COWLES COMMISSION DISCUSSION PAPER: ECONOMICS NO. 2060

A Theory for the Value of a Message and Systems for Codifying Possible Messages[1]

Leo Törnqvist

December 11, 1952

I Notations:

(1) x = the "exact" world history until t. $x \in \mathfrak{X}$.

(2) λ = a system for stratification of $\mathfrak{X}$ in n strata.

(3) $X_{\lambda i}$ = stratum i in system λ for messages about x.

(4) $x \in X_{\lambda i}$ = the message (λ, i).

(5) $a \in A_{\lambda i}$ = possible decisions after (λ, i) is received.

(6) $\hat{a}_{\lambda o} = 0$, the set of best decisions when no message = (λ, o) is received.

(7) $\hat{a}_{\lambda i}$ = the set of optimal decisions $\hat{a}_{\lambda i} \in \hat{a}_{\lambda i}$ if (λ, i) is received.

(8) $F_{\lambda i}(\cdot)$ = the conditional probability distribution for x, if (λ, i) is received.

(9) $U_{\lambda i}(a)$ = value of decision a if (λ, i) is received.

(10) $\hat{w}_{\lambda i} = \max_a [U_{\lambda i}(a) - U_{\lambda i}(o)] = U_{\lambda i}(\hat{a}_{\lambda i}) - U_{\lambda i}(o) = \max_a w_{\lambda i}(a)$ = value of the message (λ, i).

(11) $P_{\lambda i}$ = probability of receiving the message (λ, i).

(12) $\hat{w}_\lambda = \sum_{i=1}^{n_\lambda} P_{\lambda i} \hat{w}_{\lambda i}$ = value of the codifying system λ.

(13) $w(x,a) = U(x,a) - U(x,o)$ = value of decision a if x is true.

---

1. Research undertaken by the Cowles Commission for research in economics under contract Nonr-358(01), NR 047-006 with the Office of Naval Research.

Theory. Consider a as a measurable variable and put tentatively

$$w(x,a) = Q(x,a) \cdot [a_x^2 - (a_x - a)^2] = Q(x,a)(2a_x a - a^2)$$

$$= U(x,a) - U(x,o); \quad \frac{\partial Q(x,a)}{\partial a} = 0 \text{ for } a = a_x.$$

where a_x denote the variable value a corresponding to the set of optimal decisions given x. Thus

$w(x,a_x) = Q(x,a_x)\, a_x^2$ and $a_x^2 = \frac{w(x,a_x)}{Q(x,a_x)}$. By a suitable choice of the scaling and origin for the space of decisions A, it is likely that we can make $Q(x,a)$ approximately independent of a.

The value $w_{\lambda i}(a)$ is a mean value of $w(x,a)$ calculated according to the formula

$$w_{\lambda i}(a) = \int w(x,a)\, dF_{\lambda i}(x) = \int Q(x,a)(2a_x a - a^2) dF_{\lambda i}(x)$$

$$= 2a \int Q(x,a)\, a_x dF_{\lambda i}(x) - a^2 \int Q(x,a)\, dF_{\lambda i}(x).$$

$$= Q_{\lambda i}(a)(2\alpha_{\lambda i}(a)a - a^2) = Q_{\lambda i}(a)(\alpha_{\lambda i}(a)^2 - (\alpha_{\lambda i}(a) - a)^2)$$

Here $Q_{\lambda i}(a)$ denote the mean value of $Q(x,a)$ over the stratum i in system λ and

$$\alpha_{\lambda i}(a) = \frac{\int a_x\, Q(x,a)\, dF_{\lambda i}(x)}{\int Q(x,a)\, dF_{\lambda i}(x)} = \int a_x\, dF_{\lambda ia}(x)$$

a weighted mean value of a_x calculated by mean of the weights

$\frac{Q(x,a)}{Q_{\lambda i}(a)}\, dF_{\lambda i}(x) = dF_{\lambda ia}(x)$. The value $w_{\lambda i}(a)$ is always smaller or equal to

$$w_{\lambda i}(a) \leq Q_{\lambda i}(a)(\alpha_{\lambda i}(a))^2$$

This is also the case for that $a = \hat{a}_{\lambda i}$ which give the result

$$\hat{w}_{\lambda i} = w_{\lambda i}(\hat{a}_{\lambda i}) = \max w_{\lambda i}(a)$$

$$\hat{w}_{\lambda i} \leq Q_{\lambda i}(\hat{a}_{\lambda i})(\alpha_{\lambda i}(\hat{a}_{\lambda i}))^2 = \hat{Q}_{\lambda i} \cdot \hat{\alpha}_{\lambda}^2$$

The maximum value $\hat{w}_{\lambda i}$ of $w_{\lambda i}(a)$ is larger or equal to

$$w_{\lambda i} = Q_{\lambda i} \cdot a^2_{\lambda i}$$

where

$$a_{\lambda i} = \alpha_{\lambda i}(a_{\lambda i})$$

and

$$Q_{\lambda i} = Q_{\lambda i}(a_{\lambda i}).$$

Thus

$$\underline{\text{II.1.}}\quad Q_{\lambda i}a^2_{\lambda i} \leq \hat{w}_{\lambda i} = \hat{Q}_{\lambda i}(\hat{\alpha}^2_{\lambda i} - (\alpha_{\lambda i} - \hat{a}_{xi})^2) \leqq Q_{\lambda i}\hat{\alpha}^2_{\lambda i}.$$

If as assumed the scaling and origo for the space A has been chosen in a suitable way the function $Q(x,a)$ is always positive and approximately independent of a. In this case $Q_{\lambda i} \approx \hat{Q}_{\lambda i}$, and $\hat{\alpha}_{\lambda i} \approx a_{\lambda i}$

The value of a message is thus

$$\underline{\text{II.2}}\quad \hat{w}_{\lambda i} \approx Q_{\lambda i}a^2_{\lambda i}$$

and the value of the codifying system λ becomes

$$\underline{\text{II.3}}\quad \hat{w}_\lambda = \sum_{i=1}^{n_\lambda} P_{\lambda i}\hat{w}_{\lambda i} \approx \sum_{i=1}^{n_\lambda} P_{\lambda i}Q_{\lambda i}a^2_{\lambda i} = \sum_{i=1}^{n_\lambda} P_{\lambda i a'_{\lambda i}} a^2_{\lambda i}.$$

We can now state the main results of our theory of the value of a system for codifying messages in the form:

<u>III.</u> The value of a system (λ) for codifying messages can be considered as an <u>inter class variance</u> w_λ for the variable $a_{\lambda i}$.

$$a_x = \sqrt{\frac{w(x,a_x)}{Q(x,a)}}, \quad \hat{a}_{\lambda i} \approx a_{\lambda i} \approx \int a_x \, dF_{\lambda i a_{\lambda i}}(x,$$

according to the theory of stratification we can continue: The inter-class variance w_λ is a maximum if the corresponding intraclass variance

$$\sigma_{\lambda}^{2} = \sum_{i=1}^{n} p_{\lambda i a_{\lambda i}} \sigma_{\lambda i}^{2} \qquad \sigma_{\lambda i}^{2} = \int (a_x - a_{\lambda i})^2 dF_{\lambda : a_{\lambda i}}(x)$$

is a minimum. This result can be utilized for instance for the statement:

If the optimal decisions $a_{\lambda ij}$ are the same for all substrata j to a stratum, i in the system λ, then no increase in the value of the codification system λ can be obtained by the use of substratifications of the stratum (λ, i) into more precise substrata. The number n_{λ} needs thus not be larger than the number of different possible decisions to exhaust the whole value of a precise knowledge of the position of x in the set $\mathfrak{X}$.

<u>Critical Comment:</u> The value of a message can be positive (or even negative) even if it does not change any decisions compared with the decisions if it is not received if the message in itself is pleasurable or disappointing, satisfies our human curiosity, or is valuable because of other reasons than those of influencing decision making. The total value $v_{\lambda i}$ of a message (λi) is in fact a sum of the value

$$w_{\lambda i}(o) = \int [u(x,o) - u(\mathfrak{X},o)] \, dF_{\lambda i}(x)$$

of the messages in itself, independent of its value for decision making and its "decision making" value $\hat{w}_{\lambda i}$. Thus

$$v_{\lambda i} = w_{\lambda i}(o) + \hat{w}_{\lambda i}.$$

The <u>net</u> money value of the messages (λ, i) is equal to a money gain $g(v_{\lambda i})$ giving the increase in utility equal to $v_{\lambda i}$ and the money equivalent c_{xi} of the costs for getting the message.

$$\acute{g}_{\lambda i} = g(v_{\lambda i}) - c_{\lambda i}.$$

By deciding if it pays or not to buy the message (λ, i) we have to evaluate the ex ante probability distribution of $\acute{g}_{\lambda i}$. If the utility value of this distribution of gains is positive it is "profitable" to make the decision to buy the information (λ, i).