{1-q}}$. Como $v_1 \dots v_n \in L^\infty(p \cdot \mu)$ se obtiene que $(v_1 \dots v_n e_q^u) \in L^\infty(p \cdot \mu)$ y

$$\|\Lambda_n(u)(v_1 \dots v_n)\|_{p,\infty} \leq (2 - q)^{\frac{1}{1-q}} \|v_1\|_{p,\infty} \dots \|v_n\|_{p,\infty}. \quad (4.3)$$

Como $(v_1 \dots v_n)$ es arbitrario en $\prod_{i=1}^n L^\infty(p \cdot \mu)$, entonces $\Lambda_n(u)$ está bien definida. De el criterio para operadores multilineales acotados, (4.3) implica que $\Lambda_n(u)$ es continuo. \square

Ahora para $\|u\|_{p,\infty} < 1$, escribiremos $\hat{\Lambda}_0(u) := e_q^u$. Como $\|u\|_{p,\infty} < 1$, por (4.2) se tiene que $e_q^{(u)} = \hat{\Lambda}_0(u) \in L^\infty(p \cdot \mu)$ y $\|\hat{\Lambda}_0(u)\|_{p,\infty} \leq e_q^1 = (2 - q)^{\frac{1}{1-q}} < \infty$. Para $n \geq 1$, sea

$\hat{\Lambda}_n(u)$ el polinomio n -homogéneo determinado por la forma polar $\Lambda_n(u)$, es decir, para toda $v \in L^\infty(p \cdot \mu)$, $\hat{\Lambda}_n(u) \cdot (v) = \Lambda_n(u)(v, v, \dots, v)$ n -veces, denotamos $\hat{\Lambda}_n(u) \cdot (v) = v^n e_q^u$.

Sea $A : L^\infty(p \cdot \mu) \rightarrow L^\infty(p \cdot \mu)$ tal que

$$v \rightarrow A(v) := 1 + \sum_{i=1}^{\infty} \frac{1}{i!} Q_{i-1}(q)(v)^i, \quad (4.4)$$

donde para cada $q \in (0, 1)$ denotamos $Q_i(q) = q(2q-1)(3q-2) \cdots (iq - (i-1)) = \prod_{n=1}^i (nq - (n-1))$ y $Q_0(q) = 1$.

Proposición 4.1.2: Si $p \in \mathfrak{M}_\mu$, la función A es una serie de potencia de $L^\infty(p \cdot \mu)$ en $L^\infty(p \cdot \mu)$, con radio de convergencia $\hat{\rho} > 1$.

Prueba: Note que

$$A(v) := 1 + [v + \frac{1}{2!} Q_1(q)v^2 + \frac{1}{3!} Q_2(q)v^3 + \cdots] = 1 + \sum_{i=0}^{\infty} \frac{1}{i!} Q_i(q) \hat{\Lambda}_i(0) \cdot (v).$$

Ahora, la función $A = \sum_{i=0}^{\infty} \frac{1}{i!} Q_i(q) \hat{\Lambda}_i(0)$ es una serie de potencia de $L^\infty(p \cdot \mu)$ en $L^\infty(p \cdot \mu)$ para la cual se discute el radio de convergencia a continuación. Si $v \in L^\infty(p \cdot \mu)$ con $\|v\|_{p,\infty} \leq 1$ y fijo $u \in B_{p,\infty}(0, 1)$, vemos que

$$\|\frac{1}{i!} Q_i(q) \hat{\Lambda}_i(0) \cdot (v)\|_{p,\infty} = \|\frac{1}{i!} Q_i(q)(v)^i e_q^u\|_{p,\infty} \leq \frac{1}{i!} Q_i(q)(2-q)^{\frac{1}{1-q}}.$$

Sea $\mathcal{P}^n : \mathcal{P}^n(L^\infty(p \cdot \mu), L^\infty(p \cdot \mu))$ el espacio lineal normado de polinomios n -homogéneos continuos determinado por la forma polar $\Lambda_n(u)$ como en la definición (4.1). Entonces $\|\hat{\Lambda}\|_{\mathcal{P}^n} = \sup_{\|v\|_{p,\infty}=1} \|\hat{\Lambda}(v)\|_{p,\infty} \leq (2-q)^{\frac{1}{1-q}}$ y

$$\|\frac{1}{i!} Q_i(q) \hat{\Lambda}_{a,i}(0)\|_{\mathcal{P}^n} \leq \frac{1}{i!} Q_i(q)(2-q)^{\frac{1}{1-q}} \quad (4.5)$$

Por la fórmula de Cauchy-Hadamard [33], si $\hat{\rho}$ es el radio de convergencia de la serie, se sigue que

$$\frac{1}{\hat{\rho}} = \overline{\lim}(\|\hat{\Lambda}\|_{\mathcal{P}^n})^{\frac{1}{n}} = \overline{\lim}(\|\Lambda\|_{\mathcal{L}_s^n})^{\frac{1}{n}} \leq \lim_{n \rightarrow \infty} (2-q)^{\frac{1}{n(1-q)}} = 1,$$

y así $\hat{\rho} > 1$. Entonces la serie (4.4) converge absolutamente y uniformemente sobre la bola cerrada de radio r , $\overline{B}_{p,\infty}(0, r)$ para cada $r < \hat{\rho}$. \square

Antes de demostrar el siguiente teorema, veamos que la serie (4.4) es igual a $e_q^{(v)}$. Por [16], se sabe que $1 + \sum_{i=0}^{\infty} \frac{1}{i!} Q_i(q)(v)^i = e_q^{(v)}$ puntualmente. Como $p \cdot \mu$ es una medida finita,

$1 + \sum_{i=0}^k \frac{1}{i!} Q_i(q)(v)^i \rightarrow e_q^{(v)}$ en $(p \cdot \mu)$ -medible. La proposición anterior prueba la convergencia $1 + \sum_{i=0}^k \frac{1}{i!} Q_i(q)(v)^i \rightarrow 1 + \sum_{i=0}^{\infty} \frac{1}{i!} Q_i(q)(v)^i$ en $L^\infty(p \cdot \mu)$ -norma y en consecuencia también converge en media. Por la unicidad se sigue que $1 + \sum_{i=0}^{\infty} \frac{1}{i!} Q_i(q)(v)^i = e_q^{(v)}$.

Proposición 4.1.3: Sea $p \in \mathfrak{M}_\mu$ y $v \in B_{p,\infty}(0, 1)$. La serie $v \rightarrow 1 + \sum_{i=0}^{\infty} \frac{1}{i!} Q_i(q)(v)^i$ es una función analítica.

Prueba: Dado que $\Lambda_n(u) \in \mathcal{L}_s^n(L^\infty(p \cdot \mu), L^\infty(p \cdot \mu))$, basta con demostrar la analiticidad en una vecindad del cero. Se tiene la siguiente desigualdad $\|\hat{\lambda}\|_{\mathcal{P}^n} \leq \|\lambda\|_{\mathcal{L}_s^n} \leq \frac{n^n}{n!} \|\hat{\lambda}\|_{\mathcal{P}^n}$, [33]. Por la anterior desigualdad y (4.5), tenemos

$$\left\| \frac{1}{n!} Q_n(q) \Lambda_n(0) \right\|_{\mathcal{L}_s^n} \leq \frac{n^n Q_n(q) (2-q)^{\frac{1}{1-q}}}{(n!)^2}.$$

Aplicando la fórmula de Cauchy-Hadamard y dado que $Q_n(q) \leq 1$, obtenemos

$$\frac{1}{\hat{\rho}} = \lim_{n \rightarrow \infty} \left(\left\| \frac{1}{n!} \frac{Q_n(q)}{(1-q)^n} \Lambda_n(0) \right\|_{\mathcal{L}_s^n} \right)^{\frac{1}{n}} \leq \lim_{n \rightarrow \infty} \left(\frac{n^n Q_n(q) (2-q)^{\frac{1}{1-q}}}{(n!)^2} \right)^{\frac{1}{n}} \leq e.$$

Luego, $\hat{\rho} \geq \frac{1}{e}$ y así el radio de convergencia restringido de la serie es positivo y mayor o igual a $\frac{1}{e}$. Aplicando el procedimiento anterior al polinomio n -homogéneo $\frac{1}{n!} Q_n(q) \hat{\Lambda}_n(v_0)$, donde $v_0 \in B_{p,\infty}(0, 1)$, el radio de convergencia restringido satisface $0 < \frac{(1-\|v_0\|_{p,\infty})}{e} \leq \hat{\rho}$ y así en una vecindad de v_0 el mapeo

$$v \rightarrow 1 + \sum_{n=0}^{\infty} \frac{1}{n!} \frac{Q_n(q)}{(1-q)^n} \hat{\Lambda}_n(v_0)(v)$$

es analítico. \square

Definición 4.1.1: Denotando $D_{M_p} := \{u \in L^\infty(p \cdot \mu) : \frac{-1}{1-q} < u, E_p[e_q^{(u)}] < \infty\}$, definimos el mapeo $M_p : L^\infty(p \cdot \mu) \rightarrow [0, \infty]$ por $M_p(u) = E_p[e_q^{(u)}]$.

El dominio D_{M_p} de M_p contiene la bola unidad abierta $B_{p,\infty}(0, 1) \subset L^\infty(p \cdot \mu)$. También si restringimos su dominio a $B_{p,\infty}(0, 1)$, veremos que esta función es analítica e infinitamente Fréchet diferenciable.

Teorema 4.1.1: El funcional M_p satisface:

1. $M_p(0) = 1$ y si $u \neq 0$, $M_p(u) > 1$.
2. M_p es analítica sobre la bola unidad abierta $B_{p,\infty}(0, 1)$.
3. M_p es infinitamente Fréchet diferenciable y su n -ésima derivada en $u \in B_{p,\infty}(0, 1)$ evaluada en $(v_1, v_2, \dots, v_n) \in B_{p,\infty}(0, 1) \times \dots \times B_{p,\infty}(0, 1)$ está dada por

$$D^n M_p(u)(v_1, v_2, \dots, v_n) = Q_n(q) E_p[v_1 \dots v_n e_q^{(u)}].$$

Prueba: (1) Es inmediato de la definición.

(2) Para $u \in B_{p,\infty}(0, 1)$, $\frac{1}{n!} Q_n(q) E_p[\hat{\Lambda}_n(u)] \in \mathcal{P}^n(L^\infty(p \cdot \mu), \mathfrak{R})$ y así $\sum_{n=0}^{\infty} \frac{1}{n!} Q_n(q) E_p[\hat{\Lambda}_n(u)]$ es una serie de potencias de $L^\infty(p \cdot \mu)$ en \mathfrak{R} con radio de convergencia positivo. Para cada u en una vecindad de $u_0 \in B_{p,\infty}(0, 1)$, y por la proposición 4.1.3 tenemos

$$e_q^{(u)} = 1 + \sum_{n=0}^{\infty} \frac{1}{n!} \frac{Q_n(q)}{(1-q)^n} \Lambda_n(u_0)(u - u_0)^n,$$

de esto, obtenemos

$$M_p(u) = \sum_{n=0}^{\infty} \frac{1}{n!} \frac{Q_n(q)}{(1-q)^n} E_p[\hat{\Lambda}(u_0)(u - u_0)].$$

En consecuencia, M_p es analítica en $B_{p,\infty}(0, 1)$.

(3) Se sabe que $D^n M_p(u)(v_1, v_2, \dots, v_n) = n! \lambda_n(u)(v_1, v_2, \dots, v_n)$ [33], donde $\lambda_n(u)(v_1, v_2, \dots, v_n) = \frac{1}{n!} Q_n(q) E_p[\Lambda_n(u)(v_1, v_2, \dots, v_n)]$ y así

$$\begin{aligned} D^n M_p(u)(v_1, v_2, \dots, v_n) &= Q_n(q) E_p[\Lambda_n(u)(v_1, v_2, \dots, v_n)] \\ &= Q_n(q) E_p[v_1 \dots v_n e_q^{(u)}]. \end{aligned}$$

□

A continuación se definirá el funcional K_p .

Definición 4.1.2: Para cada $u \in B_{p,\infty}(0, 1)$, definimos $K_p : B_{p,\infty}(0, 1) \rightarrow [0, \infty]$ por

$$K_p(u) = \ln_q[M_p(u)].$$

Como se ha dicho antes, este funcional es importante para definir mapeos de coordenadas para la variedad de información q -exponencial. Algunas de sus propiedades permiten obtener expresiones explícitas para el espacio tangente de la variedad. Note que $K_p(0) = 1$ y está claro, por propiedades logarítmicas q -deformadas, que si $u \neq 0$, entonces K_p es estrictamente positiva, continua, creciente y cóncava.

Teorema 4.1.2: El funcional K_p satisface:

1. Para cada $u \in B_{p,\infty}(0, 1)$, la función $z = e_q^{u \ominus_q K_p(u)} p$ es una densidad de probabilidad sobre \mathfrak{M}_μ .
2. K_p es infinitamente Fréchet diferenciable y su n -derivada evaluada en u en la dirección $(v_1, \dots, v_n) \in B_{p,\infty}(0, 1) \times \dots \times B_{p,\infty}(0, 1)$, (i.e. la forma n -lineal continua $D^n K_p(u) \cdot (v_1, \dots, v_n)$) está dada por

$$D^n K_p(u) \cdot (v_1, \dots, v_n) = [M_p(u)]^{1-q} Q_n(q) E_z[v_1, \dots, v_n].$$

3. K_p es analítica en $B_{p,\infty}(0, 1)$.

Prueba: 1) Como $p \in \mathfrak{M}_\mu$, se sigue que $z > 0$. Ahora, dado que $z = \frac{e_q^u}{e_q^{K_p(u)}} p$ y $e_q^{K_p(u)} = M_p(u)$, tenemos que $\int z d\mu = \frac{1}{e_q^{K_p(u)}} \int e_q^u p d\mu = 1$.

- 2) Tomando derivadas de K_p (en vista del teorema 4.1.1 y proposición 1.4.6), se sigue

$$\begin{aligned} D^n K_p(u) \cdot (v_1, \dots, v_n) &= \frac{1}{[M_p(u)]^q} D^n M_p(u) \cdot (v_1, \dots, v_n) \\ &= \frac{1}{[M_p(u)]^q} Q_n(q) E_p[v_1, \dots, v_n e_q^u]. \end{aligned}$$

Puesto que $z = e_q^{u \ominus_q K_p(u)} p$, entonces $e_q^u \cdot p = z \cdot M_p(u)$ y por tanto

$$D^n K_p(u) \cdot (v_1, \dots, v_n) = \frac{1}{[M_p(u)]^q} Q_n(q) M_p(u) E_z[v_1, \dots, v_n].$$

- 3) K_p es composición de funciones analíticas. \square

Teorema 4.1.3: Si $v \in B_p$ entonces $DK_p(u) \cdot v = E_q[v]$.

Prueba:

Sea $p \in \mathfrak{M}_\mu$ y consideremos $q = e^{u - K_p(u)} p$. Aplicando el operador $E_q[\cdot]$ a la serie $e^{tv} = \sum_{k=0}^{\infty} \frac{1}{k!} t^k v^k$, obtenemos $E_q[e^{tv}] = \sum_{k=0}^{\infty} \frac{1}{k!} E_q[v^k] t^k$.

Derivando respecto a t , vemos que

$$\frac{d}{dt} E_q[e^{tv}] = \sum_{k=1}^{\infty} \frac{1}{(k-1)!} E_q[v^k] t^{k-1}.$$

Cuando $t = 0$, se tiene que $\frac{d}{dt} E_q[e^{tv}]|_{t=0} = E_q[v]$. Utilizando, el siguiente resultado:

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

para $n = 1$, obtenemos

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

Esto concluye la prueba [10]. \square

4.2. Definición de la variedad

En esta sección se define un modelo geométrico similar al modelo k -exponencial de Pistone [35], donde los mapeos coordenados para la variedad inducen una topología más fuerte que la topología de $L^1(p \cdot \mu)$. Al final, se construye el fibrado tangente para dicha variedad, mostrando su identificación natural con modelos q -exponencial paramétrico unidimensional. También, se demuestra la relación entre el vector tangente de la variedad y una q -deformación de la función score.

Sea (Ω, Σ, μ) un espacio de probabilidad y q un número real con $0 < q < 1$. Para cada $p \in \mathfrak{M}_\mu$ denotamos por \mathcal{V}_p la bola unitaria abierta en el espacio de Banach que contiene las variables aleatorias centradas esencialmente acotadas B_p , esto es,

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

Definimos los mapeos $e_{q,p} : \mathcal{V}_p \rightarrow \mathfrak{M}_\mu$ por

$$e_{q,p}(u) := e_q^{(u \ominus_q K_p(u))} p, \quad (4.6)$$

que está bien definido dado que $\|u\|_{p,\infty} < 1$ implica que $\frac{-1}{1-q} < u$ y por lo tanto

$$\frac{-1}{1-q} < \frac{u - K_p(u)}{1 + (1-q)K_p(u)} = u \ominus_q K_p(u).$$

Proposición 4.2.1: Para cada $p \in \mathfrak{M}_\mu$, el mapeo $e_{q,p}$ es inyectivo.

Prueba: Sea $p \in \mathfrak{M}_\mu$ y $u_1, u_2 \in \mathcal{V}_p$ tal que $e_{q,p}(u_1) = e_{q,p}(u_2)$; es decir

$$\frac{u_1 - K_p(u_1)}{1 + (1-q)K_p(u_1)} = \frac{u_2 - K_p(u_2)}{1 + (1-q)K_p(u_2)}. \quad (4.7)$$

Aplicando el operador lineal $E_p[\cdot]$ a ambos lados de (4.7) se obtiene

$$\frac{E_p[u_1] - K_p(u_1)}{1 + (1-q)K_p(u_1)} = \frac{E_p[u_2] - K_p(u_2)}{1 + (1-q)K_p(u_2)}.$$

Ahora bien, como $u_1, u_2 \in \mathcal{V}_p \subset B_p$, se tiene que $E_p[u_1] = E_p[u_2] = 0$ y la expresión anterior se transforma en

$$-K_p(u_1) (1 + (1 - q)K_p(u_2)) = -K_p(u_2) (1 + (1 - q)K_p(u_1)).$$

es decir,

$$-K_p(u_1) - (1 - q)K_p(u_1)K_p(u_2) = -K_p(u_2) - (1 - q)K_p(u_1)K_p(u_2).$$

Así pues, $K_p(u_1) = K_p(u_2)$ y finalmente, sustituyendo en (4.7), se sigue que $u_1 = u_2$. \square

De aquí en adelante, el rango de $e_{q,p}$ se denotará por \mathcal{U}_p , esto es, $\mathcal{U}_p := e_{q,p}(\mathcal{V}_p)$. Para cada $p \in \mathfrak{M}_\mu$ el mapeo $s_{q,p} : \mathcal{U}_p \rightarrow \mathcal{V}_p$ se define por

$$s_{q,p}(z) := \ln_q\left(\frac{z}{p}\right) \ominus_q E_p[\ln_q\left(\frac{z}{p}\right)] = \frac{\ln_q\left(\frac{z}{p}\right) - E_p[\ln_q\left(\frac{z}{p}\right)]}{1 + (1 - q)E_p[\ln_q\left(\frac{z}{p}\right)]} \quad (4.8)$$

Estos mapeos se usarán para definir cartas, dado que la proposición anterior demuestra que estos mapeos son inyectivos.

Proposición 4.2.2: Para cada $p \in \mathfrak{M}_\mu$, el mapeo $s_{q,p}$ es la función inversa de $e_{q,p}$.

Prueba: Sean $p, q \in M_\mu$. Veamos que $u \in \mathcal{V}_p$ implica $s_{q,p}[e_{q,p}(u)] = u$. En efecto, de las expresiones (4.6) y (4.8), se sigue que

$$\begin{aligned} s_{q,p}[e_{q,p}(u)] &= s_{q,p} \left[e_q^{(u \ominus_q K_p(u))} p \right] \\ &= s_{q,p} \left[\exp_q \left(\frac{u - K_p(u)}{1 + (1 - q)K_p(u)} \right) p \right] \\ &= \frac{\left(\frac{u - K_p(u)}{1 + (1 - q)K_p(u)} \right) - E_p \left(\frac{u - K_p(u)}{1 + (1 - q)K_p(u)} \right)}{1 + (1 - q)E_p \left(\frac{u - K_p(u)}{1 + (1 - q)K_p(u)} \right)} \\ &= \frac{u - K_p(u) - E_p[u - K_p(u)]}{1 + (1 - q)K_p(u) + (1 - q)E_p[u - K_p(u)]} \\ &= \frac{u - K_p(u) - E_p(u) + E_p(K_p(u))}{1 + (1 - q)K_p(u) + (1 - q)E_p(u) - (1 - q)E_p(K_p(u))} \\ &= u. \end{aligned}$$

dado que $K_p(u)$ es una constante real y $E_p[u] = 0$. De manera análoga, probamos que $z \in \mathcal{U}_p$ implica $s_{q,p}[e_{q,p}(z)] = z$, y se obtiene el resultado deseado. \square

Los mapeos $s_{q,p}$ serán los mapeos de coordenadas para la variedad, estos mapeos asignan una única variable aleatoria $u = s_{q,p}(z) \in \mathcal{V}_p$, para un $z \in \mathcal{U}_p$. Demostraremos que la familia de pares $(\mathcal{U}_p, s_{q,p})_{p \in \mathfrak{M}_\mu}$ define un atlas sobre \mathfrak{M}_μ . Considere $p_1, p_2 \in \mathfrak{M}_\mu$ con

$\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2} \neq \emptyset$. Por cálculo directo tenemos que los mapeos $s_{p_2} \circ e_{p_1} : s_{q,p_1}(\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2}) \rightarrow s_{q,p_2}(\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2})$ están dados para cada $u \in s_{q,p_1}(\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2})$ por

$$s_{p_2}(e_{p_1}(u)) = \frac{u + [1 + (1 - q)u] \ln_q\left(\frac{p_1}{p_2}\right) - E_{p_2} \left[u + [1 + (1 - q)u] \ln_q\left(\frac{p_1}{p_2}\right) \right]}{1 + (1 - q)E_{p_2} \left[u + [1 + (1 - q)u] \ln_q\left(\frac{p_1}{p_2}\right) \right]} \quad (4.9)$$

En efecto,

$$\begin{aligned} s_{p_2}(e_{p_1}(u)) &= \frac{\ln_q\left(\frac{e_{p_1} p_1}{p_2}\right) - E_{p_2} \left[\ln_q\left(\frac{e_{p_1} p_1}{p_2}\right) \right]}{1 + (1 - q)E_{p_2} \left[\ln_q\left(\frac{e_{p_1} p_1}{p_2}\right) \right]} \\ &= \frac{\ln_q \left(e_q^{(u \ominus_q K_{p_1})} \frac{p_1}{p_2} \right) - E_{p_2} \left[\ln_q \left(e_q^{(u \ominus_q K_{p_1})} \frac{p_1}{p_2} \right) \right]}{1 + (1 - q)E_{p_2} \left[\ln_q \left(e_q^{(u \ominus_q K_{p_1})} \frac{p_1}{p_2} \right) \right]} \\ &= \frac{\frac{u - K_{p_1}(u)}{1 + (1 - q)K_{p_1}(u)} + \ln_q\left(\frac{p_1}{p_2}\right) + (1 - q) \frac{u - K_{p_1}(u)}{1 + (1 - q)K_{p_1}(u)} \ln_q\left(\frac{p_1}{p_2}\right) - E_{p_2} \left[\ln_q \left(e_q^{(u \ominus_q K_{p_1})} \frac{p_1}{p_2} \right) \right]}{1 + (1 - q)E_{p_2} \left[\ln_q \left(e_q^{(u \ominus_q K_{p_1})} \frac{p_1}{p_2} \right) \right]} \\ &= \frac{\frac{u - K_{p_1}(u) + [1 + (1 - q)K_{p_1}(u)] \ln_q\left(\frac{p_1}{p_2}\right) + (1 - q)[u - K_{p_1}(u)] \ln_q\left(\frac{p_1}{p_2}\right) - E_{p_2} \left[\ln_q \left(e_q^{(u \ominus_q K_{p_1})} \frac{p_1}{p_2} \right) \right]}{1 + (1 - q)K_{p_1}(u)}}{1 + (1 - q)E_{p_2} \left[\ln_q \left(e_q^{(u \ominus_q K_{p_1})} \frac{p_1}{p_2} \right) \right]} \\ &= \frac{\frac{u - K_{p_1}(u) + \ln_q\left(\frac{p_1}{p_2}\right) + (1 - q)K_{p_1}(u) \ln_q\left(\frac{p_1}{p_2}\right) + (1 - q)u \ln_q\left(\frac{p_1}{p_2}\right) - (1 - q)K_{p_1}(u) \ln_q\left(\frac{p_1}{p_2}\right) - E_{p_2} \left[\ln_q \left(e_q^{(u \ominus_q K_{p_1})} \frac{p_1}{p_2} \right) \right]}{1 + (1 - q)K_{p_1}(u)}}{1 + (1 - q)E_{p_2} \left[\ln_q \left(e_q^{(u \ominus_q K_{p_1})} \frac{p_1}{p_2} \right) \right]} \\ &= \frac{\frac{u - K_{p_1}(u) + [1 + (1 - q)u] \ln_q\left(\frac{p_1}{p_2}\right) - E_{p_2} \left[u - K_{p_1}(u) + [1 + (1 - q)u] \ln_q\left(\frac{p_1}{p_2}\right) \right]}{1 + (1 - q)K_{p_1}(u)}}{1 + (1 - q)E_{p_2} \left[\frac{u - K_{p_1}(u) + [1 + (1 - q)u] \ln_q\left(\frac{p_1}{p_2}\right)}{1 + (1 - q)K_{p_1}(u)} \right]} \\ &= \frac{u - K_{p_1}(u) + [1 + (1 - q)u] \ln_q\left(\frac{p_1}{p_2}\right) + E_{p_2}(K_{p_1}(u)) - E_{p_2} \left[u + [1 + (1 - q)u] \ln_q\left(\frac{p_1}{p_2}\right) \right]}{1 + (1 - q)K_{p_1}(u) + (1 - q)E_{p_2} \left[u - K_{p_1}(u) + [1 + (1 - q)u] \ln_q\left(\frac{p_1}{p_2}\right) \right]} \\ &= \frac{u + [1 + (1 - q)u] \ln_q\left(\frac{p_1}{p_2}\right) - E_{p_2} \left[u + [1 + (1 - q)u] \ln_q\left(\frac{p_1}{p_2}\right) \right]}{1 + (1 - q)K_{p_1}(u) + (1 - q)E_{p_2}(u) - (1 - q)E_{p_2}(K_{p_1}(u)) + (1 - q)E_{p_2} \left[[1 + (1 - q)u] \ln_q\left(\frac{p_1}{p_2}\right) \right]} \\ &= \frac{u + [1 + (1 - q)u] \ln_q\left(\frac{p_1}{p_2}\right) - E_{p_2} \left[u + [1 + (1 - q)u] \ln_q\left(\frac{p_1}{p_2}\right) \right]}{1 + (1 - q)E_{p_2} \left[u + [1 + (1 - q)u] \ln_q\left(\frac{p_1}{p_2}\right) \right]} \end{aligned}$$

Estos mapeos serán las funciones de transición para la variedad.

Proposición 4.2.3: Sean $p_1, p_2 \in \mathfrak{M}_\mu$. El conjunto $s_{q,p_1}(\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2})$ es abierto en la B_{p_1} -topología.

Prueba: Sea $u \in s_{q,p_1}(\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2})$. Veremos que u es un punto interior de $s_{q,p_1}(\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2})$. Claramente, $\|u\|_{p_1, \infty} < 1$ puesto que $s_{q,p_1}(\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2}) \subset \mathcal{V}_{p_1}$. Para $r < 1 - \|u\|_{p_1, \infty}$, consideremos la bola $A_r = \{v \in \mathcal{V}_{p_1} : \|v - u\|_{p_1, \infty} < r\}$. Dado que $p_1 \sim_q p_2$, por el lema 1.1.2. obtenemos $\|v - u\|_{p_2, \infty} = \|v - u\|_{p_1, \infty}$, para $v \in A_r$, y además $\|u\|_{p_1, \infty} = \|u\|_{p_2, \infty}$; de aquí que $\|v\|_{p_2, \infty} - \|u\|_{p_2, \infty} \leq \|v - u\|_{p_2, \infty} = \|v - u\|_{p_1, \infty} < r$. Luego, $\|v\|_{p_2, \infty} - \|u\|_{p_1, \infty} < r$ y $\|v\|_{p_2, \infty} < r + \|u\|_{p_1, \infty} < 1$. Esto implica $A_r \subset s_{q,p_1}(\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2})$. \square

Este resultado y el lema 1.1.2. permiten demostrar el siguiente resultado.

Proposición 4.2.4: Para cada $p_1, p_2 \in \mathfrak{M}_\mu$, la función $s_{q,p_2} \circ e_{q,p_1}$, es un homeomorfismo topológico.

Prueba: Dado que $s_{q,p_2} \circ e_{q,p_1} : s_{q,p_1}(\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2}) \rightarrow s_{q,p_2}(\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2})$. Por la proposición 4.2.2, se tiene que $s_{q,p_2} \circ e_{q,p_1}$ es inyectiva. Por la proposición 4.2.3 y dado que $s_{q,p}(\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2})$ es un abierto, se sigue que $s_{q,p_2} \circ e_{q,p_1}$ es sobreyectiva y continua y además, $(s_{q,p_2} \circ e_{q,p_1})^{-1} = s_{q,p_1} \circ e_{q,p_2}$ es continua. \square

Los mapeos dados en (4.9) pueden ser escritos en términos de las operaciones q -deformadas de la siguiente manera. Sea $u \in \mathcal{V}_{p_1}$, $\bar{u} \in \mathcal{V}_{p_2}$ y $z \in (\mathcal{U}_{q,p_1} \cap \mathcal{U}_{q,p_2})$ con $s_{q,p_2} \circ e_{q,p_1}(z) = \bar{u}$. Entonces

$$\begin{aligned} \bar{u} &= \ln_q \left(e_q^{u \ominus_q K_{p_1}(u)} \frac{p_1}{p_2} \right) \ominus_q E_{p_2} \left[\ln_q \left(e_q^{u \ominus_q K_{p_1}(u)} \frac{p_1}{p_2} \right) \right] \\ &= (u \ominus_q K_{p_1}(u)) \oplus_q \ln_q \left(\frac{p_1}{p_2} \right) \ominus_q E_{p_2} \left[(u \ominus_q K_{p_1}(u)) \oplus_q \ln_q \left(\frac{p_1}{p_2} \right) \right] \\ &= \left[(u \ominus_q K_{p_1}(u)) \oplus_q \ln_q \left(\frac{p_1}{p_2} \right) \right] \ominus_q E_{p_2} \left[(u \ominus_q K_{p_1}(u)) \oplus_q \ln_q \left(\frac{p_1}{p_2} \right) \right] \end{aligned}$$

por lo tanto,

$$\bar{u} = f(u) \ominus_q E_{p_2}[f(u)], \quad (4.10)$$

donde

$$f(u) = [(u \ominus_q K_{p_1}(u)) \oplus_q \ln_q \left(\frac{p_1}{p_2} \right)]. \quad (4.11)$$

Las expresiones (4.10) y (4.11) permiten simplificar cálculos para probar la proposición siguiente.

Proposición 4.2.5: Dado $u \in s_{q,p_1}(\mathcal{U}_{q,p_1} \cap \mathcal{U}_{q,p_2})$, la derivada del mapeo $s_{q,p_2} \circ s_{q,p_1}^{-1}$ evaluado en u en la dirección de $v \in L^\infty(p_1 \cdot \mu)$ es de la forma

$$D(s_{q,p_2} \circ s_{q,p_1}^{-1})(u) \cdot v = A(u) - B(u)E_{p_2}[A(u)],$$

donde $A(u), B(u)$ son constantes que dependen de u .

Prueba: Sabemos que $\bar{u} = s_{q,p_2} \circ s_{q,p_1}^{-1}(u) = \frac{f(u) - E_{p_2}[f(u)]}{1 + (1-q)E_{p_2}[f(u)]}$ y por consiguiente

$$D(s_{q,p_2} \circ s_{q,p_1}^{-1})(u) \cdot v = \frac{[1 + (1-q)E_{p_2}[f(u)]] \frac{\partial f(u)}{\partial u} - [1 + (1-q)f(u)] \frac{\partial E_{p_2}[f(u)]}{\partial u}}{[1 + (1-q)E_{p_2}[f(u)]]^2},$$

lo cual puede escribirse como

$$D(s_{q,p_2} \circ s_{q,p_1}^{-1})(u) \cdot v = \frac{\frac{\partial f(u)}{\partial u}}{[1 + (1-q)E_{p_2}[f(u)]]} - \left[\frac{[1 + (1-q)f(u)]}{[1 + (1-q)E_{p_2}[f(u)]]^2} \right] \frac{\partial E_{p_2}[f(u)]}{\partial u} \quad (4.12)$$

Note que

$$\begin{aligned} E_{p_2}[f(u)] &= E_{p_2} \left[(u \ominus_q K_{p_1}(u)) \oplus_q \ln_q \left(\frac{p_1}{p_2} \right) \right] \\ &= E_{p_2} \left[(u \ominus_q K_{p_1}(u)) + \ln_q \left(\frac{p_1}{p_2} \right) + (1-q)(u \ominus_q K_{p_1}(u)) \ln_q \left(\frac{p_1}{p_2} \right) \right] \\ &= E_{p_2} \left[\ln_q \left(\frac{p_1}{p_2} \right) \right] + \left[1 + (1-q) \ln_q \left(\frac{p_1}{p_2} \right) \right] E_{p_2}[u \ominus_q K_{p_1}(u)], \end{aligned}$$

aplicando derivada parcial con respecto a u , obtenemos

$$\frac{\partial E_{p_2}[f(u)]}{\partial u} = E_{p_2} \left[\frac{[1 + (1-q)K_{p_1}(u) - [1 + (1-q)u]DK_{p_1}(u) \cdot v]W}{[1 + (1-q)K_{p_1}(u)]^2} \right],$$

donde $W = 1 + (1-q)\ln_q\left(\frac{p_1}{p_2}\right)$. Ahora, tomando derivada parcial respecto a u en (4.11) tenemos que el argumento en E_{p_2} de esta última expresión es precisamente

$$\frac{\partial f(u)}{\partial u} = \left[\frac{[1 + (1-q)K_{p_1}(u) - DK_{p_1}(u)v[1 + (1-q)u]]W}{[1 + (1-q)K_{p_1}(u)]^2} \right].$$

De todo lo anterior, usando (4.12), se sigue que

$$D(s_{q,p_2} \circ s_{q,p_1}^{-1})(u) \cdot v = \frac{\frac{\partial f(u)}{\partial u}}{[1 + (1-q)E_{p_2}[f(u)]]} - \frac{[1 + (1-q)f(u)]}{[1 + (1-q)E_{p_2}[f(u)]]^2} E_{p_2} \left[\frac{\partial f(u)}{\partial u} \right]$$

y puesto que $E_{p_2}[f(u)]$ es constante podemos escribir esta última expresión como

$$D(s_{q,p_2} \circ s_{q,p_1}^{-1})(u) \cdot v = \frac{\frac{\partial f(u)}{\partial u}}{[1 + (1-q)E_{p_2}[f(u)]]} - \frac{[1 + (1-q)f(u)]}{[1 + (1-q)E_{p_2}[f(u)]]} E_{p_2} \left[\frac{\frac{\partial f(u)}{\partial u}}{[1 + (1-q)E_{p_2}[f(u)]]} \right].$$

Ahora, tomando

$$A(u) = \frac{\frac{\partial f(u)}{\partial u}}{[1 + (1 - q)E_{p_2}[f(u)]]} \quad \text{y} \quad B(u) = \frac{[1 + (1 - q)f(u)]}{[1 + (1 - q)E_{p_2}[f(u)]]},$$

la derivada puede ser escrita finalmente en la forma

$$D(s_{q,p_2} \circ s_{q,p_1}^{-1})(u) \cdot v = A(u) - B(u)E_{p_2}[A(u)]. \quad (4.13)$$

□

Podemos entonces establecer el teorema principal de esta sección.

Teorema 4.2.1: La colección de pares $\{(\mathcal{U}_p, s_{q,p})\}_{p \in \mathfrak{M}_\mu}$ es un C^∞ -atlas modelado sobre B_p .

Prueba: Consideremos la familia de pares $\{(\mathcal{U}_p, s_{q,p})\}_{p \in \mathfrak{M}_\mu}$. Obviamente la familia de \mathcal{U}_p es un cubrimiento abierto de \mathfrak{M}_μ , cada \mathcal{U}_p es abierto en \mathfrak{M}_μ y $s_{q,p}$ es biyectiva por proposición 4.2.1 y 4.2.2. Esto significa que cada $(\mathcal{U}_p, s_{q,p})$ es una carta y el correspondiente mapa de transición para $p_1, p_2 \in \mathfrak{M}_\mu$ (con $\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2} \neq \emptyset$) es dado por $s_{q,p_2} \circ e_{q,p_1} : s_{q,p_1}(\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2}) \rightarrow s_{q,p_2}(\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2})$, como se define en (4.9). Por proposición 4.2.3, los conjuntos $s_{q,p}(\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2})$ son abiertos en la topología de B_p , lo cual a su vez implica que los mapas coordenados $s_{q,p}$ son homeomorfismos. La proposición 4.2.4 garantiza que los espacios $s_{q,p_1}(\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2})$ y $s_{q,p_2}(\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2})$ son topológicamente homeomorfos (por lo tanto, el mapeo definido por (4.9) es un isomorfismo topológico). Sólo resta demostrar que el mapa de transición $s_{q,p_2} \circ e_{q,p_1}$ es C^∞ . Para hacer esto observemos que el mapa definido por (4.9) puede escribirse en la forma $s_{q,p_2} \circ e_{q,p_1}(u) = \frac{1}{g(u)}h(u)$, donde

$$h(u) = u + [1 + (1 - q)u] \ln_q \left(\frac{p_1}{p_2} \right) - E_{p_2} \left[u + [1 + (1 - q)u] \ln_q \left(\frac{p_1}{p_2} \right) \right]$$

y

$$g(u) = 1 + \left\{ [1 + (1 - q)K_{p_2}(u)] (1 - q)E_{p_2} \left[u + [1 + (1 - q)u] \ln_q \left(\frac{p_1}{p_2} \right) \right] \right\}.$$

Así, finalmente es suficiente probar que h y g son C^∞ . Existe $z \in (\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2})$ tal que $p_2 = e_{q,z}(u)$ y así $u + [1 + (1 - q)u] \ln_q \left(\frac{p_1}{p_2} \right) \in \mathcal{V}_z$. Denotando $v = u + [1 + (1 - q)u] \ln_q \left(\frac{p_1}{p_2} \right)$ y aplicando (2) del teorema 4.1.2, se obtiene que $DK_z(u)v = [M_p(u)]^{1-q}E_{p_2}[v]$. Ahora, dado que K_z es C^∞ en \mathcal{V}_z , se sigue que $E_{p_2} \left[u + [1 + (1 - q)u] \ln_q \left(\frac{p_1}{p_2} \right) \right]$ es C^∞ . Puesto que cada función afin es C^∞ y h puede ser expresada como la suma de una función afin y una C^∞ -función, h es ciertamente C^∞ . Análogamente, puesto que K_{p_2} es C^∞ en \mathcal{V}_{p_2} , se sigue que g es también C^∞ . La proposición 4.2.4 también garantiza que el atlas definido está modelado sobre espacios de Banach. □

Los mapeos coordenados, y sus inversos, simplifican a los modelos exponenciales usados en [26], una dirección de trabajo es encontrar un espacio con el modelo de Orlicz para

esta variedad cuya existencia está asegurada por el hecho que si $u \in L^\infty(p \cdot \mu)$, debe haber un espacio de Orlicz tal que $u \in L^\Psi(p \cdot \mu)$ para alguna función de Young Ψ [6][39].

A continuación se presenta la construcción del fibrado tangente para la variedad.

Sea $p \in \mathfrak{M}_\mu$, una curva a través de p es un modelo q -exponencial paramétrico unidimensional: $g : I \subset \mathfrak{R} \rightarrow \mathfrak{M}_\mu$ tal que $t \rightarrow g(t) \in \mathfrak{M}_\mu$, donde $g(t_0) = p$ para algún $t_0 \in I$. Usualmente se asume que $t_0 = 0$; no obstante esto es innecesario en esta estructura. Sean $(\mathcal{U}_{p_1}, s_{q,p_1})$ y $(\mathcal{U}_{p_2}, s_{q,p_2})$ cartas alrededor de $p \in \mathfrak{M}_\mu$; entonces $g(t) = e_{q,p_1}(u_1) = e_q^{u_1 \ominus_q K_{p_1}(u_1)} p_1$, donde para cualquier $t \in I$ se sigue que $u_1(t) = s_{q,p_1}(g(t))$. Adicionalmente, $g(t) = e_{q,p_2}(u_2)$ y $u_2(t) = s_{q,p_2}(g(t))$. Las variables aleatorias $u_1(t_0)$ y $u_2(t_0)$ están relacionadas por $u_2(t_0) = (s_{q,p_2} \circ e_{q,p_1})(u_1(t_0))$. Puesto que $u_1(t_0) = s_{q,p_1}(p)$ y por la regla de la cadena $u_2'(t_0) = (s_{q,p_2} \circ e_{q,p_1})'(u_1(t_0)) \cdot u_1'(t_0)$ lo cual es igual a $u_2'(t_0) = (s_{q,p_2} \circ e_{q,p_1})'(s_{q,p_1}(p)) \cdot u_1'(t_0)$. La expresión anterior es una relación de la equivalencia que llena los requisitos para definir el espacio tangente sobre una variedad [40], por lo tanto el espacio tangente para $p \in \mathfrak{M}_\mu$ es un espacio vectorial topológico con la topología inducida de cualquiera de los espacios $L^\infty(z \cdot \mu)$, tal que $p \in \mathcal{U}_z$, y es dada por $\mathcal{T}_p(\mathfrak{M}_\mu) = \{[(\mathcal{U}_p, s_{q,p}, u')] : p \in \mathfrak{M}_\mu\}$.

Proposición 4.2.6: Sea $g(t)$ una curva regular para \mathfrak{M}_μ donde $g(t_0) = p$, y $u(t) \in \mathcal{V}_z$ es la representación en coordenadas sobre $s_{q,z}$. Entonces $g(t) = e_q^{[u(t) \ominus_q K_z(u(t))]}_z$ y también:

1. $\frac{d}{dt} \ln_q \left(\frac{g(t)}{p} \right)_{t=t_0} = Tu'(t) - Q[M_p(u(t))]^{1-q} E_p[u'(t)]$ para algunas constantes T y Q .
2. Si $z = p$, i.e. las cartas se centran en el mismo punto; los vectores tangentes se identifican con la función q -score en t dada por $\frac{d}{dt} \ln_q \left(\frac{g(t)}{p} \right)_{t=t_0} = Tu'(t_0)$.
3. Considere un modelo q -exponencial bidimensional

$$f(t, q) = e_q^{(tu \ominus_q K_p(tu))} p \quad (4.14)$$

donde si q tiende a 1 en (4.14), se obtienen los modelos exponenciales unidimensionales $e^{(tu - K_p(tu))} p$. La función q -score $\frac{d}{dt} \ln_q \left(\frac{f(t)}{p} \right)_{t=t_0}$ para un modelo q -exponencial unidimensional (4.14) corresponde al espectro $[u]$ en $t = t_0$, donde $u \in \mathcal{V}_p$.

Prueba:

$$\begin{aligned} 1. \frac{d}{dt} \ln_q \left(\frac{g(t)}{p} \right)_{t=t_0} &= \frac{d}{dt} \ln_q \left[\frac{e_q^{[u(t) \ominus_q K_z(u(t))]}_z}{e_q^{[u(t_0) \ominus_q K_z(u(t_0))]}_z} \right]_{t=t_0} \\ &= \frac{d}{dt} \left[\frac{[u(t) - K_z(u(t))] \left[\frac{1 + (1-q)K_z(u(t_0))}{1 + (1-q)K_z(u(t))} \right]^{-[u(t_0) - K_z(u(t_0))]}}{[1 + (1-q)K_z(u(t_0))]} \right]_{t=t_0} \\ &= \frac{d}{dt} \left[\frac{u(t) - K_z(u(t))}{1 + (1-q)K_z(u(t))} \right]_{t=t_0} \end{aligned}$$

$$\begin{aligned} &= \frac{1}{[1 + (1 - q)K_z(u(t_0))]} u'(t_0) - \frac{[1 + (1 - q)u(t_0)]}{[1 + (1 - q)K_z(u(t_0))]^2} \left(\frac{d}{dt} K_z(u(t_0)) \right) \\ &= T u'(t_0) - Q \left(\frac{d}{dt} K_z(u(t_0)) \right), \end{aligned}$$

donde $T = \frac{1}{[1 + (1 - q)K_z(u(t_0))]}$ y $Q = \frac{[1 + (1 - q)u(t_0)]}{[1 + (1 - q)K_z(u(t_0))]^2}$.

Del teorema 4.1.2 se sigue que $\frac{d}{dt} K_z(u(t)) = [M_p(u(t))]^{1-q} E_{g(t)}[u'(t)]$.

Por tanto

$$\frac{d}{dt} \ln_q \left(\frac{g(t)}{p} \right)_{t=t_0} = T u'(t_0) - Q [M_p(u(t))]^{1-q} E_{g(t_0)}[u'(t_0)].$$

2. Si $z = p$ entonces $E_p[u'(t_0)] = E_z[u'(t_0)]$; puesto que $u(t)$ está centrada entonces $E_p[u'(t_0)] = E_z[u'(t_0)] = 0$ y el resultado deseado se sigue.

3. Tomando $u(t) = tu$ en (1) un simple cálculo lleva a

$$\frac{d}{dt} \ln_q \left(\frac{f(t)}{p} \right)_{t=t_0} = T u.$$

□

Así el espacio tangente $T_p(\mathfrak{M}_\mu)$ es identificado con la colección de funciones q -score o los modelos q -exponenciales unidimensionales (4.14).

El modelo \mathfrak{M}_μ unido a la colección de espacios tangentes es el fibrado tangente, denotado $\mathcal{T}(\mathfrak{M}_\mu)$. así,

$$\mathcal{T}(\mathfrak{M}_\mu) := \{(f, u) : f \in \mathcal{U}_p \subset \mathfrak{M}_\mu \text{ y } u \text{ es la clase de vectores tangentes a } f\}.$$

Se sabe que esto es una variedad, donde las cartas (mapeos triviales) están dadas por $(g, u) \in \mathcal{T}(\mathcal{U}_p) \rightarrow (s_{q,p}(g), A(u) - B(u)E_p[A(u)])$, definida en la colección de subconjuntos abiertos $\mathcal{U}_p \times \mathcal{V}_p$ de $\mathfrak{M}_\mu \times L^\infty(p \cdot \mu)$.

Los mapeos de transición, para esta variedad están dados para $(u_1, v_1) \in \mathcal{V}_p \times L^\infty(p \cdot \mu)$ por

$$(u_1, v_1) \rightarrow ((s_{q,z} \circ e_{q,p})(u_1), A(v_1) - B(v_1)E_z[A(v_1)]) \in \mathcal{V}_z \times L^\infty(z \cdot \mu),$$

que da la representación local del fibrado. Se espera que la definición explícita del fibrado tangente sea pertinente en el cómputo de aproximación de modelos estadísticos paramétricos. Además, debe investigarse si los modelos paramétricos construidos en [1] están incluidos en esta variedad como curvas.

4.3. Geometría inducida por q -Divergencias

A continuación se define el funcional q -divergencia, buscando una métrica y las conexiones de la variedad, para caracterizar la geometría de la variedad de información q -exponencial.

En primer lugar, se verá que el funcional de entropía relativa es un caso especial de la Φ -divergencia de Csiszár, y entonces algunas propiedades son aseguradas.

Sea f una función definida para $t \neq 0$ y $0 < q < 1$ como

$$f(t) = -t \ln_q \left(\frac{1}{t} \right), \quad (4.15)$$

claramente, $f(t) < 0$ para $0 < t < 1$; y $f(t) \geq 0$ para $t \geq 1$. Por cálculo directo se sigue que $\frac{df}{dt} = \frac{1}{1-q}[1 - qt^{q-1}]$ y $\frac{d^2f}{dt^2} = qt^{q-2}$. Entonces f es creciente para $t > q^{\frac{1}{1-q}}$ y estrictamente convexa para toda t ; también f alcanza un valor mínimo de $-q^{\frac{q}{1-q}}$ cuando $t = q^{\frac{1}{1-q}}$. La prueba de la siguiente proposición es directa de la definición.

Proposición 4.3.1: Sea f dada por (4.15), entonces

1. $\lim_{t \rightarrow 0^+} f(t) = 0$ y $f(1) = 0$.
2. Si $a > 0$, $\lim_{\epsilon \rightarrow 0^+} f\left(\frac{a}{\epsilon}\right) = a \lim_{t \rightarrow \infty} \frac{f(t)}{t} = \frac{a}{1-q}$.

Sean $p, z \in \mathfrak{M}_\mu$, la q -divergencia de z con respecto a p es dada por

$$I^{(q)}(z//p) = \int_{\Omega} p f\left(\frac{z}{p}\right) d\mu. \quad (4.16)$$

De las propiedades de f , es claro que (4.16) es una Φ -divergencia de Csiszár; es más, por la definición de q -logaritmo se sigue que

$$I^{(q)}(z//p) = \frac{1}{1-q} \left[1 - \int_{\Omega} z^q p^{1-q} d\mu \right], \quad (4.17)$$

que es el funcional de divergencia de Tsallis [41]. Algunas propiedades de este funcional son bien conocidas, por ejemplo que es igual al funcional α -divergente y se puede llevar un factor constante donde $\alpha = 1 - 2q$, satisfaciendo el criterio de invarianza.

Además, cuando $q \rightarrow 0$ entonces $I^{(q)}(z//p) = 0$ y si $q \rightarrow 1$ entonces $I^{(q)}(z//p) = K(z//p)$ que es el funcional de divergencia de Kullback-Leibler [42].

Proposición 4.3.2: Dados $p, z \in \mathfrak{M}_\mu$ entonces

1. $I^{(q)}(z//p) \geq 0$ y la igualdad se cumple si $p = z$.
2. $I^{(q)}(z//p) \leq \int_{\Omega} (z - p) f'\left(\frac{z}{p}\right) d\mu$.

Prueba: Sea $p, z \in \mathfrak{M}_\mu$,

1. $I^{(q)}(z//p) = \int_{\Omega} p f\left(\frac{z}{p}\right) d\mu = \int_{\Omega} f\left(\frac{z}{p}\right) dp \cdot \mu$, por la desigualdad de Jensen y la proposición 4.3.1-1, se sigue que

$$\int_{\Omega} f\left(\frac{z}{p}\right) dp \cdot \mu \geq f\left[\int_{\Omega} \left(\frac{z}{p}\right) dp \cdot \mu\right] = 0,$$

Así $I^{(q)}(z//p) \geq 0$. La igualdad se da inmediatamente.

2. Puesto que f es convexa es fácil mostrar que

$$f'\left(\frac{z}{p}\right)\left(\frac{z}{p} - 1\right) \geq f\left(\frac{z}{p}\right) - f(1) \geq f'(1)\left(\frac{z}{p} - 1\right);$$

Si multiplicamos por p y por integración sobre Ω y puesto que $f(1) = 0$, podemos deducir

$$\int_{\Omega} (z - p) f'\left(\frac{z}{p}\right) d\mu \geq I^{(q)}(z//p) \geq \int_{\Omega} (z - p) f'(1) d\mu,$$

y puesto que $f'(1) = 1$ el resultado deseado se sigue. \square

Note que el término superior en la parte 2. de esta proposición, es simplificado por

$$\int_{\Omega} (z - p) f'\left(\frac{z}{p}\right) d\mu = \frac{q}{1-q} \int_{\Omega} (p - z) \left(\frac{p}{z}\right)^{1-q} d\mu = \frac{q}{1-q} \{E_p\left[\left(\frac{p}{z}\right)^{1-q}\right] - E_z\left[\left(\frac{p}{z}\right)^{1-q}\right]\}.$$

Proposición 4.3.3: Sea $p \in \mathfrak{M}_{\mu}$, $z \in \mathcal{U}_p$ y $u = s_{q,p}(z)$. Si tomamos $*u = \frac{z}{p} - 1$ y el mapeo $H_p(*u) = E_p\left[\left(1 + *u\right) \ln_q\left(\frac{1}{1+*u}\right)\right]$ entonces,

$$I^{(q)}(z//p) = \frac{K_p(u)}{1+(1-q)K_p(u)} \text{ y } H_p(*u) = -I^{(q)}(z||p).$$

Prueba: La definición de q -divergencia implica $I^{(q)}(p||z) = -E_p[\ln_q(\frac{z}{p})]$. De $z = e_q^{[u \ominus_q K_p^q(u)]} p$, obtenemos

$$\begin{aligned} I^{(q)}(p||z) &= -E_p \left[\ln_q\left(\frac{z}{p}\right) \right] = - \left(E_p \left[\frac{u - K_p(u)}{1+(1-q)K_p(u)} \right] \right) \\ &= - \left(\frac{E_p[u] - K_p(u)}{1+(1-q)K_p(u)} \right) = \frac{K_p(u)}{1+(1-q)K_p(u)}. \end{aligned}$$

Finalmente, tenemos

$$H_p(*u) = E_p \left[\left(1 + *u\right) \ln_q\left(\frac{1}{1+*u}\right) \right] = E_p \left[\left(\frac{z}{p}\right) \ln_q\left(\frac{p}{z}\right) \right] = -I^{(q)}(z||p). \quad \square$$

Hay dos aspectos importantes a ver; uno es que los mapeos de coordenadas

$$s_{q,p}(z) = \ln_q\left(\frac{z}{p}\right) \ominus_q E_p \left[\ln_q\left(\frac{z}{p}\right) \right],$$

puede ser escrito como

$$s_{q,p}(z) = \left(\frac{1}{1+(q-1)I^{(q)}(p//z)} \right) \left(\ln_q\left(\frac{z}{p}\right) + I^{(q)}(p//z) \right);$$

así la relación entre la q -divergencia y la variedad construida está clara en el sentido que la representación de coordenada de la densidad $z \in \mathcal{U}_p$ contiene la información de p con respecto a z . Por otro lado, de $s_{q,p}(z) = u$, obtenemos

$$E_z[u] = \left(\frac{1}{1+(q-1)I^{(q)}(p//z)} \right) (I^{(q)}(p//z) - D^{(q)}(z//p))$$

donde $D^{(q)}(z//p)$ es una q -deformación de el funcional de divergencia de Kullback-Leibler. El funcional $I^{(q)}(p//z)$ induce una estructura Riemanniana sobre \mathfrak{M}_μ .

El resultado siguiente, es necesario para garantizar la existencia de la métrica Riemanniana de la variedad.

Proposición 4.3.4: Sea $p, z \in \mathfrak{M}_\mu$ entonces

$$(d_u)_z I^q(z || p)|_{z=p} = 0,$$

donde los subíndice p, z significan que la derivada direccional se toma con respecto al primero y los segundos argumentos en $I^q(z || p)$, respectivamente, a lo largo de la dirección $u \in T_z(\mathfrak{M}_\mu)$ o $v \in T_p(\mathfrak{M}_\mu)$.

Prueba: Escribiendo (4.17) como

$$I^q(z || p) = \int_{\Omega} \frac{[(1-q)p+(q)z-p^{1-q}z^{(q)}]}{1-q} d\mu,$$

se sigue

$$(d_u)_z I^q(z || p) = \frac{1}{1-q} \int_{\Omega} [q - qp^{(1-q)}z^{q-1}] u d\mu$$

y

$$(d_v)_z I^q(z || p) = \frac{1}{1-q} \int_{\Omega} [(1-q) - (1-q)p^{(-q)}z^{(q)}] v d\mu$$

cuando $z = p$ se cumple el resultado deseado. \square

Según la proposición (16) en [16], el funcional $I^q(z || p)$ es acotado, puesto que

$$I^q(z || p) \geq 0 \text{ y la igualdad se cumple si } p = z$$

y

$$I^q(z || p) \leq \int_{\Omega} (z - p) f'(\frac{z}{p}) d\mu.$$

Entonces, junto con la proposición anterior, el funcional q -divergencia induce una métrica Riemanniana g y un par de conexiones, vea Eguchi [43], dadas por

$$g(u, v) = -(d_u)_z (d_v)_p I^q(z || p)|_{z=p} \quad (4.18)$$

$$\langle \nabla_w u, v \rangle = -(d_w)_z(d_u)_z(d_v)_p I^q(z \parallel p)|_{z=p} \quad (4.19)$$

donde $v \in T_p(\mathfrak{M}_\mu)$, $u \in T_p(\mathfrak{M}_\mu)$ y w es un campo vectorial.

Denotemos por $\sum(\mathfrak{M}_\mu)$ el conjunto de campos vectoriales $u : \mathcal{U}_p \rightarrow T_p(\mathcal{U}_p)$, y $F(\mathfrak{M}_\mu)$ el conjunto de \mathcal{C}^∞ funciones $f : \mathcal{U}_p \rightarrow \mathcal{R}$. El siguiente resultado establece la métrica.

Proposición 4.3.5: Sea $p, z \in \mathfrak{M}_\mu$ y v, u campos vectoriales, el tensor métrico (campo) $g : \sum(\mathfrak{M}_\mu) \times \sum(\mathfrak{M}_\mu) \rightarrow F(\mathfrak{M}_\mu)$ es dado por

$$g(u, v) = q \int_{\Omega} \frac{uv}{p} d\mu.$$

Prueba: Por cálculo directo sobre $I^q(z \parallel p)$, se sigue que

$$(d_v)_p I^q(z \parallel p) = \frac{1}{1-q} \int_{\Omega} [(1-q) - (1-q)p^{(-q)} z^{(q)}] v d\mu$$

y

$$(d_u)_z(d_v)_p I^q(z \parallel p) = -q \int_{\Omega} [p^{(-q)} z^{(q-1)}] uv d\mu,$$

así por (4.18), se sigue que

$$g(u, v) = q \int_{\Omega} \frac{uv}{p} d\mu.$$

□

Note que cuando $q \rightarrow 1$, esta métrica se reduce a la inducida por el funcional (α, β) -divergencia que induce la métrica de Fisher sobre modelos paramétricos.

Las conexiones se caracterizan como sigue.

Proposición 4.3.6: La familia de derivadas covariantes (conexiones)

$$\nabla_w^{(q)} u : \sum(\mathfrak{M}_\mu) \times \sum(\mathfrak{M}_\mu) \rightarrow \sum(\mathfrak{M}_\mu),$$

se dan como

$$\nabla_w^{(q)} u = d_w u - \left(\frac{1-q}{p}\right) uw.$$

Prueba: Considerando $(d_u)_z(d_v)_p I^q(z \parallel p)$ como en la prueba de la proposición 4.3.5, obtenemos

$$-(d_w)_z(d_u)_z(d_v)_p I^q(z \parallel p) = q \int_{\Omega} p^{(-q)} [(q-1)z^{(q-2)} uw + z^{(q-1)} d_w u] v d\mu.$$

por la proposición previa, se sigue que

$$g(\nabla_w^{(q)} u, v) = q \int_{\Omega} \frac{\nabla_w^{(q)} u}{p} v d\mu$$

y entonces

$$g(\nabla_w^{(q)} u, v) = \langle \nabla_w^{(q)} u, v \rangle.$$

Entonces

$$qp^{-1}[(q-1)p^{-1}uw + d_w u] = q \frac{\nabla_w^{(q)} u}{p},$$

así

$$\nabla_w^{(q)} u = d_w u - \left(\frac{1-q}{p}\right)uw. \quad \square$$

Es fácil probar que la conexión conjugada asociada es dada por $\nabla_w^{*(q)} u = d_w u - \frac{q}{p}uw$. Note que la toma de $q = \frac{1-\alpha}{2}$ rinde a la familia uniparametrica de Amari de α -conexiones en la forma

$$\nabla_w^{(\alpha)} u = d_w u - \left(\frac{1+\alpha}{2p}\right)uw;$$

y tomando $q = \frac{1}{2}$ resulta la conexión de Levi-Civita. Por último, la geometría de la variedad se caracteriza por el cálculo de los tensores de torsión y curvatura, para el que se demostró que es igual a cero.

Proposición 4.3.7: Para la variedad q -exponencial y la conexión dada en la proposición previa el tensor de curvatura y el tensor de torsión satisfacen $R(u, v, w) = 0$ y $T(u, v) = 0$.

Prueba: Recuerde que

$$R(u, v, w) = \nabla_u \nabla_v w - \nabla_v \nabla_u w - \nabla_{[u,v]} w. \quad (4.20)$$

$$T(u, v) = \nabla_u v - \nabla_v u - [u, v]. \quad (4.21)$$

Usando la forma general

$$\nabla_v w = d_v w + \Gamma(v, w),$$

donde

$$\Gamma : \sum(\mathfrak{M}_\mu) \times \sum(\mathfrak{M}_\mu) \rightarrow \sum(\mathfrak{M}_\mu)$$

es una forma bilineal (la contraparte del símbolo de Christoffel); y puesto que

$$d_u(\nabla_v w) = d_u(d_v w) + \Gamma(d_u v, w) + \Gamma(v, d_u w) + d_u \Gamma,$$

donde $d_u\Gamma = \Gamma'u$ es la derivada de la forma bilineal Γ , se sigue que

$$\nabla_u \nabla_v w = d_u(d_v w) + \Gamma(d_u v, w) + \Gamma(v, d_u w) + d_u\Gamma + \Gamma(d_v u, w) + \Gamma(u, \Gamma(v, w)),$$

y

$$\nabla_v \nabla_u w = d_v(d_u w) + \Gamma d_v u, w) + \Gamma(u, d_v w) + d_v\Gamma + \Gamma d_u v, w) + \Gamma(v, \Gamma(u, w)).$$

Por otra parte,

$$\nabla[u, v]w = d[u, v]w + \Gamma(d_u v - d_v u, w) = d[u, v]w + \Gamma(d_u v, w) - \Gamma(d_v u, w),$$

donde $[u, v] = d_u v - d_v u$, y entonces

$$\nabla[u, v]w = d_u(d_v w) - d_v(d_u w) + \Gamma(d_u v, w) - \Gamma(d_v u, w).$$

sustituyendo en (4.20) y (4.21), se sigue que $R(u, v, w) = \Gamma(v, \Gamma(u, w)) - \Gamma(u, \Gamma(v, w)) + d_u\Gamma(v, w) - d_v\Gamma(u, w)$, y

$$T(u, v) = \Gamma(u, v) - \Gamma(v, u).$$

Puesto que $\Gamma(u, v) = -\frac{1-q}{p}uv$, y $d_u\Gamma(v, w) = \frac{1-q}{p^2}uvw$, obtenemos

$$\begin{aligned} R(u, v, w) &= \Gamma(u, -\frac{1-q}{p}vw) - \Gamma(v, -\frac{1-q}{p}uw) + \frac{1-q}{p^2}uvw - \frac{1-q}{p^2}vuw = \\ &= -[\frac{1-q}{p}]u[-\frac{1-q}{p}vw] + [\frac{1-q}{p}]v[-\frac{1-q}{p}uw] = 0 \end{aligned}$$

Conclusiones

Amari y Ohara emplean la familia q -exponencial en la construcción de una variedad de dimensión finita y deducen que tal familia tienen una estructura geométrica dualmente plana derivada de transformación de Legendre. En *Geometry of q -exponential family of probability distributions*, para $q \rightarrow 1$, se obtiene la matriz de información de Fisher.

La variedad de Banach elaborada por Loaiza-Quiceno permite relacionarla con la entropía relativa según Tsallis, y tiene por objeto servir como modelo geométrico para caracterizar sistemas no extensivos. Cada modelo paramétrico q -exponencial se identifica con el espacio tangente. Constituye una generalización de la variedad de información de Pistone y Sempi, pues, ésta se obtiene bajo la condición, $q \rightarrow 1$.

Los modelos q -exponenciales también se puede obtener mediante el empleo del principio de máxima entropía optimizando el funcional de entropía de Tsallis.

El funcional q -divergencia permite definir una métrica, y en consecuencia generar una geometría riemanniana que generaliza la dada por Fisher y Levi-Civita.

Los diferentes funcionales señalados permiten proyectar la entropía de un sistema proporcionando un avance relevante en la exposición de los fenómenos.

Problemas abiertos

Puesto en la geometría de la variedad q -exponencial para el caso no paramétrico, el mapeo $\nabla^{(q)} \leftrightarrow \nabla^{(\alpha)}$ es suave, se espera que las curvas geodésicas y transportes paralelos obtenidos a partir de la q -conexiones conserva un isomorfismo suave con las curvas dadas por α -conexiones. Además, se debe investigar si el campo tensor métrico en la proposición 4.3.5 viene dado por una transformación conformal o la métrica de información de Fisher.

Muchos de los desarrollos realizados para las variedades estándar aún no se han tratado para las variedades q -deformadas, por ejemplo las posibles relaciones de dualidad en densidades y densidades mixtas, entre otros problemas.

Falta por discutir el caso no paramétrico modelado sobre espacios de funciones que no sean esencialmente acotadas, podrían ser espacios de funciones de Orlicz.

Bibliografía

- [1] Amari, S. Ohara, A.: *Geometry of q-exponential family of probability distributions*. Entropy. Vol. 13, (2011), pág 1170-1185.
- [2] Loaiza, Gabriel y Quiceno, Héctor: A q-exponential statistical Banach Manifold. Journal of Mathematical Analysis and Applications. Vol. 398, (2013), pág 466-476.
- [3] Loaiza, Gabriel y Quiceno, Héctor: A Riemannian geometry in the q-exponential Banach manifold induced by q-divergences. Lecture Notes in Computer Science. Vol. 8085, (2013), pág 737-742.
- [4] M.R Grasselli, Dual connections in nonparametric classical information geometry, Ann. Inst. Stat. Math. Vol. 62 (2010), pág 873–896.
- [5] D.J.H. Garling. Inequalities A journey into Linear Analysis. Cambridge University Press, New York, (2007), pág 74,75.
- [6] M.M. Rao, Z.D. Ren, *Theory of Orlicz Spaces*, Marcel Dekker Inc., New York, NY, (1991).
- [7] M.M. Rao, Z.D. Ren, *Applications of Orlicz Spaces*, Marcel Dekker Inc, (2002).
- [8] Royden H. L. Real Analysis, Segunda edición. The Macmillan Company, (1968).
- [9] Blanco L. *Probabilidad.*, Universidad Nacional de Colombia, Bogotá, (2004).
- [10] Arango, J.: Una k-Deformación para la variedad de información estadística. Tesis de maestría, Universidad EAFIT. Medellín (2012).
- [11] Rao, C.R: *Information and accuracy attainable in estimation of statistical parameters*. Bull. Calcutta Math. Soc. Vol. 37, (1945), pág 81-91.
- [12] Jeffreys, H.: *An invariant form for the prior probability in estimation problems*. Proceedings of Royal Society A. Vol. 186,(1946), pág 453-461.
- [13] Efron, B.: *Defining the curvature of a statistical problem (with applications to second order efficiency)*. Annals of Statistics. Vol. 3, (1975), pág 1189-1242.
- [14] Dawid, A.P: *On the concepts of sufficiency and ancillarity in the presence of nuisance parameters*. Journal of the Royal Statistical Society B. Vol.37, (1975), pág.248-258.

- [15] Tsallis, C.: *Possible generalization of Boltzmann-Gibbs statistics*. J.Stat. Phys. Vol. 52,(1988), pág 479-487.
- [16] Borges, E.P.: *Manifestações dinâmicas e termodinâmicas de sistemas não-extensivos*. Tese de Doutorado, Centro Brasileiro de Pesquisas Físicas, Rio de Janeiro (2004).
- [17] Lee, J.M. Riemannian manifolds. An introduction to curvature. GTM 176. ISBN 0-387-98271-X.
- [18] Robert M. Wald, General Relativity, Chicago University Press, ISBN 0-226-87033-2.
- [19] Gibilisco, P., Pistone, G.: *Connections on non-parametric statistical manifolds by Orlicz space geometry*. Infinite Dimensional Analysis Quantum Probability and Related Topics. 1, 325-347 (1998).
- [20] Amari, S.: *Differential-geometrical methods in statistics*. Springer, New York (1985).
- [21] Amari, S., Nagaoka, H.: *Methods of information Geometry*. RI: American Mathematical Society. Translated from the 1993 Japanese original by Daishi Harada, Providence (2000).
- [22] Kufner, A.: *Function Spaces*. Prague, 1977.
- [23] Colin Bennett and Robert Sharpley, Interpolation of operators, Pure and Applied Mathematics, vol. 129, Academic Press Inc., Boston, MA, 1988, ISBN 0-12-088730-4. MR 89e:46001.
- [24] S. Amari and A. Cichocki. *Information geometry of divergence functions* Bulletin of the polish academy of sciences technical sciences Vol. 58, No. 1, 2010
- [25] Amari, S. *Diferential geometry derived from divergence functions: information geometry approach*. Mathematics of Distances and Applications. <http://www.mis.mpg.de/de/veranstaltungen/konferenzen-und-workshops/2010/infgeo/abstract-1891-shun-ichi-amari.html>
- [26] 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). pág.2, 1543–1561.
- [27] Cena, Alberto: Geometric Structures on the non-parametric statistical manifold. Tesi di Dottorato. Dottorato in Matematica, Università di Milano. Eguchi, S. (2005), pág. 15-17.
- [28] Naudts, J: The q-exponential family in statistical physics. Journal of Physics: Conference Series 201, (2010).
- [29] Tsallis, C. Nonadditive entropy and nonextensive statistical mechanics-an overview after 20 years. Braz. J. Phys. Vol. 39, (2009), pág 337–356.

- [30] G. Livadiotis¹ and D. J. McComas. Beyond kappa distributions: Exploiting Tsallis statistical mechanics in space plasmas. *Journal of Geophysical Research*, Vol. 114,(2009), A11105, doi:10.1029/2009JA014352.
- [31] Anirban DasGupta. *The Exponential Family and Statistical Applications*. Probability for Statistics and Machine Learning Springer Texts in Statistics (2011), pág 583-612.
- [32] Gavin E. Crooks, "Measuring thermodynamic length"(2007), *Physical Review Letters* (2009,) pp100602. DOI: 10.1103/PhysRevLett.99.100602
- [33] M.I. Kadets, V.M Kadets, *Series in Banach Spaces. Conditional and Unconditional Convergence*, Birkhäuser Verlag, Basel, (1997), Translated from the Russian by Andrei Iacob.
- [34] Cena, A., Pistone, G.: *Exponential statistical manifold*. *Annals of the Institute of Statistical Mathematics*. Vol. 59, (2006), pág 27-56.
- [35] Pistone, G.: *k-exponential models from the geometrical viewpoint*. *The European Physical Journal B*. Springer Berlin. Online 70 (2009), pág 29-37.
- [36] Rényi, A. On measures of entropy and information. In *Proceedings of the 4th Berkeley Symposium on Mathematics, Statistics and Probability*, Berkeley, CA, USA, 20 June–30 July 1960; pág 547–561.
- [37] Yamano, T. Information theory based on non-additive information content. *Phys. Rev. E* (2001), 63, 046105.
- [38] Rajagopal, A.K.; Abe, S. Implications of form invariance to the structure of nonextensive entropies. *Phys. Rev. Lett.* 1999, 83, 1711–1714.
- [39] W. H. Young. On classes of summable functions and their Fourier series. *Proc. Roy. Soc.* Vol.87, (1912), pág 225-229.
- [40] S. Lang: *Differential Manifolds*. Springer Verlang. New York, (1985).
- [41] S. Furuichi: Fundamental properties of Tsallis relative entropy. *J. Math. Phys.* Vol. 45, (2004), pág 4868-4877.
- [42] S. Kulback, and R. A. Leibler: On Information and Sufficiency. *Annals of Mathematics and Statistics*. Vol. (22), (1951), pág 79-86.
- [43] Eguchi, S.: Second order efficiency of minimum coontrast estimator in a curved exponential family. *Annals of Statistics*. Vol. 11, (1983), pág 793-803.