Z. Wahrscheinlichkeitstheorie verw. Geb. 21, 215–224 (1972)

Axiomatic Characterizations of Some Measures of Divergence in Information

J. ACZÉL and P. NATH

Dedicated to the Memory of A. Rényi on his 50th Birthday

1.

Let $(\Omega, \mathscr{B}, \mathscr{P})$ be a probability space, that is, Ω is an abstract set, $\mathscr{B}$ a σ-algebra of subsets of Ω, $\Omega \in \mathscr{B}$ and $\mathscr{P}$ a probability measure, i.e., a nonnegative countably additive set function defined on $\mathscr{B}$, such that $\mathscr{P}(\Omega)=1$. Let Δ_n denote the set of all generalized probability distributions involving n events, as defined by Rényi [6]. We will deal with these *generalized* finite discrete probability distributions $P=\{p_1, p_2, \ldots, p_n\}$, $\sum_{k=1}^{n} p_k \leqq 1$, $p_k>0$ $(k=1,2,\ldots,n)$, rather than the ordinary complete ones, among others because not all outcomes of an experiment or market situation, for instance, need to be relevant, significant and/or observable. If $\sum_{k=1}^{n} p_k=1$, then the distribution is *complete*, else *incomplete*.

Also, observers usually do not observe and even less forecast phenomena exactly and their observations and estimations differ from each other. Let A and B be two independent observers (estimators) who assert that the probabilities associated with the same experiment (system of events), are $\{q_1, q_2, \ldots, q_n\}=Q\in\Delta_n$, $\{r_1, r_2, \ldots, r_n\}=R\in\Delta_n$ $(n=1,2,3,\ldots)$, respectively. We discuss the following question.

What is the amount of the directed divergence between the estimations Q and R of observers A and B of the (generalized) probability distribution of an experiment which actually is P?

As an answer, one of us [5] proposed the following two measures of generalized directed divergences in information:

$$I_1(P\|Q|R)=\sum_{k=1}^{n} p_k \log_2 \frac{q_k}{r_k} \Big/ \sum_{k=1}^{n} p_k, \tag{1}$$

$$I_\alpha(P\|Q|R)=(\alpha-1)^{-1} \log_2 \left(\sum_{k=1}^{n} p_k q_k^{\alpha-1} r_k^{1-\alpha} \Big/ \sum_{k=1}^{n} p_k\right) \quad (\alpha \neq 1). \tag{2}$$

Remark. In [5], these were called measures of error. However, Kerridge [2] has interpreted

$$I(P\|Q)=\sum_{k=1}^{n} p_k \log_2 \frac{p_k}{q_k} \quad \left(\sum_{k=1}^{n} p_k=\sum_{k=1}^{n} q_k=1\right)$$

as error made by the observer in estimating as Q a (complete) probability distribution which really is P (sometimes it is called information gain, for instance Rényi [7]). It is obvious that for $\sum_{k=1}^{n} p_k = \sum_{k=1}^{n} q_k = \sum_{k=1}^{n} r_k = 1$ we have

$$I_1(P\|Q|R) = I(P\|R) - I(P\|Q).$$

So $I_1(P\|Q|R)$ is the difference of two errors, so rather a divergence than an error itself. Since $I(P\|Q) = I_1(P\|P|Q)$ in this case, we could call our $I_1(P\|Q|R)$ and $I_\alpha(P\|Q|R)$ also generalized errors. However, for $P=\{p\}$, $Q=\{q\}$, $R=\{r\}$ (one-event distributions) from (1) and (2)

$$I_\alpha(\{p\}\|\{q\}|\{r\}) = \log_2 \frac{q}{r} \quad \text{both for } \alpha = 1 \text{ and } \alpha \neq 1,$$

so it does not even depend on p, and could not be called therefore an error in estimating p.

On the other hand, Kullback [3] calls

$$I(P\|Q) = I_1(P\|P|Q) = \sum_{k=1}^{n} p_k \log_2 \frac{p_k}{q_k} \quad \left(\sum_{k=1}^{n} p_k = \sum_{k=1}^{n} q_k = 1\right)$$

a "directed divergence", one reason more to call our quantities *generalized directed divergences.*

The object of the present paper is to give some axiomatic characterizations of the above quantities (1) and (2). We will see that we can give characterizations involving only one-event distributions and n-event distributions for a fixed n, say $n=2$.

2.

We assume the following postulates.

Postulate 1. $I(\{p\}\|\{q\}|\{r\}) \gtreqless 0$ *according as* $q \gtreqless r$ $(p, q, r \in (0, 1])$.

This postulates that our measure of divergence be *directed*, at least for one-event distributions: positive (negative) if q is greater (smaller) than r, and 0 if $q=r$.

Postulate 2. $I(\{1\}\|\{1\}|\{\tfrac{1}{2}\}) = 1$.

This postulate simply determines the *unit* of divergence in information as the divergence between A estimating (correctly) 1 as the probability of a certain event, and B estimating it indifferently as $\frac{1}{2}$.

Postulate 3.

$$\begin{aligned} &I(\{p_1 x, \dots, p_n x\}\|\{q_1 y, \dots, q_n y\}|\{r_1 z, \dots, r_n z\}) \\ &\quad = I(\{p_1, \dots, p_n\}\|\{q_1, \dots, q_n\}|\{r_1, \dots, r_n\}) + I(\{x\}\|\{y\}|\{z\}), \\ &\qquad (\{p_1, \dots, p_n\}, \{q_1, \dots, q_n\}, \{r_1, \dots, r_n\} \in \Delta_n,\ x, y, z \in (0, 1]). \end{aligned}$$

We suppose this only for $n=1$ and $n=2$.

This is an *additivity* postulate (weaker than the requirement of general additivity): If there is an n-event distribution P (in our case $n=2$) and an independent one-event distribution $\{x\}$ and observers A and B give the estimations Q, $\{y\}$ and R, $\{z\}$ respectively, then the direct product of the two distributions (as the distribution of the performance of both the n-relevant-outcomes-experiment and the one-relevant-outcome-experiment) will be estimated by them with the direct products of the original estimations. The postulate states that we get the divergence of information for the combined experiment as sum of the divergences of the single experiments.

Postulate 4. *If $p_1, p_2, q_1, q_2, r_1, r_2$ are positive and such that $p_1+p_2\leqq 1, q_1+q_2\leqq 1$, $r_1+r_2\leqq 1$, then there exists a continuous and strictly monotonic real function ϕ such that*

$$I(\{p_1, p_2\}\|\{q_1, q_2\}|\{r_1, r_2\})$$
$$=\phi^{-1}([p_1\,\phi\,I(\{p_1\}\|\{q_1\}|\{r_1\})+p_2\,\phi\,I(\{p_2\}\|\{q_2\}|\{r_2\})]/[p_1+p_2]).$$

This is a *quasilinearity* condition: The divergence of information on the union of two (in general, n) one-event incomplete distributions is the quasilinear mean of the divergences of information on the single distributions weighted with the actual probabilities.

The fact, that we do not use 0-probabilities in our postulates, makes them weaker.

3.

We will prove the following theorem.

Theorem 1. *Let $I(\{p\}\|\{q\}|\{r\})$ and $I(\{p_1, p_2\}\|\{q_1, q_2\}|\{r_1, r_2\})$ be defined for $p, p_1, p_2, q, q_1, q_2, r, r_1, r_2, p_1+p_2, q_1+q_2, r_1+r_2 \in (0, 1]$. If the postulates 1, 2, 3, 4 are satisfied, then and only then*

$$I(\{p\}\|\{q\}|\{r\})=\log_2\frac{q}{r} \tag{3}$$

and ϕ in Postulate 4 is either a linear function or a linear function of an exponential function. In the first case

$$\begin{aligned}I(\{p_1, p_2\}\|\{q_1, q_2\}|\{r_1, r_2\})&=\left(p_1\log_2\frac{q_1}{r_1}+p_2\log_2\frac{q_2}{r_2}\right)\Big/(p_1+p_2)\\&=I_1(\{p_1, p_2\}\|\{q_1, q_2\}|\{r_1, r_2\})\end{aligned} \tag{4}$$

in the second

$$\begin{aligned}&I(\{p_1, p_2\}\|\{q_1, q_2\}|\{r_1, r_2\})\\&\quad=(\alpha-1)^{-1}\log_2[(p_1\,q_1^{\alpha-1}\,r_1^{1-\alpha}+p_2\,q_2^{\alpha-1}\,r^{1-\alpha})/(p_1+p_2)]\\&\quad=I_\alpha(\{p_1, p_2\}\|\{q_1, q_2\}|\{r_1, r_2\})\qquad(\alpha\neq 1).\end{aligned} \tag{5}$$

Remarks. While all our *conditions* involve only *positive* probabilities, the *results* can be *extended for* $p=0$ *and* $p_1=0$ *or* $p_2=0$ $(p_1+p_2>0)$:

$$I(\{0\}\,\|\,\{q\}\,|\,\{r\})=\log_2\frac{q}{r}=I(\{0,p\}\,\|\,\{q_1,q\}\,|\,\{r_1,r\})$$
$$=I(\{p,0\}\,\|\,\{q,q_2\}\,|\,\{r,r_2\}).$$

There is no finite extension, if one of the q's or r's is 0 without (for $n\geqq 2$) also the respective p's being 0.

Proof of Theorem 1. We denote

$$f(p,q,r):=I(\{p\}\,\|\,\{q\}\,|\,\{r\})\qquad \big(p,q,r\in(0,1]\big).\tag{6}$$

Postulate 3 gives for $n=1$

$$f(p\,x,q\,y,r\,z)=f(p,q,r)+f(x,y,z).\tag{7}$$

By applying (7) twice we get

$$f(p,q,r)=g(p)+h(q)+k(r)\tag{8}$$

where

$$g(p):=f(p,1,1),\qquad h(q):=f(1,q,1),\qquad k(r):=f(1,1,r).\tag{9}$$

Now, by Postulate 1, $f(p,q,q)=0$, so, from (8)

$$g(p)+h(q)+k(q)=0\quad\text{or}\quad g(p)=c=-h(q)-k(q)$$

and (8) goes over into

$$f(p,q,r)=c+h(q)-c-h(r)=h(q)-h(r).\tag{10}$$

Also by postulate 1 we have

$$h(q)-h(r)=f(p,q,r)>0\qquad\text{if }\ q>r$$

that is, h is an *increasing* function. But (7) gives with $p=x=r=z=1$, by (9),

$$h(q\,y)=h(q)+h(y)\qquad\text{for all }\ q,y\in(0,1].$$

So (cf. [1], p. 39), $h(q)=c\log_2 q$ and, by (10) and by Postulate 2, $1=h(1)-h(\tfrac{1}{2})=c$. Thus (10) goes over into

$$f(p,q,r)=\log_2\frac{q}{r},$$

that is [cf. (6)], into (3).

Apply now Postulate 4 (we write log for $\log_2$)

$$I(\{p_1,p_2\}\,\|\,\{q_1,q_2\}\,|\,\{r_1,r_2\})=\phi^{-1}\left(\frac{p_1\,\phi\left(\log\dfrac{q_1}{r_1}\right)+p_2\,\phi\left(\log\dfrac{q_2}{r_2}\right)}{p_1+p_2}\right)$$

and Postulate 3 for $n=2$:

$$\phi^{-1}\left(\frac{p_1 x\,\phi\left(\log\frac{q_1 y}{r_1 z}\right)+p_2 x\,\phi\left(\log\frac{q_2 y}{r_2 z}\right)}{p_1 x+p_2 x}\right)$$

$$=\phi^{-1}\left(\frac{p_1\,\phi\left(\log\frac{q_1}{r_1}\right)+p_2\,\phi\left(\log\frac{q_2}{r_2}\right)}{p_1+p_2}\right)+\log\frac{y}{z}$$

or, with $t=\log\frac{q_1}{r_1}$, $u=\log\frac{q_2}{r_2}$, $v=\log\frac{y}{z}$,

$$\phi^{-1}\left(\frac{p_1\phi(t+v)+p_2\,\phi(u+v)}{p_1+p_2}\right)=\phi^{-1}\left(\frac{p_1\,\phi(t)+p_2\,\phi(u)}{p_1+p_2}\right)+v, \qquad [t,u,v\in(-\infty,\infty)]. \tag{11}$$

Thus (by Theorem 3, p. 159 of [1]) we have

$$\phi(x)=a x+b \quad\text{or}\quad \phi(x)=a\,2^{(\alpha-1)x}+b, \qquad a\neq 0,\ \alpha\neq 1\ [x\in(-\infty,\infty)]. \tag{12}$$

If we substitute these back into Postulate 4, we get (4) and (5) respectively. Since, on the other hand, (3), (4), (5) satisfy all our postulates, this concludes the proof of Theorem 1.

We can see easily that we have used from Postulate 1 only the following two consequences of this Postulate.

Postulate 5. *$I(\{1\}\|\{q\}\mid\{r\})\geqq 0$ if $q>r$.*

Postulate 6. *$I(\{p\}\|\{r\}\mid\{r\})=0$ for all $p, r\in(0,1]$.*

So we have the following

Corollary. *Theorem 1 remains true, if the Postulate 1 is replaced by the weaker Postulates 5 and 6.*

4.

Let us see now, how Postulate 6 could be further weakened.

It surely is stronger than the following.

Postulate 7.

$$I(\{p+q\}\|\{r\}\mid\{r\})=I(\{p\}\|\{r\}\mid\{r\})+I(\{q\}\|\{r\}\mid\{r\}) \qquad (p,q,p+q,r\in(0,1]).$$

Indeed, $I(\{p\}\|\{r\}\mid\{r\})=0$ (Postulate 6) satisfies Postulate 7, but the later may also have other solutions.

Whether Postulate 7 is "natural" is a question of taste. – We prove here the following.

Theorem 2. *Postulates 2, 3, 4, 5, and 7 already characterize the divergences (3) and either (4) or (5).*

Proof. Postulate 7 means in terms of the notation (6)

$$f(p+q,r,r)=f(p,r,r)+f(q,r,r) \quad \text{whenever } p,q,p+q,r\in(0,1] \tag{13}$$

or, with $r=1$, by (9)

$$g(p+q)=g(p)+q(q) \quad (p,q,p+q\in(0,1] \text{ else arbitrary}). \tag{14}$$

On the other hand, Postulate 3 or (7) with $q=r=y=z=1$ gives

$$g(px)=g(p)+g(x) \quad (p,x\in(0,1], \text{ else arbitrary}). \tag{15}$$

By comparing these two equations we get

$$g(pq)=g(p+q) \quad \text{whenever } p,q,p+q\in(0,1]. \tag{16}$$

Let us put $q=1-p$ $[p\in(0,1)]$ into (16):

$$g[p(1-p)]=g(1)=c \quad \text{(constant)}.$$

The function $p\to p(1-p)$ takes every value in $(0,\frac{1}{4}]$ as p runs through $(0,1)$, so we have

$$g(t)=c \quad \text{for all } t\in(0,\tfrac{1}{4}]. \tag{17}$$

Now put $q=\frac{1}{4}$ into (16) in order to get

$$g(p+\tfrac{1}{4})=g\left(\frac{p}{4}\right) \quad \text{for all } p\in(0,\tfrac{3}{4}].$$

But, then $\frac{p}{4}\in(0,\frac{3}{16}]\subset(0,\frac{1}{4}]$ so, by comparison with (17),

$$g(t)=c \quad \text{for all } t\in(\tfrac{1}{4},1]$$

and together with (17) they yield

$$g(t)=c \quad \text{for all } t\in(0,1].$$

If we put this back, for instance into (14), we get $c=2c$, $c=0$ and

$$g(p)=0 \quad \text{for all } p\in(0,1].$$

So, by (8),

$$f(p,q,r)=h(q)+k(r) \tag{18}$$

and, by substitution into (13), $h(r)+k(r)=2[h(r)+k(r)]$ or $k(r)=-h(r)$ and from (18)

$$f(p,q,r)=h(q)-h(r). \tag{10}$$

From here on the proof goes as that of Theorem 1.

5.

If we supplement Postulate 5 by

Postulate 8. $I(\{1\} \| \{r\} | \{r\}) = 0$,

then, by a similar but slightly more minutious proof than that of Theorem 2, the Postulate 7 can be replaced by

Postulate 9.

$$I(\{p+q\} \| \{1\} | \{1\}) = \frac{I(\{p\} \| \{1\} | \{1\}) + I(\{q\} \| \{1\} | \{1\})}{2}$$

whenever $p, q, p+q \in (0, 1]$,

which some people might consider more natural than Postulate 7.

Theorem 3. *Postulates* 2, 3, 4, 5, 8, *and* 9 *characterize the divergences* (3) *and* (4) *or* (5).

We leave the easily reconstructable proof aside, since in our opinion anyway the following condition seems to be more natural than either of Postulates 7 and 9.

Postulate 10.

$$I(\{p+q\} \| \{1\} | \{1\}) = \frac{p\, I(\{p\} \| \{1\} | \{1\}) + q\, I(\{q\} \| \{1\} | \{1\})}{p+q} \qquad (p, q, p+q \in (0, 1]).$$

If we use Postulate 10 instead of Postulate 7, then in place of Eq. (14) we have

$$(p+q)\, g(p+q) = p\, g(p) + q\, g(q).$$

With the notation

$$d(p) := p\, g(p) \qquad (p \in (0, 1]) \tag{19}$$

this gives

$$d(p+q) = d(p) + d(q), \qquad \text{whenever } p, q, p+q \in (0, 1]. \tag{20}$$

On the other hand, Eq. (15) gives with the notation (19)

$$d(p\, q) = p\, d(q) + q\, d(p) \qquad (p, q \in (0, 1]). \tag{21}$$

Now, all "derivatives" (cf. [8]) satisfy the system of functional Eqs. (20), (21). Every continuous derivative is 0, but there do exist noncontinuous derivatives (cf. [8]). So it cannot be asserted without any further supposition that $d \equiv 0$ (that is, $g \equiv 0$). We can postulate for instance this:

Postulate 11. *There exist two values* $q_0, r_0 \in (0, 1]$ *and an (arbitrarily small) interval* $(p_0, p_1) \subseteq (0, 1]$ *such that* $I(\{p\} \| \{q_0\} | \{r_0\}) \geqq 0$ *for all* $p \in (p_0, p_1)$ *(the interval could be replaced by a set of positive measure—or by a set whose translations by all rational numbers fill the real line* [4], *and* 0 *could be replaced by any other lower or upper bound).*

After this preparation we are able to prove the following theorem.

Theorem 4. *Postulates* 2, 3, 4, 5, 8, 10, *and* 11 *characterize the divergences* (3) *and* (4) *or* (5).

Proof. As in the proof of Theorem 1 we have

$$f(p,q,r)=g(p)+h(q)+k(r). \tag{8}$$

By Postulate 11,

$$0 \leqq f(p,q_0,r_0)=g(p)+h(q_0)+k(r_0), \quad \text{so} \quad g(p) \geqq -h(q_0)-k(r_0) \text{ in } (p_0,p_1),$$

which means that g and, because of (19), also d is bounded from below on an interval. So (cf. [1], pp. 32–35, 48–49), Eq. (20) implies $d(p)=cp$ which satisfies (21) only if $c=0$, $d(p)\equiv 0$. Thus, by (19), $g\equiv 0$, (8) reduces to (18) $f(p,q,r)=h(q)+k(r)$ and Postulate 8 gives

$$h(r)+k(r)=0$$

and

$$f(p,q,r)=h(q)-h(r). \tag{10}$$

From here on everything proceeds as in the proof of Theorem 1 and, since (3) satisfies also Postulates 10 and 11, the Theorem 4 is proved.

6.

We have seen that in Theorem 1 (and all the subsequent theorems) we get I_1 or I_α ($\alpha \neq 1$) according to which of the functions ϕ in (12) we choose. If we replace Postulate 4 in any of these theorems by

$$I(\{p_1,p_2\} \| \{q_1,q_2\} | \{r_1,r_2\}) = \frac{p_1 I(\{p_1\} \| \{q_1\} | \{r_1\}) + p_2 I(\{p_2\} \| \{q_2\} | \{r_2\})}{p_1+p_2}$$
$$(p_1,p_2,p_1+p_2,q_1,q_2,q_1+q_2,r_1,r_2,r_1+r_2 \in (0,1]),$$

then we have characterized the divergence

$$I(\{p_1,p_2\} \| \{q_1,q_2\} | \{r_1,r_2\}) = I_1(\{p_1,p_2\} \| \{q_1,q_2\} | \{r_1,r_2\}).$$

Another characteristic property of I_1 is that it satisfies a "Sincov-type functional equation" (cf. [1], p. 223). We formulate it as the following postulate.

Postulate 12.

$$I(\{p_1,\dots,p_n\} \| \{q_1,\dots,q_n\} | \{r_1,\dots,r_n\}) + I(\{p_1,\dots,p_n\} \| \{r_1,\dots,r_n\} | \{s_1,\dots,s_n\})$$
$$= I(\{p_1,\dots,p_n\} \| \{q_1,\dots,q_n\} | \{s_1,\dots,s_n\})$$
$$(\{p_1,\dots,p_n\}, \{q_1,\dots,q_n\}, \{r_1,\dots,r_n\}, \{s_1,\dots,s_n\} \in \Delta_n,\ n=1,2).$$

This postulate means that the divergence between A and C is the sum of the divergences between A and B and between B and C. We prove the following theorem.

Theorem 5. *Postulates* 2, 4, 5, 12 *and Postulate* 3 *for* $n=1$ *characterize the divergence* (3) *for* $n=1$ *and the divergence* I_1 *as defined in* (4) *for* $n=2$.

Proof. By Postulate 12 for $n=1$ and by (6) and (8) (which was derived from Postulate 3)

$$g(p)+h(q)+k(r)+g(p)+h(r)+k(s)=g(p)+h(q)+k(s)$$

or

$$g(p) = -h(r) - k(r) = c \quad \text{(constant)}.$$

So we have (10). By Postulate 5 again h is non-decreasing and, by Postulate 3, $h(pq) = h(p) + h(q)$. Thus we get $h(q) = c \log_2 q$ and by Postulate 2 again $c = 1$ and $f(p, q, r) = \log_2 \frac{q}{r}$ or, with (6),

$$I(\{p\} \,\|\, \{q\} \,|\, \{r\}) = \log_2 \frac{q}{r}. \tag{3}$$

From Postulates 4 and 12 (for $n = 2$) we have now

$$\phi^{-1}\left(\frac{p_1 \phi[\log(q_1/r_1)] + p_2 \phi[\log(q_2/r_2)]}{p_1 + p_2}\right) + \phi^{-1}\left(\frac{p_1 \phi[\log(r_1/s_1)] + p_2 [\log(r_2/s_2)]}{p_1 + p_2}\right)$$
$$= \phi^{-1}\left(\frac{p_1 \phi[\log(q_1/s_1)] + p_2 \phi[\log(q_2/s_2)]}{p_1 + p_2}\right).$$

If we denote $t = \log(q_1/r_1)$, $u = \log(q_2/r_2)$, $v = \log(r_1/s_1)$, $w = \log(r_2/s_2)$ we get

$$\phi^{-1}\left(\frac{p_1 \phi(t+v) + p_2 \phi(u+w)}{p_1 + p_2}\right) = \phi^{-1}\left(\frac{p_1 \phi(t) + p_2 \phi(u)}{p_1 + p_2}\right) + \phi^{-1}\left(\frac{p_1 \phi(v) + p_2 \phi(w)}{p_1 + p_2}\right). \tag{22}$$

If we specialize $w = v$, then (22) goes over into (11) (which shows that in our case Postulate 3 for $n = 2$ follows from the others) and, as in the proof of the Theorem 1, we have (12). Of the two types of functions in (12), however, the second does not satisfy (22) for all real t, u, v, and w, so we have $\phi(t) = at + b$ and Theorem 5 is proved. An alternative proof is the following.

Denote in (22) $x = p_1/(p_1 + p_2)$, $y = p_2/(p_1 + p_2)$, and

$$F(x, y;\, t, u) := \phi^{-1}[x\,\phi(t) + y\,\phi(u)]. \tag{23}$$

Then (22) goes over into

$$F(x, y;\, t+v, u+w) = F(x, y;\, t, u) + F(x, y;\, v, w) \quad \text{for all } x, y \in (0, 1);\; t, u, v, w \in (-\infty, \infty).$$

This is a Cauchy equation in the last two variables (cf. [1], pp. 214–216). Since with ϕ also F is continuous [cf. (23)], we have necessarily

$$F(x, y;\, t, u) = A(x, y)\,t + B(x, y)\,u. \tag{24}$$

By comparing (23) and (24) we get

$$\phi[A(x, y)\,t + B(x, y)\,u] = x\,\phi(t) + y\,\phi(u).$$

This and the continuity of ϕ are more than enough to warrant the linearity of ϕ (see Theorem 2 on p. 67 of [1]), so $\phi(t) = at + b$ and (4) holds. Since (3) and (4) satisfy also Postulate 12, the Theorem 5 is proved.

This research was supported in part by the National Research Council of Canada Grant Nr. A-2972.

References

1. Aczél, J.: Lectures on functional equations and their applications. New York: Academic Press 1966.
2. Kerridge, D. F.: Inaccuracy and inference. J. roy. statist. Soc. Ser **B**, 184–194 (1961).
3. Kullback, S.: Information theory and statistics. New York: J. Wiley & Sons 1959.
4. Moszner, Z.: Solution 1 of problem 49. Aequationes math. **4**, 258 (1970).
5. Nath. P.: On the measures of errors in information. J. math. Sci. Delhi **3**, 1–16 (1968).
6. Rényi, A.: On measures of entropy and information. Proc. Fourth Berkeley Sympos. math. Statist. Probab. 1960 **1**, 547–561 (1961).
7. – Wahrscheinlichkeitsrechnung. Anhang: Einführung in die Informationstheorie. § 8. Berlin: D. V. W. 1962.
8. Zariski, O., Samuel, P.: Commutative algebra I. Chapter II. Princeton: Van Nostrand 1958.

Professor J. Aczél
University of Waterloo
Faculty of Mathematics
Waterloo, Ontario
Canada
and Istituto Nazionale di Alta Matematica
Città Universitaria
I-00100 Roma, Italy

Dr. P. Nath
University of Delhi
Faculty of Mathematics
Delhi, India
and McMaster University
Department of Mathematics
Hamilton, Ontario
Canada

(Received April 28, 1970)