Entropia e distribuições contínuas

por João Tiago Praça Nunes Mexia

1. Introdução

Em [1] foi já exposta a teoria da entropia para variáveis aliatórias discretas, pelo que nos limitaremos a recordar os principais resultados.

Sejam $A_1 \cdots A_i \cdots$ os acontecimentos elementares dum sistema completo que correspondem aos números observados $a_1 . \cdots a_i \cdots$ e $p_1 \cdots p_i \cdots$ as probabilidades respectivas. A entropia do processo casual em questão é nos dada por:

$$(1) \qquad H = - \sum_{i=1}^{\infty} p_i \log_2 p_i{}^*$$

Onde se faz a convenção de por:

$$(2) \qquad p_i \log_2 p_i = 0 \ \text{ se } \ p_i = 0$$

Seja $F(x)$ a função de distribuição da nossa variável aliatória. Vê-se fàcilmente que se trata duma função em escada com saltos de valor $p_1 \cdots p_i \cdots$ nos pontos $a_1 \cdots a_i \cdots$. Tem-se portanto um funcional $H[F]$ que à função em escada com o conjunto de saltos $p_1 \cdots p_i \cdots$ faz corresponder o valor de H dado por (1).

Em [3] mostra-se que o conjunto das funções de distribuição em escada é denso, para a topologia da convergência uniforme, no espaço das funções de distribuição. É natural portanto que surja a ideia de prolongar por continuidade o funcional $H[F]$ de modo a obter uma teoria geral da entropia.

(*) Foi Wiener que em [4] propôs tomar-se a base 2.

O objectivo desta nota é:

Mostrar que tal prolongamento, por continuidade, é impossível.

2. Considerações heurísticas

Comecemos por considerar o exemplo seguinte:

Seja $f(x)$ uma função de densidade contínua não nula no intervalo $[0, 1]$. Efectuemos uma partição do intervalo $[0, 1]$ em «n» intervalos iguais. Devido ao teorema do valor médio a probabilidade de cada intervalo é:

$$(3) \qquad p_i = \int_{\frac{i-1}{n}}^{\frac{i}{n}} f(x)\,dx = f(x_i)\frac{1}{n}$$

com

$$\frac{i-1}{n} \leqq x_i \leqq \frac{i}{n}$$

Consideremos agora uma variável aliatória discreta que toma um valor a_i no intervalo $\left]\frac{i-1}{n}, \frac{i}{n}\right]$ com a probabilidade p_i. A entropia desta variável aliatória é, devido a (1) e a (3):

$$H_n = - \sum_{i=1}^{n} f(x_i) \cdot \frac{1}{n} \log_2\left[f(x_i) \cdot \frac{1}{n}\right] =$$

$$= - \sum_{i=1}^{n} f(x_i) \cdot \frac{1}{n} \log_2 f(x_i) -$$

$$- \sum_{i=1}^{n} f(x_i) \frac{1}{n} \log_2 \left(\frac{1}{n}\right)$$

Quando o módulo da nossa decomposição do intervalo tende para zero a função de distribuição da nossa variável aliatória discreta tende uniformemente para a função de distribuição que tem $f(x)$ por derivada. Vejamos portanto o limite de H_n quando n tende para $+\infty$.

Por outro lado:

$$\sum_{i=1}^{n} f(x_i) \cdot \frac{1}{n} \cdot \log_2\left(\frac{1}{n}\right) = \\ = \log_2\left(\frac{1}{n}\right) \sum_{i=1}^{n} f(x_i) \cdot \frac{1}{n}$$

Como:

$$\sum_{i=1}^{n} f(x_i) \cdot \frac{1}{n} \underset{n \to +\infty}{\longrightarrow} \int_0^1 f(x)\, dx = 1$$

Tem-se:

$$\log_2\left(\frac{1}{n}\right) \sum_{i=1}^{n} f(x_i) \cdot \frac{1}{n} \underset{n \to \infty}{\longrightarrow} -\infty$$

Verifica-se igualmente que:

$$\lim_{n \to +\infty} \sum_{i=1}^{n} f(x_i) \cdot \frac{1}{n} \cdot \log_2 f(x_i) = \\ = E(\log_2 f(x)),$$

onde $E(\log_2 f(x))$ é a esperança matemática de $\log_2 f(x)$.

Pelo que:

$$H_n \longrightarrow +\infty$$

Actualmente costuma usar-se para o estudo da entropia no caso contínuo a expressão:

$$(4) \qquad \cdots H^1 = E(\log_2 f(x)).$$

As considerações heurísticas que se fizeram não justificam a fórmula (4).

Além disso se formos calcular o valor da entropia para a distribuição uniforme no intervalo $\left[0, \frac{1}{3}\right]$ obtemos:

$$H^1 = -\log_2 3 < 0.$$

Este exemplo mostra-nos que através do uso da fórmula (4) podemos vir a obter entropias negativas.

Se interpretarmos a entropia como uma medida de incerteza seriamos levados a admitir a existência de incertezas negativas o que não é de aceitar.

3. Tratamento rigoroso

As considerações heurísticas do parágrafo precedente necessitam de ser completadas por um tratamento rigoroso do problema.

Vamos mostrar que: *dada uma sucessão qualquer de funções de distribuição em escada que tenda uniformemente para uma função de distribuição contínua a sucessão das suas entropias diverge para $+\infty$ necessàriamente.* Este resultado é, como é fácil de ver, sugerido pelo exemplo que analisamos no parágrafo anterior.

Comecemos por considerar uma função contínua $f(x)$ e uma função $f_1(x)$ que tenha um salto «d» no ponto $\bar{x}$ tem-se:

$$\text{Max}\ \{|f(\bar{x}) - f_1(\bar{x}^+)|, |f(\bar{x}) - f_1(\bar{x}^-)|\} \geqq \frac{d}{2},$$

e portanto

$$\text{Sup}\ |f(x) - f_1(x)| \geqq \frac{d}{2}.$$

Se fôr $\{F_n(x)\}$ uma sucessão de funções de distribuição em escada que tenda uniformemente para uma função de distribuição contínua e se representarmos o maior salto de $F_n(x)$ por «s_n» terá de se ter:

$$(7) \qquad s_n \underset{n \to +\infty}{\longrightarrow} 0.$$

Por outro lado quaisquer que sejam os inteiros «a» e «b» tais que $a < b$ tem-se, desde que os p_i sejam positivos e menores que a unidade:

$$(8)\quad -\sum_{i=a}^{b} p_i \log_2 p_i \geqq -\left(\sum_{i=a}^{b} p_i\right)\log_2\left(\sum_{i=a}^{b} p_i\right),$$

pois, qualquer que seja i:

$$\log_2 p_i \leqq \log_2\left(\sum_{i=a}^{b} p_i\right).$$

Seja $\{p_i^{(n)}\}$ o conjunto dos saltos de $F_n(x)$. Devido a (7) tem-se que qualquer que seja o natural m existe um natural N tal que para n maior que N tem-se, qualquer que seja i:

$$(9)\qquad p_i^{(n)} < \frac{1}{2m}.$$

Por outro lado como, qualquer que seja n, $\sum_{i=1}^{\infty} p_i^{(n)}$ é uma série absolutamente convergente de soma 1 tem-se que para n maior que N qualquer que seja o natural «j» compreendido entre «o» e «m» existe um natural $N^{(n)}j$ tal que:

$$(10)\qquad \sum_{i=1}^{N_j^n} p_i^n \leqq \frac{j}{m} \quad \text{e} \quad \sum_{i=1}^{N_j^n+1} p_i^n > \frac{j}{m}$$

tem-se ainda

$$(11)\qquad \sum_{i=1}^{N_j^n+1} p_i^n < \frac{2j+1}{2m} \quad \text{pois} \quad p_{N_j^n+1}^n < \frac{1}{2m}.$$

Donde:

$$(12)\qquad \sum_{i=N_j^n+2}^{N_{j+1}^n+1} p_i^n = \sum_{i=1}^{N_{j+1}^n+1} p_i^n - \sum_{i=1}^{N_j^n+1} p_i^n > \frac{1}{2m},$$

devido a (11) e a (10).

Por outro lado é fácil de ver que a função $-x\log_2 x$ é crescente no intervalo $]0, e^{-1}[$.

Donde, qualquer que seja «m» maior do que 2, existe pelo menos um N tal que n maior que N implica:

$$H[F_n] = -\sum_{i=1}^{\infty} p_i^n \log_2 p_i^n \geqq \sum_{i=1}^{N_{m-1}^n} p_i^n \log_2 p_i^n \geqq$$
$$\geqq -\sum_{j=1}^{m-1}\left(\sum_{i=N_j^n+2}^{N_{j+1}^n+1} p_i^n\right)\log\left(\sum_{i=N_j^n+2}^{N_{j+1}^n+1} p_i^n\right) \geqq$$
$$\geqq \frac{m-1}{2m}\log_2 m$$

A segunda desigualdade resulta de (8) e a terceira de (12) e do facto da função $-x\log_2 x$ ser crescente no intervalo $]0, e^{-1}[$. Donde

$$(13)\qquad H[F_n] \to +\infty$$

4. Observações finais

Interessa ainda que sejam feitas as seguintes observações finais:

1. Observe-se que em nenhum passo da demonstração final do parágrafo precedente se teve de recorrer à existência duma infinidade numerável de saltos não nulos em algumas das funções $F_n(x)$; logo *a demonstração continua válida se supusermos que todas as funções* $F_n(x)$ *eram funções de distribuição simples*, isto é, tais que tomam apenas um número finito de valores.

2. Como em V a convergência uniforme implica a convergência pontual, $H[F]$ *não é prolongável por continuidade a partir da topologia da convergência pontual.*

3. Como vimos em III não se pode prolongar por continuidade o funcional que nos dá a entropia associada a funções de distri-

buição discretas. Por outro lado é impossível arranjar uma sucessão de funções de distribuição contínuas que tenda uniformemente para uma função de distribuição discreta. Portanto uma teoria da entropia para o caso contínuo que venha a ser formulada terá de ser construída independentemente da teoria já existente da entropia no caso contínuo.

É este um assunto a que esperamos poder dedicar um estudo futuro.

Résumé

L'auteur fait quelques considérations sur la possibilité de prolonger par continuité le fonctionnel qui donne les valeurs de l'entropie dans le cas discret. Il montre nottament qui étant donnée une suite $\{F_n\}$ de fonctions discrétes qui convergent uniformément pour une fonction de distribution continue, la suite des valeurs de l'entropie pour les fontions F_n diverge nécessairement pour $+\infty$. Donc le prolongement par continuité ne peut point être fait à partir de la topologie de la convergence uniforme.

L'auteur de cet article est convaincu qu'une possible future théorie de l'entropie pour le cas continu devrait être construite indépendamment de la théorie existente pour le cas discret.

BIBLIOGRAFIA

[1] Dionisio, J. J. *A definição de entropia em cálculo das probabilidades:* Gazeta de Matemátida: 74-75, 1959.

[2] Khinchin, A. I. *Mathematical Foundations of information theory:* Dover Publications — 1957.

[3] Loéve, M. *Probability theory:* Van Nostraud, 2nd edd — 1960.

[4] Wiener, N. *Cybernetics*: Herman Editeur—1949.

Sur l'introduction des mathématiques modernes dans l'enseignement secondaire (1)

par J. Sebastião e Silva

Considérations générales.

Nous sommes d'accord qu'il faut introduire, dans l'enseignement secondaire, certains sujets des mathématiques modernes et, surtout, l'esprit de ces mathématiques, non seulement pour assurer une meilleure formation intellectuelle des élèves, mais aussi pour éviter une désarticulation, qui devient de plus en plus sensible, entre les études secondaires et celles universitaires. Mais nous faisons à cela des réserves.

Nous pensons que ces innovations doivent être éxécutées avec une extrême prudence et le plus fin tact pédagogique, si l'on ne veut pas créer chez les élèves une répulsion invincible pour les mathématiques ou les conduire à l'acquisition d'un formalisme vide, tout à fait stérilisant. En effet, la moderne orientation abstraite des mathématiques est une épée à deux tranchants, d'après l'usage que l'on en fait: elle peut rendre l'enseignement beaucoup plus attirant et beaucoup plus efficace; mais, mal appliquée, elle peut aussi conduire à des résultats à peu près opposés.

En conséquence, nous proposons un enseignement qui, au moins dans les cinq premières années, reste classique dans ses lignes générales, mais qui soit fortement influencé

(1) Este relatório foi publicado pela revista italiana «Archimede». Convém salientar que o ponto de vista aqui expresso figura entre os meios moderados que têm sido ùltimamente defendidos.