6. On the Écart between Two "Amounts of Information"

By Kōmei SUZUKI

(Comm. by K. KUNUGI, M.J.A., Jan. 12, 1957)

§ 1. $d(\lambda_1, \lambda_2; \Lambda)=\sum_{i=0}^{\infty} \Delta P_i \log\left(1+\frac{\Delta P_i}{P_i}\right)$

As was shown in the preceding paper the "amount of information"[2)-4)] has been defined by a specified probability space (or distribution), $(R, \mathfrak{X}, \lambda)$, and the partition,[1)] Λ, imposed on the space R. And we have conventionally denoted it by $H(\lambda; \Lambda)$. As usual

$$\Lambda : R=\bigcup_{i=0}^{\infty} A_i,\ A_i \in \mathfrak{X},\ A_i \cap A_j=0 \quad (i \neq j).$$

For any two distributions $(R, \mathfrak{X}, \lambda_1)$ and $(R, \mathfrak{X}, \lambda_2)$, providing

(a) $\lambda_1(A_i)=P_i \geq 0,\ \lambda_2(A_i)=P_i+\Delta P_i \geq 0,\ \sum_i P_i=\sum_i (P_i+\Delta P_i)=1$

(b) the series $H(\lambda_1; \Lambda)=\sum_i P_i \log 1/P_i$ and $H(\lambda_2; \Lambda)=\sum_i (P_i+\Delta P_i) \log 1/(P_i+\Delta P_i)$ to converge

(c) $-1+\alpha \leq \Delta P_i/P_i \leq k;\ k>0,\ 1>\alpha>0$ for all i,

we have directly from the result obtained in the preceding paper

$$0 \leq \Delta H-\sum_{i=0}^{\infty} \Delta P_i \log \frac{1}{P_i+\Delta P_i} \leq \sum_{i=0}^{\infty} \Delta P_i \log\left(1+\frac{\Delta P_i}{P_i}\right)$$

where $\Delta H=H(\lambda_2; \Lambda)-H(\lambda_1; \Lambda)$.

Denoting $\sum_{i=0}^{\infty} \Delta P_i \log\left(1+\frac{\Delta P_i}{P_i}\right)$ by $d(\lambda_1, \lambda_2; \Lambda)$, we have easily

(1.1)
(a) $d(\lambda, \lambda; \Lambda)=0$
(b) $d(\lambda_1, \lambda_2; \Lambda) \geq 0$ for $\lambda_1 \neq \lambda_2$
(c) $d(\lambda_1, \lambda_2; \Lambda)=d(\lambda_2, \lambda_1; \Lambda)$.

It must be noted that we could not avoid the sign of equality in (b) of (1.1); because even though $\lambda_1 \neq \lambda_2$, we would often have that $\lambda_1(A_i)=\lambda_2(A_i)$, $i=0, 1, 2, \cdots$, for some partitions imposed on R.

To appreciate more fully we consider a distribution $(R, \mathfrak{X}, \lambda_3)$ together with the above $(R, \mathfrak{X}, \lambda_1)$ and $(R, \mathfrak{X}, \lambda_2)$.

Providing again the following

(d) $\lambda_1(A_i)=P_i^{(1)},\ \lambda_2(A_i)=P_i^{(2)}=P_i^{(1)}+\Delta P_i^{(1)},\ \lambda_3(A_i)=P_i^{(3)}=P_i^{(2)}+\Delta P_i^{(2)}$

(e) $-1+\alpha \leq \Delta P_i^{(\nu)}/P_i^{(\nu)} \leq k;\ 1>\alpha>0,\ k>0,\ i=0, 1, 2, \cdots,\ \nu=1, 2$

we have

(1.2)

$$d(\lambda_3, \lambda_1; \Lambda)-\{d(\lambda_1, \lambda_2; \Lambda)+d(\lambda_2, \lambda_3; \Lambda)\} = \sum_{i=0}^{\infty}(\Delta P_i^{(1)} \log P_i^{(3)}/P_i^{(2)}+\Delta P_i^{(2)} \log P_i^{(2)}/P_i^{(1)})$$

and

$$(\Delta P_i^{(1)} \log P_i^{(3)}/P_i^{(2)}+\Delta P_i^{(2)} \log P_i^{(2)}/P_i^{(1)})\begin{cases}>0 \leftrightarrow \Delta P_i^{(1)} \cdot \Delta P_i^{(2)}>0 \\ =0 \leftrightarrow \Delta P_i^{(1)} \cdot \Delta P_i^{(2)}=0 \\ <0 \leftrightarrow \Delta P_i^{(1)} \cdot \Delta P_i^{(2)}<0.\end{cases}$$