MATHEMATICA—REVUE D'ANALYSE NUMÉRIQUE
ET DE THÉORIE DE L'APPROXIMATION

**L'ANALYSE NUMÉRIQUE ET LA THÉORIE DE L'APPROXIMATION
Tome 6, N° 2, 1977, pp. 155–161**

THE PROBLEMS OF CALCULUS AND APPROXIMATION OF THE AMOUNT OF INFORMATION

by

ION MIHOC

(Cluj-Napoca)

1. Let $\{\Omega, K, P\}$ be a probability space generated by a experiment A and let $A_1, A_2, \ldots, A_N$ be the possible autcomes of random experiment A.

If the set $\{A_1, A_2, \ldots, A_N\}$ formes a partition Π for the sure event Ω, then [2], [4]

$$H(\Pi) = H(X) = -k \sum_{i=1}^{N} p_i \log_e p_i \tag{1.1}$$

where

$$p_i = P(A_i) \geqslant 0,\ i = 1, 2, \ldots, N;\ \sum_{i=1}^{N} p_i = 1,\ k = \log_2 e, \tag{1.2}$$

represents the amount of information furnished by the experiment A.

The quantity (1.1) is called the entropy of the partition Π and it measures either the uncertainty of the experiment A, if this experiment not yet performed or the amount of information of the experiment A, if this experiment has been performed. Also, we can speak as well that the quantity (1.1) represents the amount of information contained by the random variable X generated by the experiment A.

2. Application of Taylor's series to evaluation of H(X)

THEOREM 1. *If X is a discret random variable then*

$$H(X) = -\frac{k}{n}\left\{(n-N) - n\cdot\log_e n + \sum_{i=1}^{N}\sum_{s=2}^{\infty}\frac{1}{s^{(2)}}\sum_{t=0}^{s} C_s^t(-np_i)^t\right\}, \tag{2.1}$$

where n is a natural number.

Proof. Consider

$$p_i = \pi_i + \frac{1}{n},\ i = 1, 2, \ldots, N, \tag{2.2}$$

where $\pi_i \in \left(-\frac{1}{n}, \frac{1}{n}\right]$ then when $p_i \in (0, 1]$.

Because the function $\log_e p_i$ is infinite derivable in the neighbourhood of the point $\frac{1}{n}$, it follows that the function

$$f(n\pi_i) = \log_e\left(\pi_i + \frac{1}{n}\right) = \log_e(1 + n\pi_i) - \log_e n \tag{2.3}$$

admits the expansion in the power series

$$f(n\pi_i) = -\log_e n + (n\pi_i) - \frac{(n\pi_i)^2}{2} + \frac{(n\pi_i)^3}{3} - \ldots, \tag{2.4}$$

where the power series

$$\log_e(1 + n\pi_i) = (n\pi_i) - \frac{(n\pi_i)^2}{2} + \ldots + (-1)^{s-1}\frac{(n\pi_i)^s}{s} + \ldots \tag{2.5}$$

is convergent for $|n\pi_i| < 1$, respectively, for $|\pi_i| < \frac{1}{n}$.

Therefore, the function $p_i \log_e p_i$ has the following expansion

$$p_i \log_e p_i = -\frac{1 + n\pi_i}{n}\log_e n + \frac{1}{n}\left\{n\pi_i + \sum_{s=2}^{\infty}(-1)^s\frac{(n\pi_i)^s}{s^{(2)}}\right\}, \tag{2.6}$$

where $x^{(\nu)} = x(x-1)(x-2)\ldots(x-\nu+1)$.

Having in view that

$$\sum_{i=1}^{N} np_i = n - N \tag{2.7}$$

$$(-1)^s(n\pi_i)^s = (1 - np_i)^s = \sum_{t=0}^{s} C_s^t(-np_i)^t, \tag{2.8}$$

and the development (2.6), we obtain for $H(X)$ the following form

$$H(X) = -\frac{k}{n}\left\{(n - N) - n\cdot\log_e n + \sum_{i=1}^{N}\sum_{s=2}^{\infty}\frac{1}{s^{(2)}}\sum_{t=0}^{s} C_s^t(-np_i)^t\right\}. \tag{2.9}$$

Remark 1.2. Because the series (2.5) is convergent both for $|n\pi_i| < 1$ and for $n\pi_i = 1$, it follows that the Taylor's series which correspond to the function $\log_e p_i$, will be convergent if $p_i \in \left(0, \frac{2}{n}\right]$. This series will be convergent for $p_i \in (0, 1]$ only if $n = 2$.

C o r o l l a r y 2.1. *If $n = 2$, then $H(X)$ has the form*

$$H(X) = 1 - \frac{k}{2}\left\{(2 - N) + \sum_{i=1}^{N}\sum_{s=2}^{\infty}\frac{1}{s^{(2)}}C_s^t(-2p_i)^t\right\} \tag{2.10}$$

where $k = \log_2 e$.

Remark 2.2. According to the Remark 2.1 it seems that the formula (2.9) is available only for $n = 2$, then when $p_i \in (0, 1]$.

If we have in view the condition (1.2) it follows that among the probabilities $p_1, p_2, \ldots, p_N$ exist al the most a probability p_h so that $p_h \geq \frac{1}{2}$, $(N > 2)$.

If the probability p_h is sufficient near to one, then, evident, the others probabilities $p_i, i = 1, 2, \ldots, N, i \neq h$, will be situated sufficient near to zero. In other words, if $p_h \to 1$, then is possible to find a natural number n so that all probabilities p_i, $i = 1, 2, \ldots, N$; $i \neq h$, to be situate in the interval $\left(0, \frac{2}{n}\right]$.

C o r o l l a r y 2.2. *If X is a discret random variable which satisfies the conditions*

$$p_i > 0,\ i = 1, 2, \ldots, N;\ \sum_{i=1}^{N} p_i = 1$$

and if

1^0 $p_h = \max\{p_1, p_2, \ldots, p_N\}$, $p_h \in [1 - \varepsilon, 1]$, $\varepsilon > 0$,

2^0 *there is a natural number n so that $p_i \in \left(0, \frac{2}{x}\right]$, $i = 1, 2, \ldots, N$; $i \neq h$, then, for the measure $H(X)$, we have the following form*

$$H(X) = -\frac{k}{n}\left\{(1 - N) + np_h \log_e p_h + n(1 - p_h)\,[\log_e n - 1] + \sum_{i=1}^{N}\sum_{s=2}^{\infty}\frac{1}{s^{(2)}}\sum_{t=0}^{s} C_s^t(-np_i)^t\right\}. \tag{2.11}$$

Corollary 2.3. *If $n = N$, N — the number of the possible values of discret random variable X, then the amount of information, $H(X)$, has the form*

$$H(X) = \log_2 N - \frac{k}{N}\left\{\sum_{i=1}^{N}\sum_{s=2}^{\infty}\frac{1}{s^{(2)}}\sum_{t=2}^{s} C_s^t(-Np_i)^t\right\}. \tag{2.12}$$

More, if X_1 is a random variable uniformly distributed, then $H(X_1) = \log_2 N$.

Corollary 2.4. *For any random variable X, we have*

$$H(X) \leq \log_2 N = H(X_1), \tag{2.13}$$

where X_1 is uniformly distributed.

Proof. This corollary is a fundamental property of the measure $H(X)$, [1]. In ours case, the proof of the inequality (2.13) to came back to show that $D \geqq 0$, where

$$D = \sum_{i=1}^{N}\sum_{s=2}^{\infty}\frac{1}{s^{(2)}}\sum_{t=0}^{s} C_r^t(-Np_i)^t = \sum_{i=1}^{N}\sum_{s=2}^{\infty}(-1)^s\frac{(N\pi_i)^s}{s^{(2)}}. \tag{2.14}$$

Indeed, the sign of D depends by the sign of power series

$$\sum_{s=2}^{N}(-1)^s\frac{(N\pi_i)^s}{s^{(2)}},$$

which, for $|N\pi\ | < 1$, is convergent and consequently, his sum, $S(N, \pi_i)$, satisfies the condition

$$0 < S(N\pi_i) < \frac{(N\pi_i)^2}{1 \cdot 2}.$$

This last inequality proves Corallary 2.4.

3. The Aproximation of H(X)

It is know that if the function $f(x)$ has the derivates tell the order $m + 1$ inclusively, in the vecinity of the point x_0, then it is expandable in Taylor series [3] and we have

$$f(x) = f(x_0) + \frac{x - x_0}{1!}f'(x_0) + \ldots + \frac{(x - x_0)^m}{m!}f^m(x_0) + r_m(x), \tag{3.1}$$

where

$$r_m(x) = \frac{(x - x_0)^{m+1}}{(m+1)!}f^{m(+1)}[x + \theta(x - x_0)], \quad 0 < \theta < 1, \tag{3.2}$$

is Lagrange's form of the remainder.

THEOREM 3.1. *Let X be a discret random variable. Then the measure, $H(X)$, of the amount of information of X, has the form*

$$H(X) = H_m(X) + R_m(n\pi_i), \tag{3.3}$$

where

$$H_m(X) = -\frac{k}{n}\left\{(n - N) - n\log_e n + \sum_{i=1}^{N}\sum_{s=2}^{m}(-1)^s\frac{(n\pi_i)^s}{s^{(2)}}\right\}, \tag{3.4}$$

$$R_m(n\pi_i) = (-1)^m\frac{k}{n}\sum_{i=1}^{N}(n\pi_i)^{m+1}\left[\frac{1}{m} - \frac{1}{m+1}\cdot\frac{1 + n\pi_i}{1 + \theta n\pi_i}\right], \tag{3.5}$$

$$0 < \theta < 1, \text{ and } n\pi_i = np_i - 1.$$

Proof. As an application of the formula (3.1), let us expand into the power series the function $p_i \log_e p_i$.

We have

$$p_i\log_e p_i = -\frac{1 + n\pi_i}{n}\log_e n + \frac{1}{n}\left\{n\pi_i + \frac{(n\pi_i)^2}{1 \cdot 2} - \frac{(n\pi_i)^3}{2 \cdot 3} + \ldots + (-1)^m\frac{(n\pi_i)^m}{(m-1)\cdot m} + \bar{r}_m(n\pi_i)\right\}, \tag{3.6}$$

where

$$\bar{r}_m(n\pi_i) = (-1)^{m-1}(n\pi_i)^{m+1}\left[\frac{1}{m} - \frac{1}{m+1}\cdot\frac{1 + n\pi_i}{1 + \theta n\pi_i}\right]. \tag{3.7}$$

The expansion (3.6) with the remainder (3.7), give us just the form (3.3) of the measure $H(X)$.

Corollary 3.1. *Because $n\pi_i = np_i - 1$, it follows that the measure $H(X)$ has and the form*

$$H(X) = -\frac{k}{n}\left\{(n - N) - n\log_e n + \sum_{i=1}^{N}\sum_{s=2}^{m}\frac{1}{s^{(2)}}\sum_{t=0}^{N} C_s^t(-np_i)^t\right\} + R_m(np_i - 1) \tag{3.8}$$

where

$$R_m(np_i - 1) = -\frac{k}{n}\sum_{i=1}^{N}\sum_{t=0}^{m+1} C_{m+1}^t(-np_i)^t\left[\frac{1}{m} - \frac{1}{m+1}\cdot\frac{np_i}{1 + \theta(np_i - 1)}\right], \tag{3.9}$$

$$0 < \theta < 1$$

Remark 3.1. If X is a random variable uniformly distributed, then $H(X)$, given in the relation (3.8), to came return to $H(X) = \log_2 N$.

C o r o l l a r y 3.2. *The relation* (3.3) *give us the possibility to approximate the measure* $H(X)$ *in the following form*

$$(3.10)\qquad H(X) \approx H_m(X) = -\frac{k}{n}\left\{(n-k) - n\log_e n + \sum_{i=1}^{N}\sum_{s=2}^{m}(-1)^s\frac{(n\pi_i)^s}{s^{(2)}}\right\}.$$

The error in this case is just the value of the remainder (3.5).

Remark 3.2. The remainder $R_m(n\pi_i)$, given in the relation (3.5), can be write in the following form

$$(3.11)\qquad R_m(n\pi_i) = (-1)^n\frac{k}{n}\sum_{i=1}^{N}(n\pi_i)^{m+1}\left[\frac{1}{m} - \frac{1}{m+1}\cdot g(\theta, \pi_i)\right],$$

where the function

$$(3.12)\qquad g(\theta, \pi_i) = \frac{1+n\pi_i}{1+\theta n\pi_i},\quad (i = 1, 2, \ldots, N,$$

is defined for $\theta \in (0, 1)$ and $n\pi_i = np_i - 1$.

Now, insted of the function $g(\theta, \pi_i)$, we introduce a new function namely

$$(3.13)\qquad \bar{g}(\theta, p_i) = \frac{np_i}{1+\theta(np_i-1)} = \begin{cases} 1\ ,if\ \theta = 1,\ p_i \in (0,1] \\ np_i = \begin{cases} 0,\ if\ \theta = 0,\ p_i = 0 \\ n,\ if\ \theta = 0,\ p_i = 1 \end{cases} \\ g(\theta,\ p_i),\ if\ \theta \in (0, 1) \end{cases}$$

where $n\pi_i = np_i - 1$, $i = 1, 2, \ldots, N$.

If $\theta = 1$, then for any $p_i \in (0,1]$, respectively, for any $\pi_i \in \in \left[-\frac{1}{n}, \frac{1}{n}\right]$, we have $\bar{g}(1, p_i) = 1$. Making this change the remainder $R_m(n\pi_i)$ can be written as

$$(3.14)\qquad \bar{R}_m(n\pi_i) = (-1)^m\frac{k}{n}\sum_{i=1}^{N}\frac{(n\pi_i)^{m+1}}{m(m+1)}.$$

Now, replacing in the relation (3.3), the remainder $R_m(n\pi_i)$ through the new form (3.14), we obtain the following approximation for $H(X)$.

$$(3.15)\qquad H(X) \approx H_{m+1}(X) = -\frac{k}{n}\left\{(n-N) - n\log_e n + \sum_{i=1}^{N}\sum_{s=2}^{m+1}\frac{1}{s^{(2)}}\sum_{t=2}^{m}C_s^t(-np_i)^t\right\}.$$

R E F E R E N C E S

[1] F e i n s t e i n, A., *Foundations of information theory*, New York, 1968.
[2] G u i a ş u, S., T h e o d o r e s c u, R., *Teoria matematică a informaţiei*, Editura Acad. R.S.R., Bucureşti, 1966.
[3] P i s k o u n o v, N., *Calcul différentiel et intégral*, Tome 1, Editions MIR, Moscou, 1970.
[4] S h a n n o n, C. E., *A mathematical theory of communications*, Bell System Tech. J., **27**, 379–423; 625–656 (1948).

Received 8. VI. 1977.

Universitatea Babeş–Bolyai
Facultatea de ştiinţe economice
Cluj-Napoca