

Appendix:

In this section we supply the proofs of several assertions made earlier.

Proof of Theorem 1: ($\{X, Y, \dots, Z\} > 0$ unless independent)

Abbreviate $P(X_i, Y_j, \dots, Z_k)$ by $P_{ij\dots k}$, and let

$$(A.1) \quad P_{ij\dots k} = \begin{cases} \frac{P_{ij\dots k}}{P_i P_j \dots P_k} & \text{if } P_i P_j \dots P_k > 0 \\ 1 & \text{if } P_i P_j \dots P_k = 0 \end{cases}$$

so that

$$(A.2) \quad P_{ij\dots k} = Q_{ij\dots k} P_i P_j \dots P_k$$

then

$$(A.3) \quad \{X, Y, \dots, Z\} = \exp \left[\ln \frac{P_{ij\dots k}}{P_i P_j \dots P_k} \right] = \exp \left[\ln Q_{ij\dots k} \right] = \sum_{ij\dots k} P_i P_j \dots P_k Q_{ij\dots k} \ln Q_{ij\dots k}$$

making use of the inequality

$$(A.4) \quad x \ln x \geq x - 1 \quad (\text{except for } x=1) \quad (x>0)$$

which is easily established by calculating the minimum of $x \ln x - (x-1)$ we have

$$(A.5) \quad P_i P_j \dots P_k Q_{ij\dots k} \ln Q_{ij\dots k} \geq P_i P_j \dots P_k (Q_{ij\dots k} - 1) \quad (\text{unless } Q_{ij\dots k} = 1)$$

and hence the sum

$$(A.6) \quad \sum_{ij\dots k} P_i P_j \dots P_k Q_{ij\dots k} \ln Q_{ij\dots k} \geq \sum_{ij\dots k} P_i P_j \dots P_k Q_{ij\dots k} - \sum_{ij\dots k} P_i P_j \dots P_k \quad (\text{unless all } Q_{ij\dots k} = 1)$$

$$\text{But } \sum_{ij\in K} P_i P_j P_K Q_{ij\in K} = \sum_{ij\in K} P_{ij\in K} = 1, \text{ and } \sum_{ij\in K} P_i P_j P_K = 1$$

so that the right side of (1.6) vanishes. But the left side is, by (1.3) $\{X, Y, Z\}$, and the condition that all the $Q_{ij\in K} = 1$ is precisely the independence condition (II-) so we have proved that

$$(1.7) \quad \{X, Y, Z\} > 0 \quad (\text{unless } X, Y, Z \text{ mutually independent})$$

PED

We now wish to establish some basic inequalities based upon the fact that the function $X \ln X$ is a convex function:

$$\text{Lemma 1: } k_i \geq 0, \quad p_i \geq 0, \quad \sum_i p_i = 1$$

$$\Rightarrow \left(\sum_i p_i k_i \right) \ln \left(\sum_i p_i k_i \right) \leq \sum_i p_i k_i \ln k_i$$

(This property is actually the definition of a convex function, but follows from the fact that $\frac{d^2}{dx^2}(X \ln X) = \frac{1}{x} > 0$ (for $x > 0$) which is the elementary notion of convexity.) There is an immediate corollary for the continuous case:

Corollary 1

$$f(x) \geq 0, \quad P(x) \geq 0, \quad \int P(x) dx = 1$$

$$\Rightarrow \left(\int P(x) f(x) dx \right) \ln \left(\int P(x) f(x) dx \right) \leq \int P(x) f(x) \ln f(x) dx$$

(3)

We now derive a more general and very useful inequality from

Lemma 1:

Lemma 2: $a_i \geq 0, b_i \geq 0 \quad (\forall i)$

$$\Rightarrow \left(\sum_i b_i \right) \ln \left(\frac{\sum_i b_i}{\sum_i a_i} \right) \leq \sum_i b_i \ln \frac{b_i}{a_i}$$

Proof: Let $p_i = \frac{a_i}{\sum_i a_i}$, so that $p_i \geq 0$ and $\sum_i p_i = 1$.

Then, by lemma 1:

$$(2.1) \quad \left[\sum_i p_i \left(\frac{b_i}{a_i} \right) \right] \ln \left[\sum_i p_i \left(\frac{b_i}{a_i} \right) \right] \leq \sum_i p_i \left(\frac{b_i}{a_i} \right) \ln \left(\frac{b_i}{a_i} \right)$$

substituting for p_i :

$$(2.2) \quad \left(\sum_i \frac{a_i}{(\sum a_i)} \frac{b_i}{a_i} \right) \ln \left(\sum_i \frac{a_i}{(\sum a_i)} \frac{b_i}{a_i} \right) \leq \sum_i \frac{a_i}{(\sum a_i)} \frac{b_i}{a_i} \ln \frac{b_i}{a_i}$$

so that

$$(2.3) \quad \left(\sum_i b_i \right) \ln \left(\frac{\sum_i b_i}{\sum_i a_i} \right) \leq \sum_i b_i \ln \frac{b_i}{a_i} \quad \underline{\text{QED.}}$$

We also mention the analogous continuous case:

Corr 2: $f(x) \geq 0, g(x) \geq 0$

$$\Rightarrow \left[\int f(x) dx \right] \ln \frac{\left[\int g(x) dx \right]}{\left[\int f(x) dx \right]} \leq \int f(x) \ln \frac{f(x)}{g(x)} dx$$

93 We now prove the theorems (II-) concerning the behavior of correlation and relative information under refinement. We suppose that the original (unrefined) distribution is $P_{ij...K} = P(x_i)$, and that the refined distribution is $\tilde{P}_{ij...K}^{u_i v_j ... v_K}$, so that the value x_i has been resolved into a number of values $x_i^{v_i}$, and similarly for y_j, z_k .

$$(3.1) \quad P_{ij...K} = \sum_{u_i v_j ... v_K} \tilde{P}_{ij...K}^{u_i v_j ... v_K} \quad P_i = \sum_{u_i} \tilde{P}_i^{u_i}, \text{ etc.}$$

so that

$$(3.2) \quad \{X, Y, \dots, Z\}' = \sum_{ij...K} \sum_{u_i v_j ... v_K} \tilde{P}_{ij...K}^{u_i v_j ... v_K} \ln \frac{\tilde{P}_{ij...K}^{u_i v_j ... v_K}}{\tilde{P}_i^{u_i} \tilde{P}_j^{v_j} \dots \tilde{P}_K^{v_K}}$$

but by Lemma 2:

$$(3.3) \quad \left(\sum_{u_i v_j ... v_K} \tilde{P}_{ij...K}^{u_i v_j ... v_K} \right) \ln \left(\frac{\sum_{u_i v_j ... v_K} \tilde{P}_{ij...K}^{u_i v_j ... v_K}}{\sum_{u_i v_j ... v_K} \tilde{P}_i^{u_i} \tilde{P}_j^{v_j} \dots \tilde{P}_K^{v_K}} \right) \leq \sum_{u_i v_j ... v_K} \tilde{P}_{ij...K}^{u_i v_j ... v_K} \ln \frac{\tilde{P}_{ij...K}^{u_i v_j ... v_K}}{\tilde{P}_i^{u_i} \tilde{P}_j^{v_j} \dots \tilde{P}_K^{v_K}}$$

and substitution of (3.3) into (3.2), noting that $\sum_{u_i v_j ... v_K} \tilde{P}_{ij...K}^{u_i v_j ... v_K} = (\sum_{u_i} \tilde{P}_i^{u_i})(\sum_{v_j} \tilde{P}_j^{v_j}) \dots (\sum_{v_K} \tilde{P}_K^{v_K})$ implies:

$$(3.4) \quad \{X, Y, \dots, Z\}' \geq \sum_{ij...K} \left(\sum_{u_i v_j ... v_K} \tilde{P}_{ij...K}^{u_i v_j ... v_K} \right) \ln \left(\frac{\sum_{u_i v_j ... v_K} \tilde{P}_{ij...K}^{u_i v_j ... v_K}}{(\sum_{u_i} \tilde{P}_i^{u_i})(\sum_{v_j} \tilde{P}_j^{v_j}) \dots (\sum_{v_K} \tilde{P}_K^{v_K})} \right)$$

$$= \sum_{ij...K} P_{ij...K} \ln \frac{P_{ij...K}}{P_i P_j \dots P_K} = \{X, Y, \dots, Z\}$$

and refinement never decreases a correlation. QE

Let us now consider the effects upon relative information:

Suppose that $\hat{a}_i^{u_i}, \hat{b}_j^{v_j}, \dots, \hat{c}_k^{w_k}$ are the basic measures for $\hat{P}_{ij\cdots k}^{u_i v_j \cdots w_k}$. Then the measures for $P_{ij\cdots k}$ are

$$a_i = \sum_u \hat{a}_i^{u_i}, b_j = \sum_v \hat{b}_j^{v_j}, \dots, \text{and the relative information:}$$

$$(3.5) \quad I'_{XY\cdots Z} = \sum_{i,j,\dots,k} \sum_{u_i v_j \cdots w_k} P_{ij\cdots k}^{\hat{u}_i \hat{v}_j \cdots \hat{w}_k} \ln \frac{P_{ij\cdots k}^{\hat{u}_i \hat{v}_j \cdots \hat{w}_k}}{\hat{a}_i^{u_i} \hat{b}_j^{v_j} \cdots \hat{c}_k^{w_k}}$$

and by exactly the same argument as before (Lemma 2)

$$(3.6) \quad I'_{XY\cdots Z} \geq \sum_{i,j,\dots,k} P_{ij\cdots k} \ln \frac{P_{ij\cdots k}}{\hat{a}_i^{u_i} \hat{b}_j^{v_j} \cdots \hat{c}_k^{w_k}} = I_{XY\cdots Z}$$

and refinement never decreases relative information.

Just for fun:

General Stochastic Process: (continuous)

$$P(x, s+t) = \int P(x', s) A(x', x, t) dx' \quad \text{non-negative}$$

and $\int A(x', x, t) dx'$

$$\text{where } A(x', x, s+t) = \int A(x', x'', s) A(x'', x, t) dx''$$

$$\text{Stationary} \Rightarrow P^*(x, -t) = P^*(x, t+5) \text{ all } s$$

$$\Rightarrow \text{ie } P^*(x)$$

$$\text{so that } \int P^*(x') A(x', x, t) dx' = P^*(x) \quad (\text{all } x) \quad (\text{all } t)$$

$$\text{now, } I^t = \int P(x, t) \ln \frac{P(x, t)}{P^*(x)} dx$$

$$I^{t+s} = \int \left[\int P(x', t) A(x', x, s) dx' \right] \ln \frac{\left[\int P(x', t) A(x', x, s) dx' \right]}{\left[\int P^*(x') A(x', x, s) dx' \right]}$$

$\uparrow f(x')$ $\uparrow g(x')$

by Lemma 2, cont case

$$\begin{aligned} &\leq \int \left[\int P(x', t) A(x', x, s) \ln \frac{P(x', t) A(x', x, s)}{P^*(x') A(x', x, s)} dx' \right] dx \\ &= \int \left[\underbrace{\int A(x', x, s) dx'}_{1} \right] P(x', t) \ln \frac{P(x', t)}{P^*(x')} dx' \\ &= I^t \quad \text{and} \quad I^{t+s} \leq I^t \quad \underline{\text{QED}} \end{aligned}$$