Aplikace matematiky

Harish C. Taneja

On the mean and the variance of estimates of Kullback information and relative "useful" information measures

Aplikace matematiky, Vol. 30 (1985), No. 3, 166–175

Persistent URL: `http://dml.cz/dmlcz/104139`

ON THE MEAN AND THE VARIANCE OF ESTIMATES OF KULLBACK INFORMATION AND RELATIVE "USEFUL" INFORMATION MEASURES

HARISH C. TANEJA

(Received March 17, 1983)

1. INTRODUCTION

Let $P = (p_1, p_2, \ldots, p_n)$, $0 < p_i \leqq 1$, $\sum_{i=1}^{n} p_i = 1$ be a finite discrete probability distribution of a set of n events $E = (E_1, E_2, \ldots, E_n)$ on the basis of an experiment whose predicted probability distribution is $Q = (q_1, q_2, \ldots, q_n)$, $0 < q_i \leqq 1$, $\sum_{i=1}^{n} q_i = 1$. Then Kullback's measure of relative information [3] is defined as

$$I(P/Q) = \sum_{i=1}^{n} p_i \log (p_i/q_i), \quad p_i, q_i > 0 . \tag{1.1}$$

The measure (1.1) depends only on the probabilities of the events and thus does not take into account the effectiveness of the events under consideration. Belis and Guiasu [2] introduced a 'utility distribution' $U = (u, u_2, \ldots, u_n)$ where each $u_i > 0$ accounts for the utility of the ith outcome E_i.

Thus we have two utility information schemes:

$$S = \begin{bmatrix} E_1 & E_2 & \ldots & E_n \\ p_1 & p_2 & \ldots & p_n \\ u_1 & u_2 & \ldots & u_n \end{bmatrix} \tag{1.2}$$

of a set of n events after an experiment, and

$$S^* = \begin{bmatrix} E_1 & E_2 & \ldots & E_n \\ q_1 & q_2 & \ldots & q_n \\ u_1 & u_2 & \ldots & u_n \end{bmatrix} \tag{1.3}$$

of the same set of n events before the experiment.

In both schemes (1.2) and (1.3) the utility distribution is the same because it is assumed that the utility u_i of an outcome E_i is independent of its probability of occurrence p_i, or predicted probability q_i; u_i is only a 'utility' or 'importance' of the outcome E_i for an observer relative to some specified goal, cf. Longo [4].

The measure of relative 'useful' information that the scheme (1.2) provides about the scheme (1.3), has been characterized by Taneja and Tuteja [5]. It is given by

$$(1.4) \qquad I(P/Q; U) = \sum_{i=1}^{n} u_i p_i \log (p_i/q_i), \quad u_i > 0, \quad 0 < p_i, \quad q_i \leqq 1 .$$

The quantity (1.4) measure the average 'useful' information gain in predicting a set of n events $E = (E_1, E_2, \ldots, E_n)$. In what follows, we shall denote the measures (1.1) and (1.4) by I and I_u, respectively.

The maximum lokelihood estimators (MLE) of the quantities given in (1.1) and (1.4), see Anderson [1], are given by

$$(1.5) \qquad \hat{I} = \sum_{i=1}^{n} \hat{p}_i \log (\hat{p}_i/q_i)$$

and

$$(1.6) \qquad \hat{I}_u = \sum_{i=1}^{n} u_i \hat{p}_i \log (\hat{p}_i/q_i),$$

respectively, where $\hat{p}_i$ is the MLE of p_i, so that if N_i is the frequency of occurrence of an event E_i in a random sample of size N, then

$$\hat{p}_i = N_i/N , \quad i = 1, 2, \ldots, n .$$

It may be noted that here q_i, $i = 1, 2, \ldots, n$, are the probabilities which the experimenter assigns to the various possible outcomes of the experiment. These are just predicted probabilities and have not been obtained on the basis of any experiment.

The measures of relative information find wide applications in statistics and economics, cf. [3] and [6]. Thus there is a need to study estimates of these measures. In Section 2, we obtain the mean and the variance of the MLE of (1.1) and show that it is biased and consistent. It is found that this estimate, in fact, overestimates the ture value of I. In Section 3, we obtain the mean and the variance of the MLE of (1.4) and show that it is also biased and consistent, and further that this estimate also overestimates the true value of I_u.

2. MEAN AND VARIANCE OF THE MLE OF $I(P/Q)$

The MLE of the relative information measure $I(P/Q)$ is given by

$$(2.1) \qquad \hat{I} = \sum_{i=1}^{n} \hat{p} \log (\hat{p} / q),$$

where $\hat{p}_i = N_i/N$ is the MLE of p_i and N_i, the frequency of occurrence of an event E_i in a random sample of size N, follows multinomial distribution and thus the moment generating function of the distribution of N_i's can be written as

$$(2.2) \qquad M(t_1, t_2, \ldots, t_n) = (p_1 e^{t_1} + p_2 e^{t_2} + \ldots + p_n e^{t_n})^N .$$

In our study we shall need the central moments of $\hat{p}_i$, which can easily be obtained from (2.2) by using

$$E(\hat{p}_i^r) = N^{-r}\left[\frac{\partial^r M(t_1, t_2, \ldots, t_n)}{\partial t_i^r}\right]_{t_1=t_2=\ldots=t_n=0}$$

and

$$E(\hat{p}_i^r \hat{p}_j^k) = N^{-rk}\left[\frac{\partial^r}{\partial t_i^k}\left\{\frac{\partial^k M(t_1, t_2, \ldots, t_n)}{\partial t_j^k}\right\}\right]_{t_1=t_2=\ldots=t_n=0}$$

for $i \neq j$.

The central moments of various orders of $\hat{p}_i$ that we shall need are as follows:

(2.3)
$$\begin{aligned}
E(\hat{p}_i) &= p_i\,,\\
E(\hat{p}_i - p_i)^2 &= \frac{p_i(1-p_i)}{N}\,,\\
E(\hat{p}_i - p_i)(\hat{p}_j - p_j) &= \frac{-p_i p_j}{N}\,, \quad i \neq j\\
E(\hat{p}_i - p_i)^3 &= \frac{2p_i^2 - 3p_i^3 + p_i}{N^2}\,,\\
E(\hat{p}_i - p_i)^2(\hat{p}_j - p_j) &= \frac{-p_i p_j(1-2p_i)}{N^2}\,, \quad i \neq j\\
E(\hat{p}_i - p_i)^4 &= O(N^{-2})\,,\\
E(\hat{p}_i - p_i)^3(\hat{p}_j - p_j) &= O(N^{-2})\,, \quad i \neq j\\
E(\hat{p}_i - p_i)^2(\hat{p}_j - p_j)^2 &= O(N^{-2})\,, \quad i \neq j
\end{aligned}$$

where $O(N^{-l})$ denotes the terms of magnitude N^{-l}, $l > 0$.

Theorem 1. *Given a set of n independent events $E_1, E_2, \ldots, E_n$ with probabilities of occurrence $p_1, p_2, \ldots, p_n$ and predicted probabilities $q_1, q_2, \ldots, q_n$, the mean and the variance of $\hat{I}$, the MLE of the relative information measure* (2.1), *are given respectively by*

(2.4)
$$E(\hat{I}) = I + \frac{n-1}{2N} + O(N^{-2})$$

and

(2.5)
$$V(\hat{I}) = \frac{1}{N}\left[\sum_{i=1}^{n} p_i \log^2(p_i/q_i) - I^2\right] + O(N^{-2})\,.$$

Proof. Let

(2.6)
$$I(p_i/q_i) = p_i \log(p_i/q_i)\,, \quad p_i, q_i > 0$$

and

$$(2.7) \qquad \hat{I}(p_i/q_i) = \hat{p}_i \log (\hat{p}_i/q_i)\,,$$

for $i = 1, 2, \ldots, n$.

Then the derivatives of $I(p_i/q_i)$ at p_i, $i = 1, 2, \ldots, n$ are given by

$$(2.8) \qquad I^{(1)}(p_i/q_i) = \frac{\partial I(p_i/q_i)}{\partial p_i} = 1 + \log (p_i/q_i)$$

and

$$I^{(r)}(p_i/q_i) = \frac{\partial^r I(p_i/q_i)}{\partial p_i^r} = (-1)^{r-2}\,(r-2)!\; p^{-r+1}\,, \quad r \geqq 2\,.$$

Further, it is easy to show that the mixed derivatives of all orders of $I(p\,/q\,)$vanish, and also that the derivatives in (2.8) are continuous at $p\, > 0$ for $i = 1, 2, \ldots, n$. Thus we can expand (2.7) in a convergent Taylor series about the point p_i with Lagrange's form of the remainder (considering derivatives up to the fourth order only) as follows:

$$(2.9) \quad \hat{I}(p_i/q_i) = I(p_i/q_i) + (\hat{p}_i - p_i)\, I^{(1)}(p_i/q_i) + \frac{1}{2!}(\hat{p}_i - p_i)^2\, I^{(2)}(p_i/q_i) +$$

$$+ \frac{1}{3!}(\hat{p}_i - p_i)\, I^{(3)}(p_i/q_i) + \frac{1}{4!}(\hat{p}_i - p_i)^4\, I^{(4)}[\{p_i + \theta(\hat{p}_i - p_i)\}/q_i]\,, \quad 0 < \theta < 1\,.$$

By virtue of (2.8) this gives

$$(2.10) \quad \hat{I}(p_i/q_i) = I(p_i/q_i) + \{1 + \log (p_i/q_i)\}\, (\hat{p}_i - p_i) + \frac{(\hat{p}_i - p_i)^2}{2p_i} -$$

$$- \frac{(\hat{p}_i - p_i)^3}{6p_i^2} + \frac{(\hat{p}_i - p_i)^4}{12\{p_i + \theta(\hat{p}_i - p_i)\}^3}\,, \quad 0 < \theta < 1$$

Summing this for all i's, we get

$$(2.11) \; \hat{I}(P/Q) = I(P/Q) + \sum_{i=1}^{n}\{1 + \log (p_i/q_i)\}\, (\hat{p}_i - p_i) + \frac{1}{2}\sum_{i=1}^{n}\frac{(\hat{p}_i - p_i)^2}{p_i} -$$

$$- \frac{1}{6}\sum_{i=1}^{n}\frac{(\hat{p}_i - p_i)^3}{p_i^2} + \frac{1}{12}\sum_{i=1}^{n}\frac{(\hat{p}_i - p_i)^4}{\{p_i + \theta(\hat{p}_i - p_i)\}^3}\,, \quad 0 < \theta < 1\,.$$

Using (2.3), we get

$$(2.12) \qquad E(\hat{I}) = I + \frac{1}{2N}\sum_{i=1}^{n}(1 - p_i) - \frac{1}{6N^2}\sum_{i=1}^{n}\left(2 - 3p_i + \frac{1}{p_i}\right) +$$

$$+ \frac{1}{12}\sum_{i=1}^{n} E\left[\frac{(\hat{p}_i - p_i)^4}{\{p_i + \theta(\hat{p}_i - p_i)\}^3}\right], \quad 0 < \theta < 1\,.$$

The quantity $E\left[\frac{(\hat{p}_i - p_i)^4}{\{p_i + \theta(\hat{p}_i - p_i)\}^3}\right]$ is of an order less than N^{-2}, since

$$\frac{(\hat{p}_i - p_i)^4}{\{p_i(1-\theta) + \theta\hat{p}_i\}^3} \leq \frac{(\hat{p}_i - p_i)^4}{p_i^3(1-\theta)^3}.$$

Thus (2.12) can be rewritten as

$$E(\hat{I}) = I + \frac{1}{2N}\sum_{i=1}^{n}(1 - p_i) - \frac{1}{6N^2}\sum_{i=1}^{n}\left(2 - 3p_i + \frac{1}{p_i}\right) + O(N^{-3})$$

or

$$E(\hat{I}) = I + \frac{n-1}{2N} + O(N^{-2}),$$

which is (2.4).

Next we find the variance of the estimate $\hat{I}(P/Q)$. By definition,

$$\begin{aligned} V(\hat{I}) &= E[\hat{I} - E(\hat{I})]^2 \\ &= E\left[\hat{I} - I - \frac{n-1}{2N} + O(N^{-2})\right]^2. \end{aligned}$$

Using (2.11) and restricting ourselves to derivatives up to the third order only, we get

$$\begin{aligned} V(\hat{I}) = E\Bigg[&\sum_{i=1}^{n}\{1 + \log(p_i/q_i)\}(\hat{p}_i - p_i) + \frac{1}{2}\sum_{i=1}^{n}\frac{(\hat{p}_i - p_i)^2}{p_i} - \\ &- \frac{1}{6}\sum_{i=1}^{n}\frac{(\hat{p}_i - p_i)^3}{\{p_i + \theta(\hat{p}_i - p_i)\}^2} - \frac{n-1}{2N} + O(N^{-2})\Bigg]^2, \end{aligned}$$

or

(2.13) $$V(\hat{I}) = E\sum_{i=1}^{n}(\hat{p}_i - p_i)\{1 + \log(p_i/q_i)\}]^2 + O(N^{-2}).$$

The first term of (2.13) is equal to

$$\begin{aligned} &E[\sum_{i=1}^{n}(\hat{p}_i - p_i)^2\{1 + \log(p_i/q_i)\}^2 + \\ &\quad + \sum_{i\neq j}(\hat{p}_i - p_i)(\hat{p}_j - p_j)\{1 + \log(p_i/q_i)\}\{1 + \log(p_j/q_j)\}] = \\ &= \frac{1}{N}[\sum_{i=1}^{n}p_i(1 - p_i)\{1 + \log(p_i/q_i)\}^2 - \sum_{i\neq j}p_ip_j\{1 + \log(p_i/q_i)\}\{1 + \log(p_j/q_j)\}] = \\ &= \frac{1}{N}[\sum_{i=1}^{n}p_i + \sum_{i=1}^{n}p_i\log^2(p_i/q_i) + 2\sum_{i=1}^{n}p_i\log(p_i/q_i) - \sum_{i=1}^{n}p_i^2 - \sum_{i=1}^{n}p_i^2\log^2(p_i/q_i) - \end{aligned}$$

$$- 2\sum_{i=1}^{n} p_i^2 \log (p_i/q_i) - \sum_{i \neq j} p_i p_j \log (p_i/q_i) \log (p_j/q_j) -$$

$$- \sum_{i \neq j} p_i p_j \{1 + \log (p_i/q_i) + \log (p_j/q_j)\}] =$$

$$= \frac{1}{N} [\sum_{i=1}^{n} p_i \log^2 (p_i/q_i) - I^2] + \frac{1}{N} [\sum_{i=1}^{n} [p_i + 2p_i \log (p_i/q_i) - p_i^2 - 2p_i^2 \log (p_i/q_i) -$$

$$- p_i \sum_{j=1}^{n} p_j \{1 + \log (p_i/q_i) + \log (p_j/q_j)\} - p_i^2 \{1 + 2 \log (p_i/q_i)\}]] =$$

$$= \frac{1}{N} [\sum_{i=1}^{n} p_i \log^2 (p_i/q_i) - I^2] + \frac{1}{N} [\sum_{i=1}^{n} p_i + 2\sum_{i=1}^{n} p_i \log (p_i/q_i) - (\sum_{i=1}^{n} p_i)(\sum_{j=1}^{n} p_j) -$$

$$- (\sum_{j=1}^{n} p_j) \{ \sum_{i=1}^{n} p_i \log (p_i/q_i)\} - (\sum_{i=1}^{n} p_i) \{ \sum_{j=1}^{n} p_j \log (p_j/q_j)\}] =$$

$$= \frac{1}{N} [\sum_{i=1}^{n} p_i \log^2 (p_i/q_i) - I^2] .$$

This proves (2.5).

In proving the above theorem, we have used derivatives of $I(p_i/q_i)$ taking the logarithms to the base 'e' and therefore the units of estimates in (2.4) are natural units. However, if we consider 'binary units', then the biased information content in this estimate of I is given by

$$E(\hat{I}) - I = \frac{n-1}{2N} \log_2 e + O(N^{-2}) .$$

Further, since $n > 1$, the estimate $\hat{I}$ in fact overstimates the true value of I.

We note that when $N \to \infty$, then $E(\hat{I}) \to I$ and $V(\hat{I}) \to 0$. Therefore, we conclude that $\hat{I}$ is a consistent estimator of I.

3. MEAN AND VARIANCE OF THE MLE OF $I(P/Q; U)$

The MLE of the relative 'useful' information measure $I(P/Q; U)$ is given by

$$\hat{I}_u = \sum_{i=1}^{n} u_i \hat{p}_i \log (\hat{p}_i/q_i) \tag{3.1}$$

where $\hat{p}_i = N_i/N$ is the MLE of p_i and N_i, the frequency of the occurrence of the event E_i in a random sample of size N, follows a multinomial distribution.

Now find the expression for the mean and the variance of $\hat{I}_u$. We prove the following theorem.

Theorem 2. *Given a set of n independent events $E_1, E_2, \ldots, E_n$ with probabilities of occurrence $p_1, p_2, \ldots, p_n$ and predicted probabilities $q_1, q_2, \ldots, q_n$, the mean and* *the variance of* $\hat{I}_u$, *the MLE of relative 'useful' information, are given respectively by*

(3.2)
$$E(\hat{I}_u) = I_u + \frac{U - \bar{u}}{2N} + O(N^{-2})$$

and

(3.3)
$$V(\hat{I}_u) = \frac{1}{N}\left[\sum_{i=1}^{n} u_i^2 p_i\{\log^2(p_i/q_i) + \log e(p_i/q_i)^2\} - \frac{1}{N}(I_u + \bar{u})^2 + O(N^{-2}),\right.$$

where

$$U = \sum_{i=1}^{n} u_i \quad \textit{and} \quad \bar{u} = \sum_{i=1}^{n} u_i p_i .$$

Proof. Let

(3.4)
$$I(p_i/q_i, u_i) = u_i p_i \log(p_i/q_i), \quad u_i > 0, \quad p_i, q_i > 0$$

and

(3.5)
$$\hat{I}(p_i/q_i, u_i) = u_i \hat{p}_i \log(\hat{p}_i/q_i)$$

for $i = 1, 2, \ldots, n$.

Then the derivatives of $I(p_i/q_i, u_i)$ at p_i, $i = 1, 2, \ldots, n$ are given by

(3.6)
$$I^{(1)}(p_i/q_i, u_i) = \frac{\partial I(p_i/q_i, u_i)}{\partial p_i} = u_i\{1 + \log(p_i/q_i)\}$$

and

$$I^{(r)}(p_i/q_i; u_i) = \frac{\partial^r I(p_i/q_i, u_i)}{\partial p_i^r} = u_i(-1)^{r-2}(r-2)!\, p_i^{-r+1}, \quad r \geqq 2 .$$

Further, we note that the mixed derivatives of all orders of $I(p_i/q_i, u_i)$ vanish and also the derivatives in (3.6) are continuous at $u_i > 0$, $p_i, q_i > 0$, for $i = 1, 2, \ldots, n$. Thus we can expand (3.5) in a convergent Taylor series about the point p_i with Lagrange's form of the remainder as follows:

3.7)
$$\hat{I}(p_i/q_i, u_i) = I(p_i/q_i, u_i) + (\hat{p}_i - p_i) I^{(1)}(p_i/q_i, u_i) + \frac{1}{2!}(\hat{p}_i - p_i)^2 I^{(2)}(p_i/q_i, u_i) + \frac{1}{3!}(\hat{p}_i - p_i)^3 I^{(3)}(p_i/q_i, u_i) + \frac{1}{4!}(\hat{p}_i - p_i)^4 I^{(4)}[\{p_i + \theta(\hat{p}_i - p_i)\}/q_i, u_i], \quad 0 < \theta < 1 .$$

By virtue of (3.6) this gives

(3.8)
$$\hat{I}(p_i/q_i, u_i) = I(p_i/q_i, u_i) + (\hat{p}_i - p_i) u_i\{1 + \log(p_i/q_i)\} + \frac{1}{2}\frac{u_i(\hat{p}_i - p_i)^2}{p_i} - \frac{u_i(\hat{p}_i - p_i)^3}{6p_i^2} + \frac{u_i(\hat{p}_i - p_i)^4}{12\{p_i + \theta(\hat{p}_i - p_i)\}^3}, \quad 0 < \theta < 1 .$$

Summing this for all i's, we get

$$(3.9) \qquad \hat{I}(P/Q;U) = I(P/Q;U) + \sum_{i=1}^{n} u_i(\hat{p}_i - p_i)\{1 + \log(p_i/q_i)\} +$$

$$+ \frac{1}{2}\sum_{i=1}^{n}\frac{u_i(\hat{p}_i - p_i)^2}{p_i} - \frac{1}{6}\sum_{i=1}^{n}\frac{u_i(\hat{p}_i - p_i)^3}{p_i^2} + \frac{1}{12}\sum_{i=1}^{n}\frac{u_i(\hat{p}_i - p_i)^4}{\{p_i + \theta(\hat{p}_i - p_i)\}^3}, \quad 0 < \theta < 1 .$$

Now using (2.3), we get

$$(3.10) \qquad E(\hat{I}_u) = I_u + \frac{1}{2N}\sum_{i=1}^{n} u_i(1 - p_i) - \frac{1}{6N^2}\sum_{i=1}^{n} u_i\left(2 - 3p_i + \frac{1}{p_i}\right) +$$

$$+ \frac{1}{12}\sum_{i=1}^{n} u_i E\left[\frac{(\hat{p}_i - p_i)^4}{\{p_i + \theta(\hat{p}_i - p_i)\}^3}\right], \quad 0 < \theta < 1 .$$

As in Theorem 1, (3.10) can be rewritten as

$$E(\hat{I}_u) = I_u + \frac{1}{2N}\sum_{i=1}^{n} u_i(1 - p_i) - \frac{1}{6N^2}\sum_{i=1}^{n} u_i\left(2 - 3p_i + \frac{1}{p_i}\right) + O(N^{-3})$$

or

$$E(\hat{I}_u) = I_u + \frac{U - \bar{u}}{2N} + O(N^{-2}),$$

which is (3.2).

Next we find the variance of the estimate $\hat{I}(P/Q;U)$. By definition,

$$V(\hat{I}_u) = E[\hat{I}_u - E(\hat{I}_u)]^2$$
$$= E\left[\hat{I}_u - I_u - \frac{U - \bar{u}}{2N} + O(N^{-2})\right]^2 .$$

Using (3.9) and restricting ourselves to derivatives up to the third order only, we get

$$V(\hat{I}_u) = E\left[\sum_{i=1}^{n} u_i(\hat{p}_i - p_i)\{1 + \log(p_i/q_i)\} + \frac{1}{2}\sum_{i=1}^{n}\frac{u_i(\hat{p}_i - p_i)^2}{p_i} - \right.$$
$$\left. - \frac{1}{6}\sum_{i=1}^{n}\frac{u_i(\hat{p}_i - p_i)^3}{\{p_i + \theta(\hat{p}_i - p_i)\}^2} - \frac{U - \bar{u}}{2N} + O(N^{-2})\right]^2$$

or

$$(3.11) \qquad V(\hat{I}_u) = E[\sum_{i=1}^{n} u_i(\hat{p}_i - p_i)\{1 + \log(p_i/q_i)\}]^2 + O(N^{-2}) .$$

The first term of (3.11) is equal to

$$E[\sum_{i=1}^{n} u_i^2(\hat{p}_i - p_i)^2\{1 + \log(p_i/q_i)\}^2 +$$
$$+ \sum_{i \neq j} u_i u_j(\hat{p}_i - p_i)(\hat{p}_j - p_j)\{1 + \log(p_i/q_i)\}\{1 + \log(p_j/q_j)\}] =$$

$$= \frac{1}{N}\left[\sum_{i=1}^{n} u_i^2 p_i(1-p_i)\{1+\log(p_i/q_i)\}^2 - \sum_{i \neq j} u_i u_j p_i p_j \{1+\log(p_i/q_i)\}\{1+\log(p_j/q_j)\}\right] =$$

$$= \frac{1}{N}\left[\sum_{i=1}^{n} u_i^2 p_i + \sum_{i=1}^{n} u_i^2 p_i \log^2(p_i/q_i) + 2\sum_{i=1}^{n} u_i^2 p_i \log(p_i/q_i) - \sum_{i=1}^{n} u_i^2 p_i^2 - \right.$$

$$- \sum_{i=1}^{n} u_i^2 p_i^2 \log^2(p_i/q_i) - 2\sum_{i=1}^{n} u_i^2 p_i^2 \log(p_i/q_i) -$$

$$\left. - \sum_{i \neq j} u_i u_j p_i p_j \log(p_i/q_i)\log(p_j/q_j) - \sum_{i \neq j} u_i u_j p_i p_j \{1 + \log(p_i/q_i) + \log(p_j/q_j)\}\right] =$$

$$= \frac{1}{N}\left[\sum_{i=1}^{n} u_i^2 p_i \log^2(p_i/q_i) - I_u^2\right] +$$

$$+ \frac{1}{N}\left[\sum_{i=1}^{n} \{u_i^2 p_i + 2u_i^2 p_i \log(p_i/q_i) - u_i^2 p_i^2 - 2u_i^2 p_i^2 \log(p_i/q_i) - \right.$$

$$\left. - u_i p_i \sum_{j=1}^{n} u_j p_j [1 + \log(p_i/q_i) + \log(p_j/q_j)] + u_i^2 p_i^2 [1 + 2\log(p_i/q_i)]\}\right] =$$

$$= \frac{1}{N}\left[\sum_{i=1}^{n} u_i^2 p_i \log^2(p_i/q_i) - I_u^2\right] +$$

$$+ \frac{1}{N}\left[\sum_{i=1}^{n} u_i^2 p_i + 2\sum_{i=1}^{n} u_i^2 p_i \log(p_i/q_i) - \left(\sum_{i=1}^{n} u_i p_i\right)\left(\sum_{j=1}^{n} u_j p_j\right) - \right.$$

$$\left. - \left\{\sum_{i=1}^{n} u_i p_i \log(p_i/q_i)\right\}\left(\sum_{j=1}^{n} u_j p_j\right) - \left\{\sum_{j=1}^{n} u_j p_j \log(p_j/q_j)\right\}\left(\sum_{i=1}^{n} u_i p_i\right)\right] =$$

$$= \frac{1}{N}\left[\sum_{i=1}^{n} u_i^2 p_i \log^2(p_i/q_i) - I_u^2\right] + \frac{1}{N}\sum_{i=1}^{n} u_i^2 p_i \{1 + 2\log(p_i/q_i)\} - \frac{\bar{u}^2}{N} - \frac{2\bar{u}I_u}{N} =$$

$$= \frac{1}{N}\left[\sum_{i=1}^{n} u_i^2 p_i \log^2(p_i/q_i)\right] + \frac{1}{N}\left[\sum_{i=1}^{n} u_i^2 p_i \log e(p_i/q_i)^2\right] - \frac{1}{N}\left[I_u^2 + \bar{u}^2 + 2\bar{u}I_u\right] =$$

$$= \frac{1}{N}\left[\sum_{i=1}^{n} u_i^2 p_i \{\log^2(p_i/q_i) + \log e(p_i/q_i)^2\}\right] - \frac{(I_u + \bar{u})^2}{N}$$

(Taking log e $= 1$, natural units.) Thus (3.3) follows.

Now in (3.2) we have $U > \bar{u}$, thus the estimate $\hat{I}_u$ overestimates the true value of I_u. Also it is clear from (3.2) and (3.3) that when $N \to \infty$, then $E(\hat{I}_u) \to I_u$ and $V(\hat{I}_u) \to 0$. Thus we conclude that $\hat{I}_u$ is a consistent estimator of I_u. Further, when $u_i = 1$ for all i, then obviously (3.2) and (3.3) reduce to (2.4) and (2.5), respectively.

Acknowledgements. The author is thankful to Dr. R. K. Tuteja, M. D. University, Rohtak, for his guidance in the preparation of this paper. He is thankful to Dr. Ashok Kumar, M. D. University, Rohtak, for discussion at various stages. Thanks are also extended to the referee for his valuable suggestions on the earlier versions of the paper.

References

[1] *T. W. Anderson:* An Introduction to Multivariate Statistical Analysis. John Wiley and Sons Inc., New York, 1958.
[2] *M. Belis, S. Guiasu:* A Quantitative-Qualitative Measure of Information in Cybernetic Systems. IEEE Trans. Inform. Theory, 14 (1968), 593–594.
[3] *S. Kullback:* Information Theory and Statistics. Wiley, New York, 1959.
[4] *G. Longo:* A Quantiative-Qualitative Measure of Information. Springer-Verlag, 1972.
[5] *H. C. Taneja, R. K. Tuteja:* Characterization of a Quantitative-Qualitative Measure of Relative Information. Information Sciences, 33 (1984).
[6] *H. Theil:* Economics and Information Theory. North-Holland Publishing Co., Amsterdam, 1967.

Souhrn

O STŘEDNÍ HODNOTĚ A ROZPTYLU ODHADŮ KULLBACKOVY INFORMACE A MÍRY RELATIVNÍ „UŽITEČNÉ" INFORMACE

HARISH C. TANEJA

V článku je odvozena střední hodnota a rozptyl maximálně věrohodného odhadu Kullbackovy míry informace a míry relativní „užitečné" informace.

Author's address: Dr. *Harish C. Taneja,* Department of Mathematics, Guru Nanak Dev University, Amritsar - 143005, India.