Geometría e información

Optativo

Mariela Portesi
Pedro Walter Lamberti
Steeve Zozor

Facultad de Ciencias Exactas

Esto es una dedicatoria
del libro.

Agradecimientos

Este es el texto de agradecimiento, max una carilla. Este es el texto de agradecimiento, max una carilla. Este es el texto de agradecimiento, max una carilla. Este es el texto de agradecimiento, max una carilla. Este es el texto de agradecimiento, max una carilla. Este es el texto de agradecimiento, max una carilla.

Esto es un epígrafe con texto simulado.
Esto es un epǵrafe con texto simulado.
AUTOR DEL EPÍGRAFE, TÍTULO DE LA OBRA

PRÓLOGO

Este libro surge de la experiencia de los autores en el dictado del curso semestral ”Métodos de geometría diferencial en teoría de la información", que se imparte en la Facultad de Ciencias Exactas de la Universidad Nacional de La Plata y en la Facultad de Matemática, Astronomía y Física de la Universidad Nacional de Córdoba. ...

Los autores

ADVERTENCÍA

Este libro surge de la experiencia de los autores en el dictado del curso semestral ”Métodos de geometría diferencial en teoría de la información", que se imparte en la Facultad de Ciencias Exactas de la Universidad Nacional de La Plata y en la Facultad de Matemática, Astronomía y Física de la Universidad Nacional de Córdoba. ...

Mariela Portesi
Grenoble, Junio de 2016

Índice

Referencías

CAPÍTULO 1
Nociones de teoría de la información

Steeve Zozor

“Deberías llamarla ‘entropía’, por dos motivos.
En primer lugar su función de incerteza
ha sido usada en la mecánica estadística
bajo ese nombre, y por ello, ya tiene un nombre.
En segundo lugar, y lo que es más importante,
nadie sabe lo que es realmenta la entropía,
por ello, en un debate, siempre llevará la ventaja.
VON NEUMANN TO SHANNON (TRIBUS & MCIRVINE, 1971)

1.1 Entropias y divergencias generalizadas

A pesar de que la entropia de Shannon y sus cantidades asociadas demostraron sus potencias tan de un punto de vista descriptivo que en termino de aplicaciones en la transmisión de la información y la compresión, varios nociones informacionales, de tipo entropias o divergencias, aparecieron luego. En esta sección no se desarollará todos los enfoques ni todas las aplicaciones tan la literatura es importante. La meta es dar los caminos conduciendo a las generalizaciones de la entropia de Shannon por un lado, y de la divergencia de Kullback-Leibler por el otro lado. No son siempre vinculados, a pesar de que sea desirable que a cada entropia sean asociados nociones de entropias condicionales y relativas.

1.1.1 Entropias y propiedades

Si la entropia de Shannon fue el punto de salida fundamental en todo el desarollo de la teoria de la información, un poco mas de una decada despues de su papel clave y muy completo, Rényi propuso una medida generalizada (Rényi, 1961). Su punto de vista fue mas matematico que fisico o ingeniero. Retomó los axiomas de Fadeev (Fadeev, 1956, 1958; Khinchin, 1957) a probabilidades incompletas $p = \begin{bmatrix} p_1 & \cdots & p_n \end{bmatrix}^t, \quad p_i \ge 0, \quad w_p = \sum_i p_i \le 1$: (i) la invarianza de $H(p)$ por permutación de os p_i, (ii) la continuidad de la incerteza elemental $H(p_i)$ (p_i visto como probabilidad incompleta), (iii) $H\left(\frac{1}{2}\right) = 1$, (iv) la aditividad $H(p \otimes q) = H(p) + H(q)$ donde $p \otimes q$ es el producto de Kronecker [1], *i. e.,* probabilidad conjunta de dos variables independientes, y consideró en lugar de la recursividad un axioma dicho de valor promedio, axioma muy parecido a la recursividad. Para p y q probabilidades incompletas tales que $p \cup q = \begin{bmatrix} p_1 & \cdots & p_n & q_1 & \cdots & q_m \end{bmatrix}^t$ sea incompleta ($w_p + w_q \le 1$), el axioma (v) es $H(p \cup q) = \frac{w_p\, H(p) + w_q\, H(q)}{w_p + w_q}$. Demostró que con (v) en lugar de la recursividad, el conjunto de axiomas conduce de nuevo a la entropia de Shannon. La generalización propuesta por Rényi era de generalizar el axioma (v) remplazando la media aritmético por una media generalizada (v') $H^{\mathrm{r}}(p \cup q) = g^{-1}\left(\frac{w_p\, g\big(H^{\mathrm{r}}(p)\big) + w_q\, g\big(H^{\mathrm{r}}(q)\big)}{w_p + w_q}\right)$ con g estrictamente monotona y continua, llamado media *cuasi-aritmética, o quasi-lineal, o de Kolmogorov-Nagumo*. De las propiedades de la media cuasi-aritmetica (Nagumo, 1930; Kolmogorov, 1930, 1991; Hardy, Littlewood & Pólya, 1952), eso es equivalente a buscar una entropia elemental $H^{\mathrm{r}}(p_i)$ y remplazar la media aritmética $\sum_i p_i H^{\mathrm{r}}(p_i)$ por una media de Kolmogorov-Nagumo, $g^{-1}\left(\sum_i p_i g\big(H^{\mathrm{r}}(p_i)\big)\right)$. Rényi propusó la función de Kolmogorov-Nagumo $g_\alpha(x) = 2^{(\alpha-1)x}, \quad \alpha > 0, \quad \alpha \ne 1$, probando que la entropia que los axiomas (i)-(ii)-(iii)-(iv)-(v') se cumplen y conduce a la entropia de Rényi de un vector de probabilidad p,

$$H^{\mathrm{r}}_\alpha(p) = \frac{1}{1-\alpha} \log_2 \left(\sum_{i=1}^n p_i^\alpha \right)$$

Relaxando el axioma (iii), se puede elegir $g_\alpha(x) = a^{(\alpha-1)x}, \quad a > 0, \quad a \ne 1$; el logaritmo será de la base a cualquiera; En lo que sigue, usaremog $\log$ sin precisar la elección de base. Rényi nombró esta medida de incerteza *entropia de orden* α. Notablemente,

$$H^{\mathrm{r}}_1(p) \equiv \lim_{\alpha \to 1} H^{\mathrm{r}}_\alpha(p) = H(p) \quad \text{entropia de Shannon}$$

En otros terminos, la clase de Rényi contiene como caso particular la entropia de Shannon. En su papel, Rényi introdujo una ganancia de información, parecida a una entropia relativa, probando que las solas entropias admisibles son la de Shannon y la que introdujo. Volveremos en la sección siguiente sobre esta entropia relativa, o divergencia de Rényi. **Preciser les quelques proprietes qui se conservent, dont la concavite pour $\alpha \le 1$.**

Unos años después de Rényi, de la famosa escuela matematica **checa**, J. Havrda & F. Charvát en (Havrda & Charvát, 1967) volvieron a los axiomas de Khintchin, para extender la entopia de Shannon, *i. e.,* considerando (i) la invarianza por permutación, (ii) la continuidad, (iii) la expensividad, (iv) $H^{\mathrm{hc}}(1) = 0$ y $H^{\mathrm{hc}}\left(\frac{1}{2}, \frac{1}{2}\right) = 1$, pero generalisando la recursividad por (v) $H^{\mathrm{hc}}(p_1, \ldots, p_n) = H^{\mathrm{hc}}(p_1, \ldots, p_{n-2}, p_{n-1} + p_n) + \alpha(p_{n-1} + p_n)^\alpha H^{\mathrm{hc}}\left(\frac{p_{n-1}}{p_{n-1}+p_p}, \frac{p_n}{p_{n-1}+p_p}\right), \quad \alpha > 0$ [2]. Con $\alpha = 1$ se recupera la recursividad estandar, pero con

---

[1] $\begin{bmatrix} p_1 & \cdots & p_n \end{bmatrix}^t \otimes \begin{bmatrix} q_1 & \cdots & q_m \end{bmatrix}^t = \begin{bmatrix} p_1 q_1 & \cdots & p_1 q_m & \cdots & p_n q_1 & \cdots & p_n q_m \end{bmatrix}^t$.

[2] En sus papel, lo imponen para cualquier pars p_i, p_j sin imponer la invarianza por permutación, pero es equivalente a la exposición

$\alpha \neq 1$ eso permite dar un peso diferente a la incerteza del estado interno *i. e.,* probabilidades que se juntan (la describen como clasificación refinada). Estos axiomas conducen necesariamente a la entropia (teorema 1)

$$H_\alpha^{\mathrm{hc}}(p) = \frac{1}{1-2^{1-\alpha}}\left(1 - \sum_i p_i^\alpha\right)$$

que nombraron α-*entropia structural*. De nuevo, relaxando el axioma (iv), se puede remplazar en el coeficient $2^{1-\alpha}$ por $a^{1-\alpha}, \quad a > 0, \quad a \neq 1$. De nuevo, parae que la entropia de Shannon es un caso particular,

$$H_1^{\mathrm{hc}}(p) \equiv \lim_{\alpha\to 1} H_\alpha^{\mathrm{hc}}(p) = H(p) \quad \text{entropia de Shannon}$$

Luego, proban que $H_\alpha^{\mathrm{hc}}(p)$ es concava con respeto a los p_i y maxima para una distribución uniforma (teorema 2). Aun que no aparece así en el papel, satisface la propiedad de Schur-concavidad (teorema 3). A pesar de que mencionan que H_α^{hc} sea diferente que H_α^{r}, es sencillo ver que hay un mapa uno-uno entre las dos entropias.

Independiente de Havrda & Charvát, todavía en la escuela **checa**, Z. Daróczy en (Daróczy, 1970) defino la entropia H^f a partir de una *función información* f satifaciendo (i) $f(0) = f(1)$, (ii) $f\left(\frac{1}{2}\right) = 1$ y la ecuación funcional (ii) $f(x) + (1-x) f\left(\frac{y}{1-x}\right) = f(y) + (1-y) f\left(\frac{x}{1-y}\right)$ sobre $\{(x,y) \in [0;1)^2, \quad x + y \leq 1\}$, siendo $H^f(p) = \sum_{i=2}^n s_i f\left(\frac{p_i}{s_i}\right), \quad s_i = \sum_{j=1}^{i-1} p_j$. Daróczy mostró que si f es medible, o continua en 0, o no negatiba y acotada, necesariamente $f(x) = h_2(x) = -x\log_2 x - (1-x)\log_2(1-x)$, conduciendo a la entropia de Shannon (teorema 1; ver también (Lee, 1964; Tverberg, 1958)). En otros terminos, su axioma (v) es alternativa a la recursividad. Para extender la entropia de Shannon, propuso extender este axioma (v) por la ecuación funcional $f_\alpha(x) + (1-x)^\alpha f_\alpha\left(\frac{y}{1-x}\right) = f_\alpha(y) + (1-y)^\alpha f_\alpha\left(\frac{x}{1-y}\right)$, lo que condujo necesariamente a la entropia (teoremas 2 y 3)

$$H_\alpha^{\mathrm{d}}(p) = \frac{1}{1-2^{1-\alpha}}\left(1 - \sum_i p_i^\alpha\right)$$

es decir nada mas que la entropia introducida por Havdra & Charvát. Sin embargo, el estudio de Daróczy fue mas intensivo que el de Havdra & Charvát. Primero, notó el mapa entre su entropia y la de Rénty. Luego, probó que se conserva la invarianza por permutación (no era un axioma), $H_\alpha^{\mathrm{d}}\left(\frac{1}{2}, \frac{1}{2}\right) = 1$ (lo llama normalización), la expansividad, una additividad extendida, una recursividad extendida precisamente del modelo de Havrda-Charvát (teorema 4). Probó también que $H_\alpha^{\mathrm{d}} \geq 0$ (alcanzado en el caso deterministico) y máxima en el caso uniforme (teorema 6), que incidentalmente $H_\alpha^{\mathrm{d}}\left(\frac{1}{d}, \ldots, \frac{1}{d}\right)$ crece con d. Muy interesante también es se puede definir una entropia condicional en el mismo modelo que en el caso de Shannon $H_\alpha^{\mathrm{d}}(X|Y) = \sum_y \left[p_{X|Y}(x,y)\right]^\alpha H_\alpha^{\mathrm{d}}(p_{X|Y}(\cdot, y))$, que existe una regla de cadena, $H_\alpha^{\mathrm{d}}(X,Y) = H_\alpha^{\mathrm{d}}(Y) + H_\alpha^{\mathrm{d}}(X|Y)$ y que condicionar reduce la entropia $H_\alpha^{\mathrm{d}}(X|Y) \leq H_\alpha^{\mathrm{d}}(X)$ (teorema 8). Mostró también que si se pierde la additividad, se obiene para X e Y independientes $H_\alpha^{\mathrm{d}}(X,Y) = H_\alpha^{\mathrm{d}}(X) + H_\alpha^{\mathrm{d}}(Y) + \left(2^{1-\alpha} - 1\right) H_\alpha^{\mathrm{d}}(X) H_\alpha^{\mathrm{d}}(Y)$. La propiedades de regla de cadena le permitió revisitar la caracterisación de un canal de transmisión y redefinir una capacidad canal extendidas (capacidad tipo α; basicamente se usa el mismo enfoque que Shannon, pero usando H_α^{d} en lugar de H, ver sección 6 del papel).

---

de este parafo.

redecouverte Tsallis 1988..., Lindhardt & Nielsen Studies in Dynamical Systems Kongelige Danske Videnskabernes Selskab, 38(9):1-42, 1971, De nombreuses autres, Varama 66, mais premier pas unifie Burbea-Rao; puis Salicru

proprietes et voir conditionnelle... moyenne adqueate

versiones dferenciales poner acá la codificación a la Renyi, y la cuantificacion fina; EPI generalizada por Madiman, etc. Lutwak, Bercher etc., Kagan

Revisite capacite a la Daroczy?

1.1.2 Divergencias y propiedades

Extension a la Renyi, a la HC/D/T, Cressie Reads, Cressie Pardo, Vajda; cf Burbea Rao: generalization Czizar et voir avec h phi avant meme Salicru. Cf aussi Bregmann

1.2 Entropias cuanticas discretas

Mas alla caso de informaciones a partir de medida; caso infinito, continuo queda en discusiones

EPILÓLOGO

Este libro surge de la experiencia de los autores en el dictado del curso semestral ”Métodos de geometría diferencial en teoría de la información", que se imparte en la Facultad de Ciencias Exactas de la Universidad Nacional de La Plata y en la Facultad de Matemática, Astronomía y Física de la Universidad Nacional de Córdoba. ...

Los autores

Referencías

Daróczy, Z. (1970). Generalized information functions. *Information and Control*, *16*(1), 36–51.

Fadeev, D. K. (1956). On the concept of entropy of a finite probabilistic scheme (russian). *Uspekhi Matematicheskikh Nauk*, *11*(1(67)), 227–231.

Fadeev, D. K. (1958). *Fundations in Information Theory*, chapter On the concept of entropy of a finite probabilistic scheme (English traduction). New-York: McGraw-Hill.

Hardy, G., Littlewood, J. E., & Pólya, G. (1952). *Inequalities* (2nd ed.). Cambridge, UK: Cambridge University Press.

Havrda, J. & Charvát, F. (1967). Quantification method of classification processes: Concept of structural α-entropy. *Kybernetika*, *3*(1), 30–35.

Khinchin, A. I. (1957). *Mathematical foundations of information theory*. New-York: Dover Publications.

Kolmogorov, A. N. (1930). Sur la notion de la moyenne. *Atti della Reale Accademia Nazionale dei Lincei*, *12*, 388–391.

Kolmogorov, A. N. (1991). On the notion of mean. In V. M. Tikhomirov (Ed.), *Selected Works of A. N. Kolmogorov*, volume I: Mathematics and Mechanics (pp. 144–146). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Lee, P. M. (1964). On the axioms of information theory. *The Annals of Mathematical Statistics*, *35*(1), 415–418.

Nagumo, M. (1930). Über eine klasse der mittelwerte. *Japanese journal of mathematics: transactions and abstracts*, *7*, 71–79.

Rényi, A. (1961). On measures of entropy and information. *in Proceeding of the 4th Berkeley Symposium on Mathematical Statistics and Probability*, *1*, 547–561.

Tribus, M. & McIrvine, E. C. (1971). Energy and information. *Scientific American*, *225*(3), 179–188.

Tverberg, H. (1958). A new derivation of the information function. *Mathematica Scandinavica*, *6*, 297–298.

Los autores

Lamberti, Pedro Walter

Este es un párrafo Normal con texto simulado, (Arial 10, interlineado de 1,5 líneas, sin sangría en la primera línea). Este es un párrafo Normal con texto simulado, (Arial 10, interlineado de 1,5 líneas, sin sangría en la primera línea). Este es un párrafo Normal con texto simulado, (Arial 10, interlineado de 1,5 líneas, sin sangría en la primera línea). Este es un párrafo Normal con texto simulado, (Arial 10, interlineado de 1,5 líneas, sin sangría en la primera línea).

Portesi, Mariela

Este es un párrafo Normal con texto simulado, (Arial 10, interlineado de 1,5 líneas, sin sangría en la primera línea). Este es un párrafo Normal con texto simulado, (Arial 10, interlineado de 1,5 líneas, sin sangría en la primera línea). Este es un párrafo Normal con texto simulado, (Arial 10, interlineado de 1,5 líneas, sin sangría en la primera línea). Este es un párrafo Normal con texto simulado, (Arial 10, interlineado de 1,5 líneas, sin sangría en la primera línea).

Zozor, Steeve

Este es un párrafo Normal con texto simulado, (Arial 10, interlineado de 1,5 líneas, sin sangría en la primera línea). Este es un párrafo Normal con texto simulado, (Arial 10, interlineado de 1,5 líneas, sin sangría en la primera línea). Este es un párrafo Normal con texto simulado, (Arial 10, interlineado de 1,5 líneas, sin sangría en la primera línea). Este es un párrafo Normal con texto simulado, (Arial 10, interlineado de 1,5 líneas, sin sangría en la primera línea).