Kybernetika

María Dolores Esteban; Domingo Morales

A summary on entropy statistics

Kybernetika, Vol. 31 (1995), No. 4, 337--346

Persistent URL: http://dml.cz/dmlcz/124679

Terms of use:

© Institute of Information Theory and Automation AS CR, 1995

Institute of Mathematics of the Academy of Sciences of the Czech Republic provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.

This paper has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* http://project.dml.cz

KYBERNETIKA — VOLUME 31 (1995), NUMBER 4, PAGES 337–346

A SUMMARY ON ENTROPY STATISTICS[1]

María Dolores Esteban and Domingo Morales

With the purpose to study as a whole the major part of entropy measures cited in the literature, a mathematical expression is proposed in this paper. In favour of this mathematical tool is the fact that most entropy measures can be obtained as a particular or a limit case of the $H_{h,v}^{\varphi_1,\varphi_2}$-entropy functional, and therefore, all those properties which are proved for the functional are also true for its particularizations. Entropy estimates are obtained by replacing probabilities by relative frequencies and their asymptotic distributions are obtained. To finish the asymptotic variance of many entropy statistics are tabulated.

1. INTRODUCTION

Let $(\mathcal{X}, \beta_{\mathcal{X}}, P)_{P\in\Delta_M}$ be an statistical space, where $\mathcal{X} = \{x_1, \dots, x_M\}$, $\Delta_M = \left\{P = (p_1, \dots, p_M)^t / p_i \geq 0 \text{ and } \sum_{i=1}^M p_i = 1\right\}$ and $\beta_{\mathcal{X}}$ is the σ-field of all the subsets of $\mathcal{X}$. For any $P \in \Delta_M$, the $H_{h,v}^{\varphi_1,\varphi_2}$-entropy is defined by the following expression:

$$H_{h,v}^{\varphi_1,\varphi_2}(P) = h\left(\frac{\sum_{i=1}^M v_i\varphi_1(p_i)}{\sum_{i=1}^M v_i\varphi_2(p_i)}\right),$$

where $v_i > 0$, $i = 1, \dots, M$, is the weight associated to the element x_i of $\mathcal{X}$. Furthermore we suppose that $\varphi_1: [0,1) \to \mathbb{R}$, $\varphi_2: [0,1) \to \mathbb{R}$ and $h: \mathbb{R} \to \mathbb{R}$ are any of the 3-uples of functions appearing in Table 1.

In Table 1, v_i and functions $h(x)$, $\varphi_1(x)$ and $\varphi_2(x)$ are given for the following entropy measures: 1. Shannon [14], 2. Rényi [12], 3. Aczel–Daróczy [1], 4. Aczel–Daróczy [1], 5. Aczel–Daróczy [1], 6. Varma [19], 7. Varma [19], 8. Kapur [8], 9. Havrda–Charvát [7], 10. Arimoto [2], 11. Sharma–Mittal [15], 12. Sharma–Mittal [15], 13. Taneja [18], 14. Sharma–Taneja [16], 15. Sharma–Taneja [17], 16. Ferreri [5], 17. Sant'anna–Taneja [13], 18. Sant'anna–Taneja [13], 19. Belis–Guiasu [3] and Gil [6], 20. Picard [10], 21. Picard [10], 22. Picard [10] and 23. Picard [10].

[1]The research in this paper was supported in part by DGICYT Grants N.PB93-0068 and by Complutense University grant N.PR161/93-4812. Their financial support is gratefully acknowledged.

Table 1.

Measure	$h(x)$	$\varphi_1(x)$	$\varphi_2(x)$	v_i
1	x	$-x\log x$	x	v
2	$(1-r)^{-1}\log x$	x^r	x	v
3	x	$-x^r\log x$	x^r	v
4	$(s-r)^{-1}\log x$	x^r	x^s	v
5	$(1/s)\arctan x$	$x^r\sin(s\log x)$	$x^r\cos(s\log x)$	v
6	$(m-r)^{-1}\log x$	x^{r-m+1}	x	v
7	$(m(m-r))^{-1}\log x$	$x^{r/m}$	x	v
8	$(1-t)^{-1}\log x$	x^{t+s-1}	x^s	v
9	$(1-s)^{-1}(x-1)$	x^s	x	v
10	$(t-1)^{-1}(x^t-1)$	$x^{1/t}$	x	v
11	$(1-s)^{-1}(e^x-1)$	$(s-1)x\log x$	x	v
12	$(1-s)^{-1}\left(x^{\frac{s-1}{r-1}}-1\right)$	x^r	x	v
13	x	$-x^r\log x$	x	v
14	$(s-r)^{-1}x$	x^r-x^s	x	v
15	$(\sin s)^{-1}x$	$-x^r\sin(s\log x)$	x	v
16	$\left(1+\frac{1}{\lambda}\right)\log(1+\lambda)-\frac{x}{\lambda}$	$(1+\lambda x)\log(1+\lambda x)$	x	v
17	x	$-x\log\left(\frac{\sin(sx)}{2\sin(s/2)}\right)$	x	v
18	x	$-\frac{\sin(xs)}{2\sin(s/2)}\log\left(\frac{\sin(sx)}{2\sin(s/2)}\right)$	x	v
19	x	$-x\log x$	x	w_i
20	x	$-\log x$	1	v_i
21	$(1-r)^{-1}\log x$	x^{r-1}	1	v_i
22	$(1-s)^{-1}(e^x-1)$	$(s-1)\log x$	1	v_i
23	$(1-s)^{-1}\left(x^{\frac{r-1}{s-1}}-1\right)$	x^{r-1}	1	v_i

Estimation of population $H_{h,v}^{\varphi_1,\varphi_2}$-entropies can be done by estimating the probability vector P with the relative frequency vector $\widehat{P}=(\widehat{p}_1,\dots,\widehat{p}_M)^t$ associated to a simple random sample of size n. In this paper we show that the asymptotic distribution of $\sqrt{n}\ [H_{h,v}^{\varphi_1,\varphi_2}(\widehat{P})-H_{h,v}^{\varphi_1,\varphi_2}(P)]$ is $\mathcal{N}(0,\sigma^2)$ where $\sigma^2=\sum_{i=1}^M t_i^2p_i-\left(\sum_{i=1}^M t_ip_i\right)^2$, and we tabulate the values of t_i appearing in the expression of its asymptotic variance. On the basis of this result, a confidence interval for $H_{h,v}^{\varphi_1,\varphi_2}(P)$ can be given and hypotheses about $H_{h,v}^{\varphi_1,\varphi_2}(P)$ can be tested.

2. ASYMPTOTIC DISTRIBUTION OF $\mathbf{H}_{h,v}^{\varphi_1,\varphi_2}$-STATISTICS

If $f\in C^i(A)$ denotes that the real valued function f has a continuous derivative of ith order in the set A, then we obtain the following result.

Theorem 2.1. Suposse that $h\in C^1(\mathbb{R})$, $\varphi_1\in C^1((0,1))$, $\varphi_2\in C^1((0,1))$ and $p_i>0$, $i=1,\dots,M$. If the relative frequency estimator of $P=(p_1,\dots,p_M)$, $\widehat{P}$, is based on a simple random sample of size n, then

$$n^{\frac{1}{2}}[H_{h,v}^{\varphi_1,\varphi_2}(\widehat{P}) - H_{h,v}^{\varphi_1,\varphi_2}(P)] \xrightarrow[n\uparrow\infty]{L} \mathcal{N}(0,\sigma^2),$$

where

$$\begin{aligned}
\sigma^2 &= T^t\Sigma T = \sum_{i=1}^{M} t_i^2 p_i - \left(\sum_{i=1}^{M} t_i p_i\right)^2 \\
\Sigma &= (p_i(\delta_{ij} - p_j))_{i,j=1,\ldots,M} = \text{diag}(P) - PP^t \\
T &= (t_1,\ldots,t_M)^t \quad \text{and} \\
t_i &= h'\left(\frac{\sum_{i=1}^M v_i\varphi_1(p_i)}{\sum_{i=1}^M v_i\varphi_2(p_i)}\right) \cdot \frac{v_i\varphi_1'(p_i)\sum_{i=1}^{M} v_i\varphi_2(p_i) - v_i\varphi_2'(p_i)\sum_{i=1}^M v_i\varphi_1(p_i)}{\left(\sum_{i=1}^M v_i\varphi_2(p_i)\right)^2}, \\
& i = 1,\ldots,M.
\end{aligned}$$

Proof. By the mean value theorem

$$H_{h,v}^{\varphi_1,\varphi_2}(\widehat{P}) = H_{h,v}^{\varphi_1,\varphi_2}(P) + \sum_{i=1}^{M} \frac{\partial H_{h,v}^{\varphi_1,\varphi_2}(P^*)}{\partial p_i}(\widehat{p}_i - p_i),$$

where $\|P^* - P\|_2 < \|\widehat{P} - P\|_2$.

We conclude that

$$\sqrt{n}\,[H_{h,v}^{\varphi_1,\varphi_2}(\widehat{P}) - H_{h,v}^{\varphi_1,\varphi_2}(P)] \quad \text{and} \quad \sqrt{n}T^t(\widehat{P} - P)$$

have asymptotically the same distribution (cf. Rao [11], p. 385). Finally applying the Central Limit Theorem, the results follows. □

In Table 2, the expressions of the values t_i obtained in Theorem 2.1 are given.

Table 2.

Measure	t_i
Shannon [14]	$-(1 + \log p_i)$
Rényi [12]	$\frac{r}{1-r} p_i^{r-1} \left[\sum_{i=1}^M p_i^r\right]^{-1}$
Aczel–Daróczy [1]	$-\left[p_i^{r-1}(1 + r\log p_i)\left(\sum_{i=1}^M p_i^r\right)^{-1} - rp_i^{r-1}\sum_{i=1}^M p_i^r \log p_i \left(\sum_{i=1}^M p_i^r\right)^{-2}\right]$
Aczel–Daróczy [1]	$(s-r)^{-1}\left(\sum_{i=1}^M p_i^r\right)^{-1}\left(\sum_{i=1}^M p_i^s\right)\left[rp_i^{r-1}\sum_{i=1}^M p_i^s - sp_i^{s-1}\sum_{i=1}^M p_i^r\right]$

Table 2. (cont.)

Measure	t_i
Aczel–Daróczy [1]	$\frac{1}{s}\left[1+\left(\frac{\sum_{i=1}^{M} p_i^r \sin(s\log p_i)}{\sum_{i=1}^{M} p_i^r \cos(s\log p_i)}\right)^2\right]^{-1}\left[\sum_{i=1}^{M} p_i^r \cos(s\log p_i)\right]^{-2}\cdot$
	$\cdot\left[p_i^{r-1}(r\ \sin(s\log p_i)+s\ \cos(s\log p_i))\sum_{i=1}^{M} p_i^r \cos(s\log p_i)-\right.$
	$\left.-p_i^{r-1}(r\ \cos(s\log p_i)-s\ \sin(s\log p_i))\sum_{i=1}^{M} p_i^r \sin(s\log p_i)\right]$
Varma [19]	$\frac{r-m+1}{m-r} p_i^{r-m}\left(\sum_{i=1}^{M} p_i^{r-m+1}\right)^{-1}$
Varma [19]	$\frac{r}{m^2(m-r)} p_i^{(r/m)-1}\left(\sum_{i=1}^{M} p_i^{r/m}\right)^{-1}$
Kapur [8]	$(1-t)^{-1}\left(\sum_{i=1}^{M} p_i^{t+s-1}\right)^{-1}\left(\sum_{i=1}^{M} p_i^{s}\right)^{-1}\cdot$
	$\cdot\left[(t+s-1)p_i^{t+s-2}\sum_{i=1}^{M} p_i^s - s p_i^{s-1}\sum_{i=1}^{M} p_i^{t+s-1}\right]$
Havrda–Charvát [7]	$(1-s)^{-1} s p_i^{s-1}$
Arimoto [2]	$(t-1)\, p_i^{(1/t)-1}\left(\sum_{i=1}^{M} p_i^{1/t}\right)^{t-1}$
Sharma–Mittal [15]	$-(1+\log p_i)\exp\left\{(s-1)\sum_{i=1}^{M} p_i\log p_i\right\}$
Sharma–Mittal [15]	$\frac{r}{1-r} p_i^{r-1}\left(\sum_{i=1}^{M} p_i^r\right)^{\frac{s-r}{r-1}}$
Taneja [18]	$-p_i^{r-1}(1+r\log p_i)$
Sharma–Taneja [16]	$\frac{1}{s-r}\left[r p_i^{r-1} - s p_i^{s-1}\right]$
Sharma–Taneja [17]	$-\frac{1}{\sin s} p_i^{r-1}\left[r\ \sin(s\log p_i)+s\ \cos(s\log p_i)\right]$
Ferreri [5]	$-(1+\log(1+\lambda p_i))$
Sant'anna–Taneja [13]	$-\left[\log\left(\frac{\sin(p_i s)}{2\ \sin(s/2)}\right)+p_i\frac{s\ \cos(p_i s)}{\sin(p_i s)}\right]$

Table 2. (cont.)

Measure	t_i
Sant'anna–Taneja [13]	$-\frac{s\cos(p_i s)}{2\sin(s/2)}\left(\sin(p_i s) + \log\frac{\sin(p_i s)}{2\sin(s/2)}\right)$
Belis–Guiasu; Gil [3], [6]	$-w_i(\log p_i + 1)\left(\sum_{i=1}^{M} w_i p_i\right)^{-1} + w_i \sum_{i=1}^{M} w_i p_i \log p_i \left(\sum_{i=1}^{M} w_i p_i\right)^{-2}$
Picard [10]	$-v_i p_i^{-1}\left[\sum_{i=1}^{M} v_i\right]^{-1}$
Picard [10]	$-v_i p_i^{r-2}\left[\sum_{i=1}^{M} v_i p_i^{r-1}\right]^{-1}$
Picard [10]	$-v_i p_i^{-1}\left[\sum_{i=1}^{M} v_i\right]^{-1} \exp\left\{(s-1)\frac{\sum_{i=1}^{M} v_i p_i}{\sum_{i=1}^{M} v_i}\right\}$
Picard [10]	$-v_i p_i^{r-2}\left[\sum_{i=1}^{M} v_i\right]^{-1}\left(\sum_{i=1}^{M} v_i p_i \left[\sum_{i=1}^{M} v_i\right]^{-1}\right)^{\frac{r-s}{s-1}}$

The following result gives a necessary and sufficient condition for $\sigma^2 = 0$.

Proposition 2.1. Let $S_n = n^{1/2}T^t(\widehat{P} - P)$ be the first order term in the Taylor's expansion of $H_{h,v}^{\varphi_1,\varphi_2}(\widehat{P})$ around P. Then,

$$S_n = 0 \text{ for all } n \text{ with probability one if and only if } \sigma^2 = 0.$$

Proof. If $S_n = 0$ a.s., then $V[S_n] = 0$ for every $n \in \mathbb{N}$ and therefore

$$\sigma^2 = \lim_{n\to\infty} V[S_n] = 0.$$

On the other hand it is easy to check that $V[S_n] = \sigma^2$, and therefore $\sigma^2 = 0$ implies $S_n = 0$ a.s. □

With regard to Theorem 2.1, it is necessary to determine the asymptotic distribution of the $H_{h,v}^{\varphi_1,\varphi_2}$-statistics when the asymptotic variance become zero. If $A = (a_{ij})_{i,j=1,\dots,M}$ with $a_{ij} = \frac{\partial^2 H_{h,v}^{\varphi_1,\varphi_2}(P)}{\partial p_i \partial p_j}$, then we obtain the following result

Theorem 2.2. Assume that $h \in C^2(\mathbb{R})$, $\varphi_1 \in C^2((0,1))$, $\varphi_2 \in C^2((0,1))$ and $p_i > 0$, $i = 1, \dots, n$. If $\sigma^2 = 0$ and the relative frequency estimator of P, $\widehat{P}$, is based on a random sample of size n, then

$$2n[H_{h,v}^{\varphi_1,\varphi_2}(\widehat{P}) - H_{h,v}^{\varphi_1,\varphi_2}(P)] \xrightarrow[n\uparrow\infty]{L} \sum_{i=1}^{M} \beta_i \chi_1^2,$$

where the χ_1^2's are independent and the β_i's are the eigenvalues of $A\Sigma$.

P r o o f. By Proposition 2.1 and the mean value theorem

$$H_{h,v}^{\varphi_1,\varphi_2}(\widehat{P}) = H_{h,v}^{\varphi_1,\varphi_2}(P) + \frac{1}{2}(\widehat{P}-P)^t \left(\frac{\partial^2 H_{h,v}^{\varphi_1,\varphi_2}(P^*)}{\partial p_i \partial p_j} \right)_{i,j=1\dots,M} (\widehat{P}-P),$$

where $\|P^* - P\|_2 < \|\widehat{P} - P\|_2$.

We conclude that

$$2n[H_{h,v}^{\varphi_1,\varphi_2}(\widehat{P}) - H_{h,v}^{\varphi_1,\varphi_2}(P)] \qquad \text{and} \qquad n(\widehat{P}-P)^t A(\widehat{P}-P)$$

have asymptotically the same distribution (cf. Rao [11], p.385).

Finally, applying the Central Limit Theorem and well known facts about quadratic forms of normal variates, the result follows.

A particular but important case of Theorem 2.2 appears when $P = U = (1/M, \dots, \dots, 1/M)$. Under this asumption, a chi-square asymptotic distribution is obtained. □

Theorem 2.3. Assume that $h \in C^2(\mathbb{R})$, $\varphi_1 \in C^2((0,1))$ and $\varphi_2 \in C^2((0,1))$. If $P = U$, $v_i = v\ \forall i$ and $\widehat{P}$ is based on a random sample of size n, then

$$\frac{2n[H_{h,v}^{\varphi_1,\varphi_2}(\widehat{P}) - H_{h,v}^{\varphi_1,\varphi_2}(U)]}{b} \xrightarrow[n\uparrow\infty]{L} \chi^2_{M-1},$$

where

$$b = h'\left(\frac{\varphi_1(1/M)}{\varphi_2(1/M)}\right)[\varphi_2(1/M)\,\varphi_1''(1/M) - \varphi_1(1/M)\,\varphi_2''(1/M)]\left[M^2\varphi_2(1/M)^2\right]^{-1}.$$

P r o o f. Following the steps of the proof of Theorem 2.2, we get that

$$2n[H_{h,v}^{\varphi_1,\varphi_2}(\widehat{P}) - H_{h,v}^{\varphi_1,\varphi_2}(U)] \qquad \text{and} \qquad n(\widehat{P}-U)^t A(\widehat{P}-U)$$

have asymptotically the same distribution, and therefore the results follows. □

In Table 3, the expressions of the values b obtained in Theorem 2.3 are given.

3. STATISTICAL APPLICATIONS

The previous result giving the asymptotic distribution of $H_{h,v}^{\varphi_1,\varphi_2}$-entropy statistics, in a simple random sampling, can be used in various settings to construct confidence intervals and to test statistical hypotheses based on one or more samples.

Table 3.

Measure	b
Shannon [14]	-1
Rényi [12]	$-r$
Aczel–Daróczy [1]	$2r-1$
Aczel–Daróczy [1]	$[r(r-1)-s(s-1)](s-r)^{-1}$
Aczel-Daróczy [1]	$2r-1$
Varma [19]	$m-r-1$
Varma [19]	$-rm^{-3}$
Kapur [8]	$(1-t)^{-1}[(t+s-1)(t+s-2)-s(s-1)]$
Havrda–Charvát [7]	$-sM^{1-s}$
Arimoto [2]	$-t^{-1}M^{t-1}$
Sharma–Mittal [15]	$-M^{1-s}$
Sharma–Mittal [15]	$-rM^{1-s}$
Taneja [18]	$-M^{1-r}[2r-1-r(r-1)\log M]$
Sharma–Taneja [16]	$M(s-r)^{-1}[r(r-1)M^{-r}-s(s-1)M^{-s}]$
Sharma–Taneja [17]	$-M^{1-r}(\sin s)^{-1}[(2rs-s)\cos(s\log M)-(r(r-1)-s^2)\sin(s\log M)]$
Ferreri [5]	$-\lambda(M+\lambda)^{-1}$
Sant'anna–Taneja [13]	$-sM^{-1}[2\cot(sM^{-1})-sM^{-1}(\csc(sM^{-1}))^2]$
Sant'anna–Taneja [13]	$\frac{1}{M}\left[\frac{s^2\sin(s/M)}{2\sin(s/2)}\left(1+\log\left(\frac{\sin(s/M)}{2\sin(s/2)}\right)\right)-\frac{s^2(\cos(s/M))^2}{2\sin(s/2)\sin(s/M)}\right]$

a) Test for a predicted value of the population entropy

To test $H_0 : H_{h,v}^{\varphi_1,\varphi_2}(P) = D_0$ against $H_1 : H_{h,v}^{\varphi_1,\varphi_2}(P) \neq D_0$, we reject the null hypothesis if

$$|T_a| = \left|\frac{n^{1/2}\left(H_{h,v}^{\varphi_1,\varphi_2}(\widehat{P}) - D_0\right)}{\widehat{\sigma}}\right| > z_{\alpha/2},$$

where $\widehat{\sigma}$ is obtained from σ^2 in Theorem 2.1 when p_i is replaced by $\widehat{p}_i$ and z_α is the $(1-\alpha)$-quantile of the standard normal distribution. In this context an approximate $1-\alpha$ level confidence interval for $H_{h,v}^{\varphi_1,\varphi_2}(P)$ is given by

$$\left(H_{h,v}^{\varphi_1,\varphi_2}(\widehat{P}) - \frac{\widehat{\sigma}z_{\alpha/2}}{n^{1/2}}, H_{h,v}^{\varphi_1,\varphi_2}(\widehat{P}) + \frac{\widehat{\sigma}z_{\alpha/2}}{n^{1/2}}\right).$$

Furthermore the minimum sample size giving a maximun error ε at a confidence level $1-\alpha$, is

$$n = \left[\frac{\widehat{\sigma}^2 z_{\alpha/2}^2}{\varepsilon^2}\right] + 1.$$

b) Test for a common predicted value of r population entropies

To test $H_0 : H_{h,v}^{\varphi_1,\varphi_2}(P_1) = \ldots = H_{h,v}^{\varphi_1,\varphi_2}(P_r) = D_0$, we reject the null hypotheses if

$$T_b = \sum_{j=1}^{r} \frac{n_j \left(H_{h,v}^{\varphi_1,\varphi_2}(\widehat{P}_j) - D_0\right)^2}{\widehat{\sigma}_j^2} > \chi_{r,\alpha}^2,$$

where n_j is the size of the independent sample in the jth population, $\widehat{\sigma}_j$'s are obtained from σ when p_i is replaced in theorem 2.1 by $\widehat{p}_i^{(j)}$, $i = 1,\ldots,M$, $j = 1,\ldots,r$, and $\chi_{r,\alpha}^2$ is the $(1-\alpha)$-quantile of the chi–square distribution with r degrees of freedom.

In this context an approximate $1-\alpha$ confidence interval for the difference of entropies corresponding to independent populations is given by

$$H_{h,v}^{\varphi_1,\varphi_2}(\widehat{P}_1) - H_{h,v}^{\varphi_1,\varphi_2}(\widehat{P}_2) \pm z_{\alpha/2}\sqrt{\frac{\widehat{\sigma}_1^2}{n_1} + \frac{\widehat{\sigma}_2^2}{n_2}}.$$

Furthermore, for $n = n_1 = n_2$, the minimum sample size giving a maximum error ε at a confidence level $1-\alpha$, is

$$n = \left[\frac{(\widehat{\sigma}_1^2 + \widehat{\sigma}_2^2) z_{\alpha/2}^2}{\varepsilon^2}\right] + 1.$$

c) Test for the equality of r population entropies

To test $H_0 : H_{h,v}^{\varphi_1,\varphi_2}(P_1) = \ldots = H_{h,v}^{\varphi_1,\varphi_2}(P_r)$, we reject the null hypotheses if

$$T_c = \sum_{j=1}^{r} \frac{n_j \left(H_{h,v}^{\varphi_1,\varphi_2}(\widehat{P}_j) - \overline{H}\right)^2}{\widehat{\sigma}_j^2} > \chi_{r-1,\alpha}^2,$$

where

$$\overline{H} = \frac{\sum_{j=1}^{r} \frac{n_j H_{h,v}^{\varphi_1,\varphi_2}(\widehat{P}_j)}{\widehat{\sigma}_j^2}}{\sum_{j=1}^{r} \frac{n_j}{\widehat{\sigma}_j^2}},$$

and n_j and $\widehat{\sigma}_j$ are defined above.

d) Test for discrete uniformity

To test $H_0: P = U$, we reject the null hypothesis if

$$T_d = \frac{2n\left[H_{h,v}^{\varphi_1,\varphi_2}(\widehat{P}) - H_{h,v}^{\varphi_1,\varphi_2}(P)\right]}{b} > \chi^2_{M-1,\alpha}.$$

Entropic test of uniformity including that considered in Example 3.1 have been studied in Feistauerová and Vajda [4]. This test is specially interesting because it can be used to test for goodness-of-fit to a completely specified distribution. In this sense, we are using the idea that Mann and Wald [9] suggested, i. e. to take intervals with equal probabilities. To finish, we give an example to illustrate this procedure.

Example 3.1. The following sample was simulated from a Normal distribution with mean 2 and standard deviation 1.1:

1.4917575	2.4957872	3.4450585	1.1089838	3.9169254
1.1129468	3.4227659	2.2783900	1.7450712	3.1313416
0.8892388	2.0998305	2.6937549	2.4090498	1.0307736
1.4372195	0.4340365	1.6559174	2.1303488	1.9678865
1.1263431	3.1135033	2.7112041	0.8490701	1.9045762
1.5705222	3.3131407	2.5999673	2.2633095	2.2082426
1.6716791	3.1731862	1.2352134	2.0345601	4.0074727
2.6539161	2.2695147	1.7740887	4.0774582	0.7336872
0.0610451	1.9614988	1.9162852	2.6076725	2.0605398
1.4447520	−.3579844	0.2110429	2.5557666	1.1575424

To test for H_0: Data from $\mathcal{N}(2, 1.1)$, we take six intervals:

$$\begin{array}{lcl} I_1 & = & (-\infty, 2 - 0.97 \times 1.1) = (-\infty, 0.933) \\ I_2 & = & (2 - 0.97 \times 1.1, 2 - 0.43 \times 1.1) = (0.933, 1.527) \\ I_3 & = & (2 - 0.43 \times 1.1, 2) = (1.527, 2) \\ I_4 & = & (2, 2 + 0.43 \times 1.1) = (2, 2.473) \\ I_5 & = & (2 + 0.43 \times 1.1, 2 + 0.97 \times 1.1) = (2.473, 3.067) \\ I_6 & = & (2 + 0.97 \times 1.1, \infty) = (3.067, \infty) \end{array}$$

with the property

$$P(\mathcal{N}(2, 1.1) \in I_i) = \frac{1}{6}, \quad i = 1, \ldots, 6.$$

We use the Shannon entropy statistic, so we reject the null hypothesis if

$$T = 2\,n\left[\log M - H(\widehat{P})\right] > \chi^2_{M-1,0.05}.$$

Now, $\widehat{p}_1 = 0.14$, $\widehat{p}_2 = 0.18$, $\widehat{p}_3 = 0.18$, $\widehat{p}_4 = 0.18$, $\widehat{p}_5 = 0.14$, $\widehat{p}_6 = 0.18$, $n = 50$, $H(\widehat{P}) = -\sum_{i=1}^{6} \widehat{p}_i \log \widehat{p}_i = 1.785$, $T = 0.676$ and $\chi^2_{5,0.05} = 11.070$. Furthermore, the classical chi-square statistic is $S = nM \sum_{i=1}^{M} \left(\widehat{p} - \frac{1}{M}\right)^2 = 0.64$. Thus both procedures behaves similarly and the conclusion is that we cannot reject the null hypothesis.

(Received June 17, 1994.)

REFERENCES

[1] J. Aczél and Z. Daróczy: Characterisierung der entropien positiver ordnung und der Shannonschen entropie. Acta Math. Acad. Sci. Hunger. *14* (1963), 95–121.

[2] S. Arimoto: Information-theoretical considerations on estimation problems. Inform. and Control *19* (1971), 181–194.

[3] M. Belis and S. Guiasu: A quantitative–qualitative measure of information in cybernetics systems. IEEE Trans. Inform. Theory *IT-4* (1968), 593–594.

[4] J. Feistauerová and I. Vajda: Testing system entropy and prediction error probability. IEEE Trans. Systems Man Cybernet. *23* (1993), 1352–1358.

[5] C. Ferreri: Hypoentropy and related heterogeneity divergence measures. Statistica *40* (1980), 55–118.

[6] P. Gil: Medidas de incertidumbre e información en problemas de decisión estadística. Rev. Real Acad. Cienc. Exact. Fís. Natur. Madrid *LXIX* (1975), 549–610.

[7] J. Havrda and F. Charvát: Concept of structural α-entropy. Kybernetika *3* (1967), 30–35.

[8] J. N. Kapur: Generalized entropy of order α and type β. The Math. Seminar *4* (1967), 78–82.

[9] H. B. Mann and A. Wald: On the choice of the number of class intervals in the application of the chi-squared test. Ann. Math. Statist. *13* (1942), 306–317.

[10] C. F. Picard: The use of Information theory in the study of the diversity of biological populations. In: Proc. Fifth Berk. Symp. *IV*, 1979, pp. 163–177.

[11] C. R. Rao: Linear Statistical Inference and its Applications. Second edition. J. Wiley, New York 1973.

[12] A. Rényi: On the measures of entropy and information. In: Proc. 4th Berkeley Symp. Math. Statist. and Prob., 1, 1961, pp. 547–561.

[13] A. P. Sant'anna and I. J. Taneja: Trigonometric entropies, Jensen difference divergences and error bounds. Inform. Sci. *35* (1985), 145–156.

[14] C. E. Shannon: A mathematical theory of communication. Bell. System Tech. J. *27* (1948), 379–423.

[15] B. D. Sharma and D. P. Mittal: New non-additive measures of relative information. J. Combin. Inform. System Sci. *2* (1975), 122–133.

[16] B. D. Sharma and I. J. Taneja: Entropy of type (α, β) and other generalized measures in information theory. Metrika *22* (1975), 205–215.

[17] B. D. Sharma and I. J. Taneja: Three generalized additive measures of entropy. Elektron. Informationsverarb. Kybernet. *13* (1977), 419–433.

[18] I. J. Taneja: A Study of Generalized Measures in Information Theory. Ph.D. Thesis, University of Delhi 1975.

[19] R. S. Varma: Generalizations of Rényi's entropy of order α. J. Math. Sci. *1* (1966), 34–48.

Professor María Dolores Esteban and Professor Domingo Morales, Departamento de Estadística e Investigacion Operativa, Facultad de Matemáticas, Universidad Complutense de Madrid, 28040 – Madrid. Spain.