Proceeding Series of the Brazilian Society of Applied and Computational Mathematics, Vol. 3, N. 1, 2015.
Trabalho apresentado no XXXV CNMAC, Natal-RN, 2014.

Quantificação de informação na teoria de Bayes em escalas temporais

Igor de S. Barbalho **Lucas Ragiotto**
Depto de Matemática, FEIS, UNESP
15385-000, Ilha Solteira, SP
E-mails: igorsouza4@gmail.com lucasragiotto@gmail.com

Berenice C. Damasceno **Luciano Barbanti**
Depto. de Matemática, FEIS, UNESP
15385-000, Ilha Solteira, SP
E-mails: berenice@mat.feis.unesp.br barbanti@mat.feis.unesp.br

RESUMO

Medida de "surpresa"

A teoria da informação permeia todos os campos da ciência. Um dos primeiros problemas que a teoria pressupõe é como quantificar a informação de um conjunto de dados.

Shannon em 1948 quantificou a informação de um conjunto de dados C, dentro de uma mensagem M, como sendo $\log_2 \mathrm{P_M}(\mathrm{C})$.

Na década de 2000 Baldi e Itti [2], diante dos conceitos envolvidos na teoria de Bayes propuseram outra definição para a quantidade de informação que renomearam como uma medida de "surpresa" (do observador).

De acordo com o subjacente psicológico na teoria em Bayes este efeito é o que provoca a transformação das crenças *à priori* para as crenças *à posteriori*. Neste sentido vemos, então, que esta quantidade de informação pode ser medida pela distância -entropia relativa - entre as distribuições *à priori* e *à posteriori* do observador sobre os espaços de hipóteses disponíveis.

Chamando de $\mathcal{M}$ o espaço de hipóteses disponíveis com $M \epsilon \mathcal{M}$ esta medida é então relacionada com $\frac{\mathrm{P(D|M)}}{\mathrm{P(D)}}$ de acordo com o teorema de Bayes: $\mathrm{P(M|D)}=\frac{\mathrm{P(D|M)}}{\mathrm{P(D)}}\mathrm{P(M)}$.

Conforme se depreende da equação o efeito de D é a mudança de $\mathrm{P(M)}$ para $\mathrm{P(M|D)}$.

Isto vale dizer que vemos o dado D como um operador agindo no espaço de distribuições sobre o espaço dos modelos. Então um dos modos de quantificar a informação, (medida de surpresa) fornecida por D é a medida da distância , $\log_2(\frac{\mathrm{P(D|M)}}{\mathrm{P(D)}})$, entre as distribuições *à priori* e *à posteriori*.

Probabilidade em escalas temporais

O cálculo em escalas temporais foi introduzido em 1988 por Stefan Hilger [1], com o intuito de unificar as teorias de sistemas, tanto contínuo quanto discreto.

Uma escala temporal $\mathbb{T}$ é um conjunto fechado não-vazio dos números reais. Exemplos de escalas temporais são $\mathbb{R}$, $\mathbb{Z}$, $\{\frac{1}{2^k}\ ; k \in \mathbb{N}\ \}$, ou o conjunto de Cantor.

Os operadores fundamentais desta teoria unificante são:

$\sigma(\mathrm{t})=\inf\{\mathrm{r;r>t}\}$ - operador de avanço

$\rho(t) = sup\ \{r; r < t\}$ - operador de retardo,

A Δ- *derivada é*

$$\mathrm{f}^{\Delta}(\mathrm{t})=\begin{cases} \mathrm{f}'(\mathrm{t}) & \text{se } \sigma(\mathrm{t})=\mathrm{t} \\ \frac{\mathrm{f}(\sigma(\mathrm{t}))-\mathrm{f}(\mathrm{t})}{\sigma(\mathrm{t})-\mathrm{t}} & \text{se } \sigma(\mathrm{t})\neq \mathrm{t} \end{cases}$$

As escalas temporais são próprias para interpretar Probabilidade, pois as variáveis aleatórias e as funções de distribuição podem ser descritas tanto no caso contínuo, como no discreto.

Seja uma escala temporal $\mathbb{T}$ e $A \subset \mathbb{T}$. Então A tem a decomposição $A = \bigcup_{i=1}^{m}[a_i,b_i] \cup \{t_1,t_2,...t_n\}$, onde $\tau_i=\sigma(\tau_i)=\rho(\tau_i)$, $\forall \tau_i \in (a_i,b_i)$ e t_j -pontos isolados em $\mathbb{T}$.

A Δ-medida de Lebesgue [3] , se $A \neq \emptyset$,é

$$\mu_\Delta(A) = \sum_{i=1}^{m} (\sigma(b_i)\text{-}a_i) + \sum_{j=1}^{n} (\sigma(t_j)\text{-}t_j)$$

e $\mu_\Delta(\emptyset)=0$.

Definição1: Seja uma escala temporal $\mathbb{T}$ e $\mathfrak{C}_\mathbb{T}$ uma álgebra de subconjuntos de $\mathbb{T}$, e P_Δ uma Δ-medida sobre $\mathfrak{C}_\mathbb{T}$.Se $P_\Delta(\mathbb{T}) = 1$, então P_Δ é uma medida de probabilidade. Nestas condições a tripla { $\mathbb{T}$, $\mathfrak{C}_\mathbb{T}$, P_Δ} é chamada um Δ-*espaço de probabilidade*.

Finalmente podemos definir a Δ-probabilidade de A, assim:

Definição2: Seja $\Omega_\mathbb{T}$ um espaço amostral e $A \subset \Omega_\mathbb{T}$. Então $P_\Delta(A)=\frac{\mu_\Delta(A)}{\mu_\Delta(\Omega_T)}$ é a Δ -probabilidade de A. Como exemplo podemos considerar $\emptyset \neq A \subset \Omega_\mathbb{T} \subset \mathbb{N}$. Então $P_\Delta(A)=\frac{\#(A)}{\#(\Omega_T)}$, como usualmente.

Objetivo e primeiros resultados

O objetivo principal deste trabalho é a extensão dos resultados de Baldi e Itti na teoria usando a medida “surpresa” para escalas temporais, e sua aplicação em experimentos de Marketing e em Economia no mercado BOVESPA.

Os primeiros resultados preliminares obtidos de imediato, entre outros, podem ser elencados:

Proposição 1: A probabilidade $P_\Delta(K|L) = \frac{\mu_\Delta(K\cap L)}{\mu_\Delta(L)}$, supondo $\mu_\Delta(L) \neq \emptyset$.

Proposição 2: Uma medida de “surpresa” neste contexto $\log_2\left(\frac{P(D|M)}{P(D)}\right)=\log_2\left(\frac{\mu_\Delta(M\cap D)\mu_\Delta(\Omega_T)}{\mu_\Delta(M)\mu_\Delta(\mu_\Delta(D))}\right)$.

Palavras-chave: *Teorema de Bayes, Teoria da Informação, Escalas Temporais, Medida de Surpresa.*

Referências

[1] S. Hilger, Analysis on measure chains - a unified approach to continuous and discrete calculus, Results Math.vol.18,pp.18-56, (1990).

[2] L. Itti & P. F. Baldi, Bayesian surprise attracts human attention. Vision Research. 49(10), 1295-1306, (2009).

[3] U. Ufuktepe, A. Deniz, Lebesgue-Stieltjes measure on time scales. Turk. J. Math. 32, 1-8 (2009).

