MATHEMATICA — REVUE D'ANALYSE NUMÉRIQUE
ET DE THÉORIE DE L'APPROXIMATION
L'ANALYSE NUMÉRIQUE ET LA THÉORIE DE L'APPROXIMATION
Tome 20, N^os 1—2, 1991, pp. 69—75

NOTE SUR L'ENTROPIE D'UNE DISTRIBUTION CONTINUE

ION PURCARU
(Bucarest)

Résumé. Le but de cette note est de donner des expressions d'approximation de l'entropie pour quelques familles de distributions continues d'où, par particularisation, on déduit des résultats correspondant aux distribution classiques.

Soit X une variable aléatoire continue à la distribution probabilistique définie par $f(x) \geqslant 0$, $x \in D \subseteq R$, $\int_D f(x)\,\mathrm{d}x = 1$. Supposons que l'entropie Shannon de X ou de $f(x)$ définie par

$$H(X) = -\int_D f(x) \ln f(x)\,\mathrm{d}x = H(f) \tag{1}$$

est connue (voir [1], [2], [4]). On sait que, généralement, même pour les distributions probabilistiques classiques, on ne peut pas donner des expressions bien déterminées de l'entropie (1) et, par conséquent, d'un cas à l'autre, il faut l'approximer. C'est le but de cette note. On supposera connues les fonctions gamma et bêta d'Euler (voir [3] et [5]) utilisées parfois dans ce qui suit.

PROPOSITION 1. *Si*

$$f(x) = \frac{na^m}{2\,\Gamma(m)}\, |x|^{mn-1}\, \mathrm{e}^{-a|x|^n}, \quad x \in R \tag{2}$$

où $a > 0$, $m > 0$, $n > 0$ et $\Gamma(m)$ est la fonction gamma d'Euler, alors

$$H(X) = \frac{m(1-mn)}{n} \sum_{k=1}^{\infty} \frac{1}{k(k+m)} + \frac{(1+m\gamma)(mn-1)}{mn} + m + \ln \frac{2}{n}\, \Gamma(m)\, a^{-\frac{1}{n}} \tag{3}$$

où γ est la constante d'Euler ($\gamma \approx 0{,}5772$).

Preuve. Si on introduit (2) en (1), alors par calcul direct on trouve sans difficulté

$$H(X) = \ln \frac{2\,\Gamma(m)}{na^m} - \frac{n(mn-1)a^m}{\Gamma(m)} \int_0^\infty x^{mn-1}\,e^{-ax^n} \ln x\, dx + m \tag{4}$$

Dénotons par

$$F(a, m, n) = \int_0^\infty x^{mn-1}\,e^{-ax^n}\, dx = \frac{\Gamma(m)}{n}\, a^{-m} \tag{5}$$

Des expressions (5) nous allons obtenir

$$\frac{\partial F}{\partial m} = n \int_0^\infty x^{mn-1}\,e^{-ax^n} \ln x a\, dx = \frac{1}{n}\, a^{-m}\,(\Gamma'(m) - \Gamma(m) \ln a) \tag{6}$$

où $\Gamma'(m)$ est la dérivée de la fonction $\Gamma(m)$ dont la convergence est également assurée pour $m > 0$ (voir [5]).
Par conséquent de (6) l'expression (4) sera

$$H(X) = \ln \frac{2\,\Gamma(m)}{n}\, a^{-\frac{1}{n}} + m - \frac{mn-1}{n}\, \frac{\Gamma'(m)}{\Gamma(m)} \tag{7}$$

En considérant la formule de Weierstrass (voir [5]).

$$me^{\gamma m} \prod_{k=1}^\infty \left(\left(1 + \frac{m}{k}\right) e^{-\frac{m}{k}} \right) \Gamma(m) = 1 \tag{8}$$

on déduit sans difficulté la relation

$$\Gamma'(m) = -\left(\frac{1}{m} + \gamma - m \sum_{k=1}^\infty \frac{1}{k(k+m)} \right) \Gamma(m) \tag{9}$$

et, par conséquent, de (7) et (9) nous allons obtenir (3) et l'affirmation de la proposition est vérifiée.

Remarque 1. Si dans (2) on prend $m = \frac{1}{2}$, $n = 2$ et $a = \frac{1}{2\sigma^2}$, $\sigma > 0$ alors on a la distribution normale $N(0, \sigma)$ et $H(X) = \ln \sigma\sqrt{2\pi e}$

Remarque 2. Si dans (2) on prend $m = n = 1$, alors on a la distribution de Laplace et $H(X) = \ln \frac{2e}{a}$.

Analoguement à la proposition 1 on peut formuler et déduire la

Proposition 2. *Si*

$$f(x) = \frac{na^m}{\Gamma(m)}\, x^{mn-1}\, e^{-ax^n}, \quad x \geqslant 0 \tag{10}$$

où $a > 0$, $m > 0$, $n > 0$, et $\Gamma(m)$ est la fonction gamma d'Euler alors

$$H(X) = \frac{m(1-mn)}{n} \sum_{k=1}^\infty \frac{1}{k(k+m)} + \frac{(1+m\gamma)(mn-1)}{mn} + m + \ln \frac{\Gamma(m)\, a^{-\frac{1}{n}}}{n} \tag{11}$$

Remarque 1. Si dans (10) on prend $m = n = 1$, alors on a la distribution exponentielle de paramètre a et $H(X) = 1 - \ln a$.

Remarque 2. Si dans (10) on prend $n = 1$, alors on a la distribution gamma généralisée et de (11) on déduit

$$H(X) = m(1-m) \sum_{k=1}^\infty \frac{1}{k(k+m)} + \frac{(1+m\gamma)(m-1)}{m} + m + \ln \frac{\Gamma(m)}{a} \tag{12}$$

d'où pour $a = 1$ on trouve l'entropie pour la distribution gamma simple.

Remarque 3. Si dans (10) on prend $n = 1$, $m = \frac{p}{2}$, $p \in N^*$ et $a = \frac{1}{2\sigma^2}$, $\sigma > 0$, alors on a la distribution $\chi_p^2(\sigma)$ et de (11) on déduit que

$$H(X) = \frac{p(2-p)}{2} \sum_{k=1}^\infty \frac{1}{k(2k+p)} + \frac{(p\gamma+2)(p-2)}{2p} + \frac{p}{2} + \ln 2\sigma^2\, \Gamma\left(\frac{p}{2}\right) \tag{13}$$

Remarque 4. Si dans (10) on prend $m = 1$, alors on a la distribution de Weibull et de (11) on déduit que

$$H(X) = 1 + \frac{n-1}{n}\gamma - \ln na^{\frac{1}{n}} \tag{14}$$

Proposition 3. *Si*

$$f(x) = \frac{(m+1)\, a^p}{2pB(p, q)}\, |x|^m (1 + a|x|^{\frac{m+1}{p}})^{-p-q}, \quad x \in R \tag{15}$$

où $a > 0,\ m > -1,\ p > 0,\ q > 0$ *et* $B(p,q)$ *est la fonction bêta d'Euler, alors*

$$H(X) = \ln \frac{2pB(p,q)}{m+1} a^{-\frac{p}{m+1}} + \frac{mpq}{m+1} \sum_{k=0}^{\infty} \frac{1}{(k+p)(k+p+q)} +$$

(16) $$+ \frac{(m+1)pq + p^2}{m+1} \sum_{k=0}^{\infty} \frac{1}{(k+q)(k+p+q)}$$

Preuve. Si on introduit (15) en (1) et si on fait le changement de variable

(17) $$ax^{\frac{m+1}{p}} = t(1-t)^{-1}$$

alors, par calcul direct, on trouve

$$H(X) = \ln \frac{2pB(p,q)}{m+1} a^{-\frac{p}{m+1}} - \frac{mp}{(m+1)B(p,q)} \int_0^1 t^{p-1}(1-t)^{q-1} \ln t\,dt -$$

(18) $$- \frac{p+(m+1)q}{(m+1)B(p,q)} \int_0^1 t^{p-1}(1-t)^{q-1} \ln(1-t)\,dt$$

Sachant que

(19) $$B(p,q) = \int_0^1 t^{p-1}(1-t)^{q-1}\,dt, \quad p > 0,\ q > 0$$

on observe que

$$H(X) = \ln \frac{2pB(p,q)}{m+1} a^{-\frac{p}{m+1}} - \frac{mp}{(m+1)B(p,q)} \frac{\partial B(p,q)}{\partial p} -$$

(20) $$- \frac{p+(m+1)q}{(m+1)Bp,q)} \frac{\partial B(p,q)}{\partial q}$$

où $\frac{\partial B}{\partial p}$ et $\frac{\partial B}{\partial q}$ sont le dérivées partielles de la fonction $B(p,q)$ par rapport à p ou à q.

Puisque (voir [5])

(21) $$B(p,q) \cdot \Gamma(p+q) = \Gamma(p) \cdot \Gamma(q)$$

à la suite de la relation (9) on peut écrire que

(22) $$\begin{cases} \dfrac{\partial B(p,q)}{\partial p} = -qB(p,q) \displaystyle\sum_{k=0}^{\infty} \frac{1}{(k+p)(k+p+q)} \\ \dfrac{\partial B(p,q)}{\partial q} = -pB(p,q) \displaystyle\sum_{k=0}^{\infty} \frac{1}{(k+q)(k+p+q)} \end{cases}$$

et, par conséquent, en introduisant (22) en (20) on trouve la relation (16) et l'affirmation de la proposition est vérifiée.

Remarque. Si dans (15) on prend $m = 0$, $a = \frac{1}{n}$, $n = N^*$, $p = \frac{1}{2}$ et $q = \frac{n}{2}$, alors on a la distribution Student $S(n)$, et de (16) on déduit que

(23) $$H(X) = \ln \sqrt{n}\, B\left(\frac{1}{2}, \frac{n}{2}\right) + (n+1) \sum_{k=0}^{\infty} \frac{1}{(2k+n)(2k+n+1)}$$

d'où, en particulièr, pour $n = 1$ on a la distribution de Cauchy, pour laquelle on déduit que $H(X) = \ln 4\pi$.

Analoguement à la proposition 3 on peut formuler et déduire la

PROPOSITION 4. *Si*

(24) $$f(x) = \frac{(m+1)a^p}{pB(p,q)} x^m \left(1 + ax^{\frac{m+1}{p}}\right)^{-(p+q)}, \quad x \geqslant 0$$

où $a > 0,\ m > -1,\ p > 0,\ q > 0$ *et* $B(p,q)$ *est la fonction bêta d'Euler, alors*

$$H(X) = \ln \frac{pB(p,q)}{m+1} a^{-\frac{m+1}{p}} + \frac{mpq}{m+1} \sum_{k=0}^{\infty} \frac{1}{(k+p)(k+p+q)} +$$

(25) $$+ \frac{(m+1)pq + p^2}{m+1} \sum_{k=0}^{\infty} \frac{1}{(k+q)(k+p+q)}$$

Remarque 1. Si dans (24) on prend $m = 0$ et $p = q = 1$, alors on a la distribution rationnelle et de (25) on déduit que $H(X) = \ln \frac{e^2}{a}$

Remarque 2. Si dans (24) on prend $p = 1$, alors on a la distribution de Burr de paramètre a et de (25) on déduit que

(26) $$H(X) = \frac{mq}{m+1} \sum_{k=0}^{\infty} \frac{1}{(k+1)(k+q+1)} + \frac{(m+1)q+1}{(m+1)q} - \ln q(m+1)a^{m+1}$$

d'ou pour $a = 1$ on trouve l'entropie de la distribution Burr simple.

Remarque 3. Si dans (24) on prend $m = \frac{n_1}{2} - 1$, $p = \frac{n_1}{2}$, $q = \frac{n_2}{2}$ et $a = \frac{n_1}{n_2}$, $n_1 \in N^*$, $n_2 \in N^*$, alors on a la distribution de Snedecor

$S(n_1, n_2)$ et de (25) on déduit que

$$H(X) = \ln \frac{n_2}{n_1} B\left(\frac{n_1}{2}, \frac{n_2}{2}\right) + 2n_2(n_1 - 2) \sum_{k=0}^{\infty} \frac{1}{(2k+n_1)(2k+n_1+n_2)} +$$

$$+ n_1(n_2 + 1) \sum_{k=0}^{\infty} \frac{1}{(2k+n_2)(2k+n_1+n_2)} \tag{27}$$

On peut formuler et déduire également

Proposition 5. Si

$$f(x) = \frac{na^p}{B(p,q)} x^{np-1}(1 - ax^n)^{q-1}, \quad x \in [0, a^{-\frac{1}{n}}] \tag{28}$$

où $a > 0$, $n > 0$, $p > 0$, $q > 0$ et $B(p,q)$ est la fonction bêta d'Euler, alors

$$H(X) = \ln \frac{B(p,q)}{n} a^{-\frac{1}{n}} + \frac{(np-1)q}{n} \sum_{k=0}^{\infty} \frac{1}{(k+p)(k+p+q)} +$$

$$+ \frac{p(q-1)}{n} \sum_{k=0}^{\infty} \frac{1}{(k+q)(k+p+q)} \tag{29}$$

Remarque. Si dans (28) on prend $a = n = 1$, alors on a la distribution bêta et de (29) on déduit que

$$H(X) = \ln B(p,q) + q(p-1) \sum_{k=0}^{\infty} \frac{1}{(k+p)(k+p+q)} +$$

$$+ p(q-1) \sum_{k=0}^{\infty} \frac{1}{(k+q)(k+p+q)} \tag{30}$$

Par calcul direct on peut déduire également

Proposition 6. *Si*

$$f(x) = \frac{ab^p}{pB(p,q)} e^{ax}(1 + be^{\frac{a}{p}x})^{-(p+q)} \quad x \in R \tag{31}$$

où $a > 0$, $b > 0$, $p > 0$, $q > 0$ et $B(p,q)$ est la fonction bêta d'Euler, alors

$$H(X) = \ln \frac{pB(p,q)}{a} + pq \sum_{k=0}^{\infty} \frac{2k+p+q}{(k+p)(k+q)(k+p+q)} \tag{32}$$

Remarque. Si dans (31) on prend $a = n_1$, $b = \frac{n_1}{n_2}$, $p = \frac{n_1}{2}$, $q = \frac{n_2}{2}$, avec $n_1 \in N^*$ et $n_2 \in N^*$, alors on a la distribution de E. A. Fisher et de (32) on déduit que

$$H(X) = \ln \frac{1}{2} B\left(\frac{n_1}{2}, \frac{n_2}{2}\right) + n_1 n_2 \sum_{k=0}^{\infty} \frac{4k + n_1 + n_2}{(2k+n_1)(2k+n_2)(2k+n_1+n_2)} \tag{38}$$

On peut constater sans difficulté que l'entropie $H(X)$ n'est pas invariable par rapport à des transformations de la variable X.

RÉFÉRENCES

1. Guiasu, S., *Applications de la théorie de l'information*, Editura Academiei, Bucarest, 1968.
2. Guiaşu, S., *Information Theory with Applications*, McGraw-Hill, New York, 1977.
3. Iosifescu, M., Mihoc, Gh., Ciucu, G., *Théorie des probabilités et statistique mathématique*, Editura Tehnică, Bucarest, 1966.
4. Purcaru, I., *Information et corrélation*, Ed. Ştiinţifică şi Enciclopedică, Bucarest, 1988.
5. Teodorescu, N., Olariu, V., *Equations de la physique mathématique*, Editura Didactică şi Pedagogică, Bucarest, 1970.

Reçu le 10.V.1990

Académie d'Études Économiques
Département de Mathématiques
Bucarest, Roumanie