情報와 意思決定
(Informations and Decision Making)

金 載 周*

0. 序 論

不確實性下에서 行動決定을 하지 않으면 안되는 상황에서 情報란 무엇이며 情報量과 情報價値는 意思決定에 어떤 역할을 하는가 情報와 主觀的인 確率에 기초한 Bayes 決定基準이 傳統的인 意思決定基準에 비하여 우위성 등을 논하고자 함이 本論文의 主된 目的이다. 제1.2절에서는 결정문제의 구조, 不確實性과 그 測度에 관하여 언급하고 제3절에서는 전통적인 결정기준과 그 모순점을 예시하고 Bayes 決定基準이란 무엇인가를 소개하여 제4절에서는 Bayes 통계학에 기초한 情報量, 情報價値의 意思決定과의 관계 등을 살펴보고자 한다.

1. 決定問題의 構造

우리들은 개인적 생활에 있어서나 혹은 직장의 책임자로서 끊임없이 의사결정을 촉구당하는 경우가 많다. 예를 들어 어떤 사업에 投資를 할 것인가, 안할 것인가, 투자를 한다면 얼마정도의 金額을 투자할 것인가 라든가 혹은 어떤 C_1회사의 판매부장의 입장에서 그회사의 제품 C_1에 대하여 강력한 판매촉진 캠페인을 실시할 것인가 안할 것인가에 대한 결단을 촉구당하고 있다고 할때 좋은 決定이란 무엇이며 또한 그것을 어떤 방법으로 구할 것인가라는 문제에 해답을 주는 것이 (意思)決定理論이다.

이와같은 결정문제에 공통된 특질을 추상하여 보면 다음과 같다.

(1) 어떤것이나 不確實性에 직면한 결정문제다.

(2) 決定者는 자기에게 유리한 결과가 얻어지도록 결정을 하고싶다.

(3) 最終決定을 내리기 전에 情報를 입수하여 不確實性을 해소하고 싶다.

위의 C_1회사의 예에서 만일 이회사의 주경쟁회사인 C_2회사가 C_1제품과 동종류인 C_2제품을 시장에 내놓을 가능성이 있다고 한다면 이때 C_1회사의 판매부장은 C_1제품에 대한 판매촉진 캠페인을 실시하여 C_2제품이 시장에 주는 충격을 약화시키거나 만일 C_2제품이 발매되지 않는다면 C_1제품은 잘팔리고 있으므로 캠페인을 실시하지 않는다는 어느 行動을 취하지 않으면 안된다. 이제 위의 예를 가지고 결정문제의 구조를 좀더 자세히 검토해 보기로 하자.

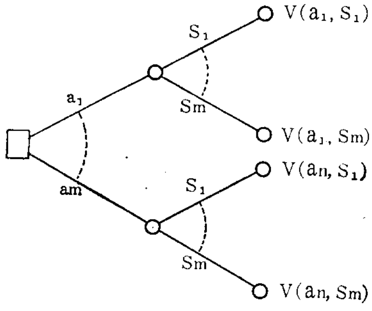
(一) 決定者가 최종적으로 정하지 않으면 안되는 것은 판매촉진 캠페인을 실시할 것인가 안할 것인가의 어느것이다. 이것을 각각 行動(action)이라고 부르고 a_1, a_2로 나타내기로 한다. 일반적인 決定問題에 있어서 行動전체가 $a_1, a_2, \cdots, a_n$ 일때 이들 전체의 집합을 A로 나타내고 이것을 行動空間(action space)라고 부른다.

(二) 行動 a_1, a_2에 기인한 결과는 C_2제품이 시장에 발매될 것인가 아닌가에 의존되며 이것은 決定者가 制御할 수 없는 要因으로 行動의 結果를 규정한다. 이와같이 決定者의 制御外에 있는 것을 명확히 규정한 命題를 自然狀態(state of nature)라 부른다. 일반적으로 일어날 수 있는 모든 자연상태를 $s_1, s_2, \cdots, s_m$라 하고 그 集合을 S라 할 때 이 S를 狀態空間(state space)라 부른다. 여기서 決定者는 眞의 자연상태는 S가운데 어느 要素일 것이다는 알고 있지만 어느 것이다는 確實히 알 수 없다. 不確實性에 직면한 決定問題라는 것은 바로 이와같은 意味에서 나온 것이다.

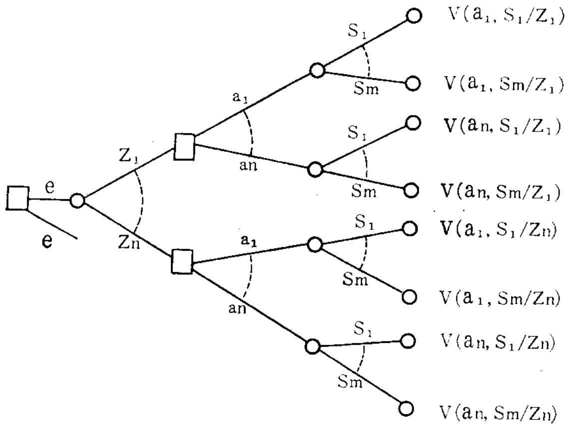
(三) 行動 $a \in A$를 취했을 때 眞의 자연상태가 $s \in S$일 때 일어나는 결과를 $V(a,s)$라고 할 때 $V(a,s)$는 金額으로 나타내지는 경우도 있고 그렇지 않는 경우도 있다.

(四) 최종결정으로 내리는 行動을 취하기전에 自然狀態 S에 관한 情報를 입수한다. 즉 C_1회사의 판매부장은 어떤 情報源(source of message)으로부터 C_2회사가 C_2제품을 발매할 것인가 발매치 않을 것인

*서울大學校 自然大學 계산통계학과.



(a) 情報를 쓰지 않는 경우



(b) 情報를 쓰는 경우

가에 대한 情報를 세일즈맨의 말, 業界의 뉴스 또는 C_2회사의 社員으로 부터 구할 수 있다. 이것을 기호로 $Z_1, \cdots, Z_N$로 나타내기로 한다. 現實의 決定問題는 복잡한 구조를 가지고 있으므로 樹形圖(decision tree)로 나타내고 이것을 통하여 좋은 결정을 쉽게 구할 수 있지만 數學的인 번잡성은 피할 수 없다. 情報를 이용하는 경우와 이용하지 않는 경우의 樹形圖를 그려보면 다음과 같다. 단, 여기서 e는 情報를 나타내는 기호이다.

2. 不確實性에 관하여

不確實性을 연구대상으로하는 數學의 한 분과인 確率論은 확립된 이론으로 널리 인정받고 있음은 주지의 사실이다. 그러나 현실현상과의 대응에 있어서 확율을 어떻게 해석할 것인가에 관한 문제로 두개의 立場이 대립되고 있다. 즉 베이지안(Baysian)立場과 非베이지안(Non-Baysian)立場이다. 간단히 말해서 前者는 主觀的인 確率 [2] [8] (subjective probability)에 기초를 두고 있고 後者는 客觀的인 確率(objective probability)에 기초를 두고 있다.

(a) 客觀的確率

어떤 事象(event)이 일어나는 확율의 시행을 반복했을 때 그 事象이 일어나는 相對度數의 極限으로 해석하는 立場을 客觀的確率의 立場이라 한다. 이 確率概念은 중요하고 적용범위도 넓은 것은 사실이지만 앞으로 5년이내 세계 3차대전이 일어날 확율이라든가, 남산에 금이 10톤이상 매장되어 있을 확율이라든가 등은 客觀的 確率로 언급한다는 것은 허락되지 않는다.

(b) 主觀的인 確率

어떤 事象이 일어나는 確率을 決定者가 판단하여 "확실한 정도"를 양적으로 표현한 것으로 해석하는 立場을 主觀的 確率의 立場 혹은 判斷確率的立場이라 부른다. 따라서 主觀的確率은 決定者가 가지고 있는 過去의 經驗·知識등에 의하여 정하여지는 것이므로 이것은 決定者에 따라 변한다. 그러나 이것은 客觀的確率이 적용되지 않는 事象 및 命題에 대해서도 主觀的確率을 적용하여 不確實性하의 決定問題를 보다 合理的으로 해결하자는 것이 최근의 결정이론이며 이것을 베이스論的 決定理論이라 불리워진다. 主觀的確率의 구성 및 측정에 관하여는 Degroot[2]와 Raiffa[8]를 참조하여 주기 바란다. 그러면 이 베이스決定理論이 現實適用性이 높은 理論이면서 실제 企業에 적용한 例가 極히 드문 理由는 무엇인가에 대하여 살펴보기로 하자

첫째로는 計算이 복잡하다는 것이다. 즉 行動 및 狀態의 수가 증가하고 거기에 動態的 要因이 첨가하면 그 복잡성은 가속도로 증가한다.

둘째의 이유는 事前確率(prior probability)를 주관적으로 판단한다고 했지만 그것을 정하는 데 곤란점이 있기 때문이다.

셋째로는 이 理論 자체가 아직 젊은 理論으로 지금부터 개발하여 가지 않으면 안되는 면이 너무 많기 때문이다.

(c) 不確實性의 測度

$(S,\mathscr{S})$를 可測인 狀態空間 ξ를 $(S,\mathscr{S})$ 위에 정의된 事前確率測度 ξ가 $(S,\mathscr{S})$ 위의 임의의 측도λ에 관하여 절대연속이라 가정한다.

(i) 未知의 상태 $s \in S$에 관한 不確實性의 測度 $H(\xi)$를 Shannon[7] Lindley[6]등은 다음과 같이 정의했다.

$$H(\xi) = -\int \xi(s) \log \xi(s) \ d\lambda$$

특히 $S = \{s_1, \cdots, s_m\}$ 일때는

$$H(\xi) = -\sum_{i=1}^{m} \xi(s_i) \log \xi(s_i)$$

로 된다.

(ii) $(A, \mathscr{A})$를 可測인 行動空間 $W(\cdot, \cdot)$를 $(S \times A, \mathscr{S} \times \mathscr{A})$ 위에 정의된 可測인 損失函數라하자. 그때 $(S, \mathscr{S})$, $\xi, (A, \mathscr{A})$, $W(\cdot, \cdot)$는 基本決定問題 D_0을 규정한다고 말하고 $D_0 \equiv \{S, \xi, A, W\}$로 나타내다 기본결정문제$D_0$에 직면한 不確實性의 측도로 다음과 같은 量을 정의한다.

$$R_0(D_0) \equiv R_0(\xi | W) \equiv \inf_{a \in A} \int W(s,a) d\xi(s).$$

여기서 $R_0(D_0)$는 Bayes 危險을 나타낸다. 기본결정문제에서 베이스 위험에 해당하는 결정을 택하면 그것이 最適決定(optimal decision)이 된다. 이것은 또한 만일 $W(\cdot, \cdot)$ 대신 效用函數 $U(\cdot, \cdot)$가 들어 갔다면 기대효용을 最大로 하는 결정과 대등한 관계가 있다.

3. 여러가지 決定基準

決定者에 있어서 주된 문제는 주어진 行動空間 A 가운데 어떤 要素를 택할 것인가 이다. 이제 S를 未知의 狀態空間 $V(a,s)$를 $A \times S$에 정의된 利得函數라고 할 때 이때 $V(a,s)$는 만일 A, S,가 有限集合: $A = \{a_1, \cdots, a_m\}$, $S = \{s_1, \cdots, s_n\}$이면 $V(a,s) = V_{ij}$가 되고 다음과 같은 利得行列을 얻게된다.

$$(V_{ij}) = \begin{pmatrix} V_{11} \cdots\cdots V_{1n} \\ \vdots \qquad\quad \vdots \\ V_{m1} \cdots\cdots V_{mn} \end{pmatrix}$$

만일 $a^* \in A$가 존재하여 모든 $s \in S$와 모든 $a' \in A$에 대하여 $V(a^*, s) \geq V(a', s)$이면 a^*를 택하면 a^*가 최적결정이 된다.

그러나 실제문제에서는 이러한 a^*가 존재하지 않는 것이 보통이므로 $a \in A$에 대하여 실수치 $V(a)$를 만들고 $\operatorname*{Max}_{a \in A} V(a) = V(a^*)$인 평가기준을 만들 수 있다.

(1) Laplace 決定基準

각 행동 $a \in A$를 실수 $\frac{1}{n} \sum_{j=1}^{n} V(a, s_j)$로 평가하고자 하는 기준이다. 즉 각 행동 a를 실수 $V_L(a) = \sum_{j=1}^{n} V(a, s_j)$ 로 평가 하는 것과 同値關係이다.

a_L^*를 $\operatorname*{Max}_{a \in A} V_L(a) = V(a_L^*)$를 만족하는 것이라면 a_L^*가 최적행동이다.

(2) A. Wald의 基準

실수치 $V(a)$를 $V_w(a) = \operatorname*{Min}_{s \in S} V(a,s)$로 정의하고 $\operatorname*{Max}_{a \in A} V_w(a) = V_w(a_w^*)$된 a_w^*를 최적결정으로 택하는 기준을 말한다. 이것을 Maxmin 利得 결정기준 이라고도 한다.

(3) Hurwicz의 基準

실함수 $V_H(a)$를 $V_H(a) = \alpha [\operatorname*{Max}_{s \in S} V(a,s)] + (1-\alpha) [\operatorname*{Min}_{s \in S} V(a,s)]$로 정의하고 $V_H(a_H^*) = \operatorname*{Max}_{a \in A} V_H(a)$인 a_H^*를 최적결정으로 택하는 기준을 말한다. 단, 여기서 α는 $0 \leqslant \alpha \leqslant 1$인 실수이다.

(4) L.J. Savage의 결정기준

이 기준은 우선 각 행동 $a \in A$에 관하여 $s \in S$가 진의 상태일 때의 손실 $W(a,s)$를

$$W(a,s) = \operatorname*{Max}_{a' \in A} V(a', s) - V(a,s)$$

를 구하여 기회손실표(opportunity loss table)을 만든 다음 각 행동 a를 $V_s(a) = \operatorname*{Max}_{s \in S} W(a,s)$로 특징지우고 이것을 최소로하는 행동 $a_s^* \in A$ 즉 $\operatorname*{Min}_{a \in A} V_s(a) = V_s(a_s^*)$인 a_s^*를 최적결정으로 택하는 기준이다. 이것은 일명 Minmax 손실기준이라고도 부른다.

이제 간단한 예로서 위의 결정기준에 의한 최적결정들을 비교해보자. 상태공간 $S = \{s_1, s_2, s_3, s_4\}$라 하고 행동공간 A를 $A = \{a_1, a_2, a_3, a_4\}$라 했을 때 利得表(payoff table)이 다음과 같이 주어진다고 하자.

利 得 表

A \ S	s_1	s_2	s_3	s_4
a_1	2	3	1	3
a_2	2	2	2	2
a_3	1	5	1	1
a_4	1	4	1	2

이때 Laplace 결정기준에 의하면 a_1이, Wald의 결정기준에 의하면 a_2가, Hurwicz의 결정기준에 의하면 $\alpha = 0$일 때는 a_2, $\alpha > \frac{1}{4}$ 일 때는 a_3, $0 < \alpha < \frac{1}{4}$ 일 때는 a_2가, Savage의 결정기준에 의하면 a_4가, 각각 최적 결정이 됨을 쉽게 알 수 있다. 여기서 어떤 결정기준을 가장 바람직한 것으로 택해야 하는가? 위의 결정기준은 어느거나 만족할 수 없는 것은 우리의 결정은 항상 불확실성에 직면하고 있으며 그 불확실성의 정도를 고려하지 않았기 때문이다.

(5) 期待金額基準

이것은 Bayes 결정기준이라고도 볼 수 있는 것으로 우선 자연상태 $S_1, S_2, \cdots, S_n$에 관하여 판단확율, $P_1, P_2, \cdots, P_n$,을 가지고 다음과 같이 실수치 $M(a)$를 $M(a) = \sum_{i=1}^{n} P_i V(a, s_i)$를 정의하고 이 $M(a)$를 최대로 하는 $a_M^* \in A$ 즉 $\operatorname*{Max}_{a \in A} M(a) = M(a_M^*)$인 a_M^*를 최적결정으로 택하는 기준이다. 윗式에서 $V(a, s_j)$대신에 效用函數

$U(a,s_j)$가 들어갔을 때를 期待効用基準이라고도 한다. 이것은 情報가 없을 때의 Bayes 決定에 속하는 것이다.

4. 情報·情報量 및 情報價値

(a) 情報란?

未知의 自然狀態 S에 관한 情報$e(\tilde{z})$라는 것은 $\tilde{Z}$가 可測空間$(Z,\mathscr{Z})$ 위에 정의된 확율변수로 $s\in S$가 주어졌을 때 $(Z,\mathscr{Z})$ 위의 측도 μ에 관한 확율밀도 $f(z|s)$가 알려질 수 있다고 가정한다. 그때 만일 決定者가 $\tilde{Z}$의 실현치 z를 알 수 있으면 $s\in S$에 대한 정보 $e(\tilde{z})$가 그에게 얻어질 수 있다고하고 $\tilde{Z}$의 실현치 z를 情報$e(\tilde{z})$의 通報라고 부른다.

만일 $S=\{s_1,\cdots, s_m\}$, $\tilde{Z}=(z_1, \cdots, z_n)$라 두면 情報$e(\tilde{z})$는 Markov 行列 $\Lambda\equiv\|\lambda_{ij}\|$, $i=1, \cdots, m; j=1, \cdots, n$에 의하여 규정되고 Λ를 情報行列이라고도 한다. 여기서

$$\lambda_{ij}=P_r\{\tilde{z}=z_j|s=s_i\}=P(z_j|s_i)$$

이다. $s\in S$에 대한 정보 $e(\tilde{z})$는 ξ와 손실 함수 $W(\cdot,\cdot)$와 독립으로 주어짐은 말할 필요도 없다. 이제 基本決定問題 $D_0=\{S,\xi,A,W\}$를 해결하기 위하여 情報 $e(\tilde{z})$가 이용된다고 가정하고 이것이 어떻게 쓰이는가 검토하여 보자.

이제 決定者가 情報$e(\tilde{z})$로 부터의 通報 $\tilde{Z}=Z_j$가 얻어졌다고 하자. 그때 그는 $S=S_i$, $i=1\cdots m$, 가 眞의 상태라는 事前判斷確率 $\xi(S_i)$, $i=1 \cdots m$을 어떻게 수정하여야 하는가 이것을 가르쳐 주는 것이 確率論의 Bayes定理이다. 즉 情報 $e(\tilde{z})$로 부터의 통보 $\tilde{Z}=Z_j$가 얻어진 후의 s_i의 事後確率 $\xi(s_i|z_j)$는 Bayes 定理에 의하여

$$\xi(s_i|z_j)=\xi_i(z_j)=\frac{\xi(s_i)p(z_j|s_i)}{p(z_j)}=\frac{\xi_i\lambda_{ij}}{p(z_j)}$$

로 주어진다. 이때 基本決定問題D_0에 있어서 ξ를 $\xi(z_j)$로 바꾼 決定問題 $D_0(z_j)\equiv\{S,\xi(z_j),A,W(\cdot,\cdot)\}$로 직변하게된다. 이때 기대손실을

$$M(a|\xi(z_j))=\sum_i \xi_i(z_j)W(a,s_i)$$

라 두면 이것을 최소로 하는 a_i^*를 구하면 최적결정이 된다. 손실함수 대신 효용함수$U(a|(z_j))$를 쓰면 이것을 최대로 하는 결정 즉 $\underset{a\in A}{\mathrm{Max}}\ U(a|\xi(z_j))=U(a_i^*|\xi(z_j))$라 두면 a_i^*를 택하는 결정이 최적결정이다. 이때 a_i^*는 情報가 있을 때의 Bayes決定에 해당된다. 여기서 주의할 점은 사전확율을 계속 정보를 입수하여 사후개정하여가면 진의 상태를 알 수 있다는 점이다.[2] 진의 상태를 안다면 필요한 계산을 함으로서 정확한 결정을 내릴 수 있음은 말할 필요도 없다.

(b) 情報量과 情報價値

최후로 情報量과 情報價値에 관하여 간단히 언급하고자 한다. 상태공간 $S=\{s_1, \cdots, s_m\}$ 상의 사전확율법칙이 $\xi=(\xi_1,\cdots, \xi_m)$ 일때 자연상태 S에 관한 不確實性의 측도 $H(\xi)$는 정보 $e(\tilde{z})$로부터 통보 $\tilde{Z}=Z_j$가 얻어진 후의 사후확율법칙을 $\xi(z_j)\equiv(\xi_1(z_j),\cdots,\xi_m(z_j))$라고 할 때 사후확율에 관한 불확실성에 관한 측도, 즉

$$H(\xi(z_j))\equiv-\sum_i\xi_i(z_j)\log\xi_i(z_j)$$

로 변하고 이것의 기대값, 즉

$$E[H(\xi(z_i))]\equiv\sum_j P(z_j)H(\xi(z_j))$$

를 Rényi [5]는 $e(\tilde{z})$를 관측한 후의 損失情報量이라 정의했다 이것을 通信理論에서는 瞬昧度[9]에 해당하는 量이다. 情報量은 DeGroot [1]에 의하면 자연상태 S에 관한 事前分布의 不確實性의 측도에서 情報 $e(\tilde{z})$에 의하여 얻어지는 事後分布의 不確實性의 期待値의 差로서 정의했다. 즉 $(S,\mathscr{S})$ 위에 事前確率測度가 ξ 일때 情報$e(\tilde{z})$가 제공하는 情報量 $I(\tilde{z}|\xi)$는

$$I(\tilde{z}|\xi)\equiv H(\xi)-E[H(\xi(\tilde{z}))]$$

로 된다.

이 量은 ξ에 의존하고 $I(\tilde{Z}|\xi)\geqslant 0$[4] 임은 증명되어 있다. 또한 基本決定問題 D_0을 풀기위한 決定者에게 情報 $e(\tilde{Z})$가 주어지면 그때 그는 決定問題 $D\equiv\{D_0: e(\tilde{z})\}$을 가진다고 말하고 決定問題 D_0에서 $\tilde{Z}=Z_j$가 관측된 후의 위험을 $R_0(\xi(z)|W)$로 나타내고 그것의 期待値:

$$R(\tilde{Z}|\xi,W)\equiv E[R_0(\xi(z)|W)]$$

를 $e(\tilde{z})$를 관측한 후에 일어나는 損失決定量[4]이라 부른다. 이것은 결정문제 D에 있어서 Bayes위험임은 말할 필요도 없다. Marschak [3]는 情報價値를 다음과 같은 量, 즉

$$V(\tilde{Z}|\xi,W)\equiv R({}_0\xi|W)-R(\tilde{Z}|\xi,W)$$

로 定義하고 이量을 基本決定問題 D_0에 대한 $e(\tilde{z})$의 情報價値라 불렀다. 물론 정보가치도 음이 아닌 값을 가지며 이것은 ξ와 W에 의존하는 量이다.[4] 여기서 문제는 만일 自然狀態 S에 관한 두개의 情報 $e(\tilde{x})$와 $e(\tilde{y})$ 가운데(비용은 같다고 가정) 情報量이 큰쪽을 택할 것인가 情報價値가 큰 쪽을 택할 것인가 이다. 결정문제에 관한한 情報價値가 큰 쪽을 택해야 됨은 말할 필요도 없다.

情報量과 情報價値의 比較에서 모든 A,W에 대하여

$V(e(\tilde{x})|\xi)\geqslant V(\tilde{y})|\xi)$이면

$I(\tilde{x}|\xi) \geqq I(\tilde{y}|\xi)$이나 그 역은 성립하지 않는다[4]는 것이 증명되어 있다. 정보량과 정보가치의 비교에서 좀더 구체적인 것은 본인의 소논문[10]을 참조 하기 바란다.

5. 統 論

전통적인 意思決定論에 比하여 Bayes 決定理論은 不確實性下에 意思決定을 하기위한 理論이고 動態的인 問題에 대처할 수 있는 優位性이 있다는 點과 事前·事後分析이 가능하다는 점에서 Bayes 決定理論은 現實適用性이 높은 理論이라 볼 수 있다. 計算의 번잡성은 콤퓨터를 이용함으로서 쉽사리 해결될 수 있을 것 같고 確率의 主觀的인 判斷도 여러사람이 모여 감도분석을 행하고 "客觀的인 主觀"을 사용한다면거의 말썽없는 主觀的確率을 만들 수 있으리라 믿는다. 특히 情報범람時代라고 볼 수있는 現代社會에 Bayes 決定理論은 意思決定理論의 主流로서 계속 발전시켜 현실 적응상의 체계화가 시급하다고 본다.

참 고 문 헌

(1) M.H. DeGroot., *Uncertainty, Information, and Sequential Experiments,* Ann. Math Stat, Vol. 32 (1962).

(2) M.H. DeGroot, *Optimal Statistical Decisions,* McGraw Hill Book Co., (1970).

(3) Jacob Marschak and Koichi Miyasawa, *Economic Comparability of Information Systems,* International Economic Review, Vol. 2 (1968).

(4) Koichi Miyasawa., *Information Value and the Entoropy Formulas,* (To be Presented at the Second World Congress of the Econometric Society in Cambridge, Englnad: September, 1970).

(5) A. Rényi, *On some basic problems of statistics from the point of view of information theory,* Proc. of Fifth Berkeley Symposium on Mathematical Statistics anb Probability, University of California Press (1967).

(6) Lindley, D.V., *On a Measure of the Information Provided by an Experiment,* Ann. Math. Stat. Vol. 27 (1956).

(7) Shannon, C.E., *A Mathematical Theory of Communication,* Bell System Technical Journal, Vol. 14, (1943).

(8) Raiffa., *Decision Analysis,* Addison-Wesley (1968).

(9) Ash, R., *Information Theory,* New York: John Wiley & Sons, Inc., (1965).

(10) 김재주, 情報量과 情報價値의 比較. Bull. Korean Math. Soc. Vol. 8 (1971).