数量化による予測の的中率と相関比との関係について

千　野　貞　子

（1963 年 5 月 受 付）

The Relation between Correlation Ratio and Success Rate in the Classification by the Quantification Method

Sadako CHINO

In the method of quantification procedure that categorizes the given patterns and gives numerical values to them so that they can be treated as several indices, we often make use of the maximization of correlation ratio as follows : Let us consider two groups (sub-populations) indicating the response patterns to certain items of the questionnaire. If the numerical values derived from these response patterns are made to maximize the correlation ratio among these two groups, they will promote the discrimination power.

This numerical value, however, is given by the total sum of indices of several quantified items and is regarded to follow the central limit theorem to some extent. Hence it may be considered that it will be normally distributed. Futhermore, as we treat the problem so as to maximize the value of correlation ratio, it is natural to assume that there is only one cutting point of two distributions.

In this stage we want to get the relation between the correlation ratio and the success rate which indicates the right judgement of discrimination for each unit taken from these groups.

In this paper we obtained the relation between them and illustrated their graphs with regard to various arguments such as means, variances and groups' sizes when the distribution of scale values of each group is distributed in normal distribution.

Institute of Statistical Mathematics

§1. 序

数量化の方法の中には，あらかじめ定められた外的規準（ここでは，分類する際のグループ）を最もよく弁別するような要因を取り上げて，グループ分けの相関比を最大にするよう，要因項目の各カテゴリーに数量を与えていく方法がある．この場合の数量化の効果というものは，客観的には分類の的中率によって表現される．即ち，相関比を最大にすることによって，的中率を上げることを目的としている数量化の方法では，相関比と成功率の関係を適確に掴む事が必要となる．この事についてはすでに max.-min. の方法を用いて，両者の関係が述べられている[1]が，ここでは各グループに属するものの比率を考慮して，両者の関係ずけを試みた．この事は，グループの大きさが，

1）林知己夫：数量化理論の応用例――予測の判断的中率と相関比との関係についての一つの考察と共に，彙報２巻１号．

一方に比べて他方が異なっていて，かつその度合が次の予測（再現）においてそう異なってこないような場合に有用である．

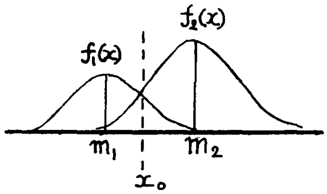
なお，個体の各要因カテゴリー（数量化されている）への反応を，それらの数量の和であらわしていることから，その総合得点は，普通（近似）の意味で，ガウス分布（ある意味での中心極限定理の考え）に従うとしても大過ないものと考えた．また，η-max. の手段をとっていることから，グループの判断的中率を最大にするような分割点を一つとすることは妥当性あることと思われる．具体的にいえば，分割点近辺で分布が単調であり，かつこれ以外に交点を有さない（実際あっても分割点としては無視できる程度）ならば，ガウス分布に近似させることは，妥当性を欠くものではないと考えられる．以上の事から，この仮定の下に話しを進めてみた．

さて，この場合の判断的中率と相関比との関係は如何になるであろうか．われわれはここで両者の関係を簡単に摑むため，図表呈示を目論んだ．ここには，その際の考え方，並びに計算法を述べ，代表的な図表を掲げてみた．

§2. 判断的中率とグループ弁別の分点 x_0 について

われわれは以下外的規準が2つのカテゴリーを持つ場合，即ち2つのグループに関する弁別の問題を考える事にする．今，個体につけられた総合得点が，各グループ内でガウス分布に従うと仮定し，これらの分布の密度函数を $f_1(x)$，$f_2(x)$ とすれば，

$$f_1(x)=\frac{1}{\sqrt{2\pi}\,\sigma_1}e^{-\frac{(x-m_1)^2}{2\sigma_1{}^2}},\quad f_2(x)=\frac{1}{\sqrt{2\pi}\,\sigma_2}e^{-\frac{(x-m_2)^2}{2\sigma_2{}^2}}$$



この時の判断的中率 P は

$$P=\int_{-\infty}^{x_0}p_1f_1(x)dx+\int_{x_0}^{\infty}p_2f_2(x)dx \qquad (1)$$

ただし，

$$\begin{cases} m_i\text{: 各グループの平均，} i=1,\ 2\text{；} m_1<m_2 \text{ とする．} \\ \sigma_i\text{: 〃 〃 標準偏差} \\ p_i\text{: 各グループに属するものの比率} \end{cases}$$

この判断的中率を最大にするような分点 x_0 を $\dfrac{dP}{dx_0}=0$ より求めると

$$2\log\frac{p_1}{p_2}\frac{\sigma_2}{\sigma_1}=\frac{(x_0-m_1)^2}{\sigma_1^2}-\frac{(x_0-m_2)^2}{\sigma_2^2}=\left\{\left(\frac{1}{\sigma_1}+\frac{1}{\sigma_2}\right)x_0-\left(\frac{m_1}{\sigma_1}+\frac{m_2}{\sigma_2}\right)\right\}$$
$$\times\left\{\left(\frac{1}{\sigma_1}-\frac{1}{\sigma_2}\right)x_0-\left(\frac{m_1}{\sigma_1}-\frac{m_2}{\sigma_2}\right)\right\} \qquad (2)$$

いま $\dfrac{p_1}{p_2}=c$……各グループに属すものの比率の比

$\dfrac{\sigma_1}{\sigma_2}=k$……標準偏差の比

$\dfrac{m_1}{m_2}=a$……平均の比

とおくと

$$2\log\frac{c}{k}=\frac{1-k^2}{\sigma_1^2}x_0^2-2(a-k^2)\frac{m_2}{\sigma_1^2}x_0+\frac{m_2^2}{\sigma_1^2}(a^2-k^2) \qquad (3)$$

となる．これより x_0 が求まるが，x_0 を原点にとると

$$2\log\frac{c}{k}=\frac{\overline{X}_1^2}{\sigma_1^2}\left(1-\frac{k^2}{A^2}\right)=\frac{\overline{X}_2^2}{\sigma_2^2}\left(\frac{A^2}{k^2}-1\right)$$

ただし
$$\bar{X}_1=m_1-x_0$$
$$\bar{X}_2=m_2-x_0$$
$$A=\frac{m_1-x_0}{m_2-x_0} \tag{4}$$

よって

$$\frac{\bar{X}_1}{\sigma_1}=A\sqrt{2\log\frac{c}{k}\cdot\frac{1}{A^2-k^2}}=\frac{A}{k}\frac{\bar{X}_2}{\sigma_2} \tag{5}$$
$$\frac{\bar{X}_2}{\sigma_2}=k\sqrt{2\log\frac{c}{k}\cdot\frac{1}{A^2-k^2}} \tag{6}$$

(1) 式の各項において

$$t=\frac{x-m_i}{\sigma_i} \quad (i=1,\ 2)$$

とおくと

$$P=p_1\int_{-\infty}^{x_0}f_1(x)dx+p_2\int_{x_0}^{\infty}f_2(x)dx=p_1\int_{-\infty}^{\frac{x_0-m_1}{\sigma_1}}\varphi(t)dt+p_2\int_{\frac{x_0-m_2}{\sigma_2}}^{\infty}\varphi(t)dt$$
$$=p_1\int_{-\infty}^{-\frac{\bar{X}_1}{\sigma_1}}\varphi(t)dt+p_2\int_{-\frac{\bar{X}_2}{\sigma_2}}^{\infty}\varphi(t)\,dt$$

$p_1+p_2=1$ より $p_2=\dfrac{1}{c+1}$ をつかって

$$P=\frac{c}{c+1}\int_{-\infty}^{-\frac{\bar{X}_1}{\sigma_1}}\varphi(t)dt+\frac{1}{c+1}\int_{-\frac{\bar{X}_2}{\sigma_2}}^{\infty}\varphi(t)dt \tag{7}$$

と書ける．ここに $\dfrac{\bar{X}_1}{\sigma_1}$, $\dfrac{\bar{X}_2}{\sigma_2}$ は上式 (5), (6) によって与えられたものである．

§3. 相関比 η による的中率の表現

ここで η について考えてみよう．

$$\eta^2=\frac{p_1p_2(m_1-m_2)^2}{p_1p_2(m_1-m_2)^2+p_1\sigma_1^2+p_2\sigma_2^2}=\frac{1}{1+\dfrac{(c+1)(ck^2+1)}{c(A-1)^2}\dfrac{\sigma_2^2}{\bar{X}_2^2}} \tag{8}$$

$\dfrac{\sigma_2^2}{\bar{X}_2^2}$ に (6) 式を代入すると

$$\eta^2=\frac{1}{1+\gamma\dfrac{(A^2-k^2)}{(A-1)^2}} \tag{9}$$

ただし

$$\gamma=\frac{(c+1)(ck^2+1)}{2ck^2\log\dfrac{c}{k}}$$

この η は一般には 0 から 1 までの任意の実数値をとる．細かいことは後章に譲るとして，まず i) $\eta\neq0$, $k\neq c$ ii) $\eta\neq0$, $k=c$ iii) $\eta=0$, $k\neq c$ iv) $\eta=0$, $k=c$ にわけて，相関比と的中率の関係を眺めてみよう．

i) $\eta\neq0$, $k\neq c$ の場合

(9) 式を変形して $\dfrac{1}{\gamma}\left(\dfrac{1}{\eta^2}-1\right)=\alpha$ とおき，A について解くと

$$A_2(1-\alpha)+2\alpha A-(\alpha+k^2)=0 \tag{10}$$

$$A=\frac{-\alpha\pm\sqrt{\alpha+k^2(1-\alpha)}}{1-\alpha} \tag{11}$$

$\bar{X}_1/\sigma_1$, $\bar{X}_2/\sigma_2$ の計算には，この A（1根を後述により選択）を (5) (6) 式に用いる．

ii) $\eta\neq0$, $k=c$ の場合

この場合は (2) 式は $\frac{1}{k^2}\frac{(x_0-m_1)^2}{2\sigma_2^2}=\frac{(x_0-m_2)^2}{2\sigma_2^2}$ となり $\left|\frac{x_0-m_1}{x_0-m_2}\right|=k$ となる．ところが (4) によって $\frac{m_1-x_0}{m_2-x_0}=A$ と定義したものであるから $A=\pm k$, §5 の b′) により $A=-k$ とする．

η との関係は (8) 式を変形して

$$\frac{\bar{X}_2}{\sigma_2}=\sqrt{\frac{(c+1)(ck^2+1)}{c(A-1)^2}\frac{\eta^2}{1-\eta^2}} \tag{12}$$

$c=k$ なる故

$$\frac{\bar{X}_2}{\sigma_2}=\sqrt{\frac{(k^3+1)}{k(k+1)}\frac{\eta^2}{1-\eta^2}}=-\frac{\bar{X}_1}{\sigma_1} \tag{13}$$

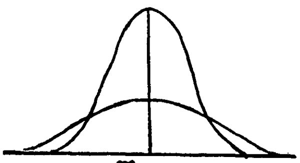
かくて与えられた $\frac{\bar{X}_1}{\sigma_1}$, $\frac{\bar{X}_2}{\sigma_2}$ より (i), (ii) とも P がきまり，相関比と的中率との関係が求められる．

$\eta=1$ であれば (12) 式より

$\frac{\bar{X}_2}{\sigma_2}=\infty$, $\frac{\bar{X}_1}{\sigma_1}=-\infty$ となり $P=1$.

さて，次に $\eta=0$ の場合を眺めてみよう．このときは $m_1=m_2$ となるから，(9) 式より $A=+1$（等根）が得られる．(第1，第2群の判別はこの場合任意)．

iii) $\eta=0$, $k\neq c$ の場合



$m_1=m_2$

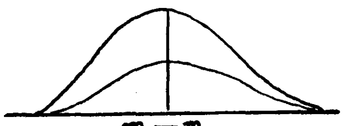
$c>k$, $1>k^2$ あるいは $c<k$, $1<k^2$ であれば（注，参照）(5), (6) 式より

$$\begin{cases}\dfrac{\bar{X}_1}{\sigma_1}=\sqrt{2\log\dfrac{c}{k}\cdot\dfrac{1}{1-k^2}} & (14)\\[2ex] \dfrac{\bar{X}_2}{\sigma_2}=k\dfrac{\bar{X}_1}{\sigma_1} & (15)\end{cases}$$

特に $k=1$ のときは交点は無限遠点にあると考えて，的中率は (1) 式より

$c<1$ ならば $P=\int_{-\infty}^{-\infty}p_1f_1(x)dx+\int_{-\infty}^{\infty}p_2f_2(x)dx=p_2$

$c>1$ ならば $P=\int_{-\infty}^{\infty}p_1f_1(x)dx+\int_{-\infty}^{\infty}p_2f_2(x)dx=p_1$



$m_1=m_2$

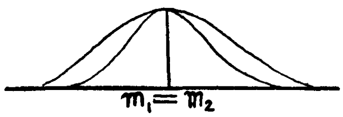
［注］ $c>k$, $1<k^2$ 或いは $c<k$, $1>k^2$ であれば (14), (15) 式は虚数となる．そこで $\frac{\bar{X}_1}{\sigma_1}$, $\frac{\bar{X}_2}{\sigma_2}$ を実数として与える η の最小値を考えると，(11) 式より

$$\eta=\sqrt{\frac{k^2-1}{k^2(1+\gamma)-1}},\quad A=k^2$$

このときの $\frac{\bar{X}_1}{\sigma_1}$, $\frac{\bar{X}_2}{\sigma_2}$ は

$$\begin{cases}\dfrac{\bar{X}_1}{\sigma_1}=k\dfrac{\bar{X}_2}{\sigma_2}\\[2ex] \dfrac{\bar{X}_2}{\sigma_2}=\sqrt{2\log\dfrac{c}{k}\dfrac{1}{k^2-1}}\end{cases}$$

iv）$\eta=0$，$k=c$ の場合



両分布が mode において，接した場合である．

$$m_1=m_2=x_0$$

このときの成功率は (1) 式の考え方（両分布を，分つ事を目的とする方式）よりすれば

$$P=\int_{-\infty}^{0} p_1 f_1(x)dx+\int_{0}^{\infty} p_2 f_2(x)dx=0.5$$

併しながら，実際には error を最小にするという考え方（すべて大きい方のグループに属するとすれば，判断の失敗は 0.5 より小さい）から

$$P=\max.\ (p_1,\ p_2)$$

としよう．

以上，$\eta\fallingdotseq 0$，$\eta=0$ にわけて，相関比と成功率の関係を大略述べたが，η が 0 とか 1 であるような特殊な場合を除いて，一般にはパラメーター A――2根あってその選択の問題が残っているもの――と，的中率の関係を摑むことが必要となる．

§4. パラメーター A について

先に述べたように $\dfrac{m_1-x_0}{m_2-x_0}=A$ と定義したが，一般に k，c を与えて η を動かせば，A は A_1，A_2 の2つの値をとる．ところで，この A_1，A_2 に対応する各グループの分布の交点（今まで x_0 として表わしたもの）を x_{01}，x_{02} で表わし

$$\frac{m_1-x_{01}}{m_2-x_{01}}=A_1,\quad \frac{m_1-x_{02}}{m_2-x_{02}}=A_2$$

とおく．このとき，いずれの A をとってその的中率を算出したらよいか．勿論，的中率をより大きくする A を採用したい．

今 (10) 式より

$$A_1+A_2=\frac{2\alpha}{\alpha-1}$$

$$A_1\cdot A_2=\frac{\alpha+k^2}{\alpha-1}$$

ただし　$\alpha=\dfrac{1}{\gamma}\left(\dfrac{1}{\eta^2}-1\right)$

をつかって，α を動かし A_1，A_2 の関係をみると第1表のようになった．即ち A_1，A_2 は2根が共に負となることはなく，かつ，その他のすべての関係をとり得る．しかも，k，c の大小関係が，α を動かす事によって得られた．なお，注目すべき事は $\alpha=0$ を中心として A_1，A_2 の関係が対蹠的となっている事である．ここで，次のように考えてみよう．

各分布の m_1，m_2 における密度函数をそれぞれ h_1，h_2 とすると

$$h_1=\frac{c}{k}h_2$$

ここでパラメーター k，c は共に負とならないから，h_1，h_2 については

$$\text{(i)}\ \ \frac{c}{k}>1,\qquad \text{(ii)}\ \ \frac{c}{k}=1,\qquad \text{(iii)}\ \ 0<\frac{c}{k}<1$$

の3つの場合について考えればよい．ところが，α と k，c の関係は第1表の通りであるから A_1，A_2 の関係が $\alpha=0$ について対蹠的であることは即ち h_1，h_2 の比 c/k が，1 より大であるか小であるかということである．このことを念頭において，第1表の分類における A_1，A_2 のいずれが的中率を大にするかを考えてみよう．

第 1 表

α の動き	類別	$A_{,1}$ A_2 の関係	
$\alpha=\infty$	a 2″	$A_1=A_2^*=1$	$\eta=0$ (k, c 任意)
$\alpha>1$	a 1	$1<A_1\leqq A_2^*$ $k>1$	$\frac{c}{k}>1$
	a 2	$0<A_1<1<A_2^*$ $k<1$	
	a 2′	$A_1=1<A_2^*$ $k=1$	
$\alpha=1$	a 3	$A_1=\frac{k^2+1}{2}$, $A_2^*=\infty$	
$1>\alpha>0$	b	$A_1^*<0<A_2$	
$\alpha=0$	b′	$A_1^*=-k$, $A_2=k$	$\frac{c}{k}=1$, 或いは $\eta=1$ (k, c 任意)
$-k^2<\alpha<0$	b	$A_1^*<0<A_2$	$0<\frac{c}{k}<1$
$\alpha=-k^2$	a 3	$A_1^*=0$, $A_2=\frac{2k^2}{k^2+1}$	
$\alpha<-k^2$	a 1	$0<A_1^*\leqq A_2<1$ $k>1$	
	a 2	$0<A_1^*<1<A_2$ $k<1$	
	a 2′	$0<A_1^*<1=A_2$ $k=1$	
$\alpha=-\infty$	a 2″	$A_1^*=A_2=1$	$\eta=0$ (k, c 任意)

*…的中率を大にする方
a …A_1, A_2 が共に正の場合
b …A_1, A_2 がそれぞれ正負の場合

§5. A_1, A_2 の選択法

今，各分布の平均については $m_1<m_2$ とし，分布の交点を x_{01}, x_{02} で表現する．m_1, m_2, x_{0i} $(i=1, 2)$ の排列を考えると

(イ)	$x_{0i}<m_1<m_2$
(ロ)	$m_1<x_{0i}<m_2$
(ハ)	$m_1<m_2<x_{0i}$
(ニ)	$x_{0i}=m_1$
(ホ)	$x_{0i}=m_2$

のいづれかとなる．

ここで，m_1, m_2, x_{01}, x_{02} の各点における分布の高さを $p_if_i(m_1)$, $p_if_i(m_2)$, $p_if_i(x_{01})$, $p_if_i(x_{02})$; $i=1, 2$ (グループ番号，ただし平均値の小さい方を第I群とする) で表わしておこう．

以下，第1表の類別に従い A_1, A_2 の選択を検討する．

(a1) $1<A_1\leqq A_2\left(\frac{c}{k}>1\right)$, $0<A_1\leqq A_2<1\left(\frac{c}{k}<1\right)$

(2根共1より大，および1より小但し正の場合)

(i) $1<A_1\leqq A_2\left(\frac{c}{k}>1\right)$ の場合

$A_1=\frac{m_1-x_{01}}{m_2-x_{01}}>1$ を満足させる x_{01} は上記 (イ)～(ホ) のうち

$\boxed{(\text{ハ})\quad m_1<m_2<x_{01}}$

同様にして $A_2>1$ に対する排列も

$\boxed{(\text{ハ})\quad m_1<m_2<x_{02}}$ となる.

x_{01}, x_{02} の排列は

$$A_2-A_1=\frac{(m_2-m_1)(x_{01}-x_{02})}{(m_2-x_{02})(m_2-x_{01})}\geqq 0$$ なるために

$$x_{01}\geqq x_{02}$$

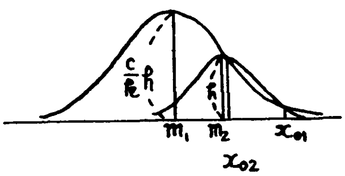
よって, m_1, m_2, x_{01}, x_{02} の排列は

$$m_1<m_2<x_{02}\leqq x_{01}$$

次に的中率

$$P_{A_1(x_{01})>1}=p_1\int_{-\infty}^{m_1}f_1(x)dx+p_1\int_{m_1}^{x_{02}}f_1(x)dx+p_1\int_{x_{02}}^{x_{01}}f_1(x)dx+p_2\int_{x_{01}}^{\infty}f_2(x)dx,$$

$$P_{A_2(x_{02})>A_1}=p_1\int_{-\infty}^{m_1}f_1(x)dx+p_1\int_{m_1}^{x_{02}}f_1(x)dx+p_2\int_{x_{02}}^{x_{01}}f_2(x)dx+p_2\int_{x_{01}}^{\infty}f_2(x)dx$$



を比べるためには $p_1\int_{x_{02}}^{x_{01}}f_1(x)dx$ と $p_2\int_{x_{02}}^{x_{01}}f_2(x)dx$ を比べればよい. 交点は x_{01} と x_{02} しかないから, 問題は区間 $[x_{02}, x_{01}]$ において $p_1f_1(x)>p_2f_2(x)$ であるか $p_1f_1(x)\geqq p_2f_2(x)$ であるかということにある. もし $p_1f_1(x)<p_2f_2(x)$ とすれば

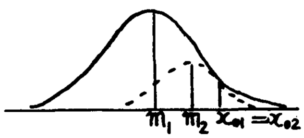
$$p_1\int_{x_{02}}^{x_{01}}f_1(x)dx<p_2\int_{x_{02}}^{x_{01}}f_2(x)dx$$

故に的中率は

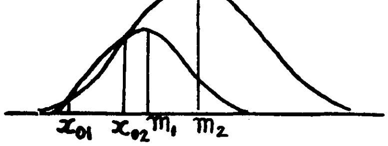
$$P_{A_1(x_{01})>1}<P^*_{A_2(x_{02})>A_1}$$

（以下 * は的中率の大なる方を示す）

$p_1f_1(x)\geqq p_2f_2(x)$ とすれば $h_1=\dfrac{c}{k}h_2>h_2$ であるから, $p_1f_1(x)$ と $p_2f_2(x)$ は x_{02} において接線を共有すべきである. この時は $p_1f_1(x)=p_2f_2(x)$, $x_{02}=x_{01}$, $A_2=A_1$ となり他に交点は存在しない. 的中率は $P_{A_1}=P_{A_2}$, 但しこの場合の P は §3 の iv と同様に考えて, $P=\max.(p_1, p_2)=p_1$ とする.



（ii） $0<A_1\leqq A_2<1$ $\left(\dfrac{c}{k}<1\right)$ のときも上と同じ方式で, x_{01}, x_{02}, m_1, m_2 の排列は, $x_{01}\leqq x_{02}<m_1<m_2$



的中率は

$$P^*_{A_1(x_{01})>1}\geqq P_{A_2(x_{02})>A_1}$$

となる.

(a2) $0<A_1<1<A_2$ の場合

（2根とも正であって1より大なるものと小なるものに別れた場合）

$$0<A_1=\frac{m_1-x_{01}}{m_2-x_{01}}<1$$ を満足させる x_{01} は（イ）～（ホ）のうち

$\boxed{(\text{イ})\quad x_{01}<m_1<m_2}$ のみ

他方 $A_2=\dfrac{m_1-x_{02}}{m_2-x_{02}}>1$ を満足させる x_{02} は

$\boxed{（ハ）\ m_1<m_2<x_{02}}$ のみ．

そこで（イ）（ハ）をまとめて

$x_{01}<m_1<m_2<x_{02}$

即ち，この場合の m_1 m_2 は必らず２つの交点 x_{01}，x_{02} にはさまれる．

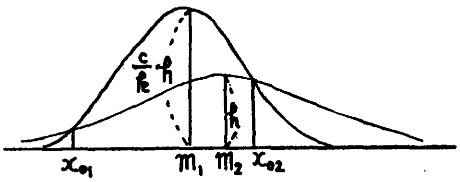
そこで第１表より

（i） $\dfrac{c}{k}>1$

(ii) $0<\dfrac{c}{k}<1$

の２つの場合を考えよう．

（i） $\dfrac{c}{k}>1$ のとき：



上の証明で m_1，m_2 は x_{01}，x_{02} にはさまれかつ，交点は x_{01}，x_{02} の他にないから，x_{01}，x_{02} にはさまれる両分布の面積は

$$p_1\int_{x_{01}}^{x_{02}} f_1(x)dx>p_2\int_{x_{01}}^{x_{02}} f_2(x)dx$$

$$\therefore\quad P_{A_1(x_{01})<1}=p_1\int_{-\infty}^{x_{01}} f_1(x)dx+p_2\int_{x_{01}}^{x_{02}} f_2(x)dx+p_2\int_{x_{02}}^{+\infty} f_2(x)dx<P_{A_2(x_{02})>1}$$

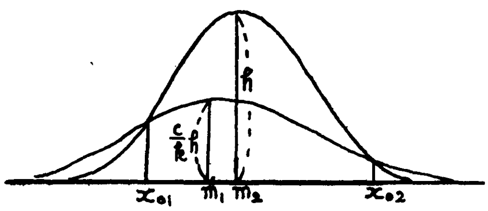
$$=p_1\int_{-\infty}^{x_{01}} f_1(x)dx+p_1\int_{x_{01}}^{x_{02}} f_1(x)dx+p_2\int_{x_{02}}^{+\infty} f_2(x)dx$$

$$P_{A_1(x_{01})<1}<P^*_{A_2(x_{02})>1}$$

(ii) $0<\dfrac{c}{k}<1$ のときも (i) と同じようなやり方で，

$$p_1\int_{x_{01}}^{x_{02}} f_1(x)dx>p_2\int_{x_{01}}^{x_{02}} f_2(x)dx$$

$$\therefore\quad P^*_{A_1(x_{01})<1}>P_{A_2(x_{02})>1}$$



(a2) の特殊な場合として，

(a2′) （i） $A_1=1<A_2$ $\left(\dfrac{c}{k}>1\right)$

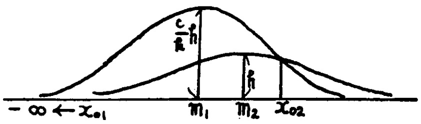
(ii) $0<A_1<1=A_2$ $\left(\dfrac{c}{k}<1\right)$

但し (10) 式より $k=1$

がある．即ち交点の１つが無限遠点となった場合である．両分布はガウス分布であって，左右対称であるから，$\dfrac{c}{k}>1$ の場合は交点は右側が有限点，左側が無限遠点 $\left(\dfrac{c}{k}<1\right.$ の場合は逆$\left.\right)$ と考えられる．即ち x_{01}，x_{02}，m_1，m_2 の排列は

（i） $x_{01}=-\infty<m_1<m_2<x_{02}$ $\left(\dfrac{c}{k}>1\right)$

(ii) $x_{01}<m_1<m_2<x_{02}=\infty$ $\left(\dfrac{c}{k}<1\right)$



的中率は (a2) と同じく

(i) $\frac{c}{k}>1$ のとき

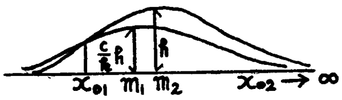
$$p_1\int_{x_{01}=-\infty}^{x_{02}} f_1(x)dx > p_2\int_{x_{01}=-\infty}^{x_{02}} f_2(x)dx$$

$$P_{A_1(x_{01}=-\infty)}=p_2\int_{x_{01}=-\infty}^{x_{02}} f_2(x)dx+p_2\int_{x_{02}}^{+\infty} f_2(x)dx<P_{A_2(x_{02}>0)}$$

$$=p_1\int_{x_{01}=-\infty}^{x_{02}} f_1(x)dx+p_2\int_{x_{02}}^{\infty} f_2(x)dx$$

$$\therefore\quad P_{A_1(x_{01}=-\infty)}<P^{*}_{A_2(x_{02})=1}$$

(ii) $\frac{c}{k}<1$ のとき



上記 (i) と同じやり方で

$$P^{*}_{A_1(x_{01})>1}>P_{A_2(x_{02})=1}$$

(a2″) $A_1=A_2=1$ の場合は $m_1=m_2$, $\eta=0$ となり成功率は §3 に述べた通りである.

次に A_1, A_2 のうち特殊点 (0, ∞, 1) をもつ場合について考察すると,

(a3) $A_1=\frac{k^2+1}{2}$, $A_2=\infty$ $\left(\frac{c}{k}>1\right)$

$A_1=0$, $A_2=\frac{2k^2}{k^2+1}$ $\left(\frac{c}{k}<1\right)$

(2つの交点のうち, 1つが mode と一致した場合)

(1) $A_2=\infty$ となった場合 $\left(\frac{c}{k}>1\right)$

(11) 式より $\alpha=1$ は $A=\infty$ を満足させる. そこで $\alpha=1$ とすると (10) 式より

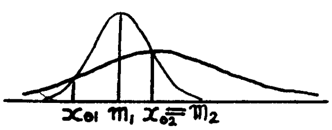
$$A_1=\frac{1+k^2}{2}>0 \quad ここに\ k>0$$

即ち片方が ∞ となる場合は他方は常に正となる. $A_1>0$, $A_2=\infty$ となるような m_1, m_2, x_{01}, x_{02} の排列は

(イ)	$x_{01}<m_1<m_2$
(ハ)	$m_1<m_2<x_{01}$
(ホ)	$x_{02}=m_2$

より (i) $x_{01}<m_1<x_{02}=m_2$
(ii) $m_1<m_2=x_{02}<x_{01}$

となる.

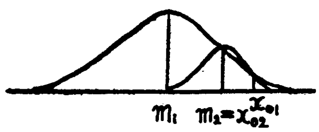


(i) $x_{01}<m_1<x_{02}=m_2$

(a2) の (i) と同じ証明法で的中率は

$$P_{A_1(x_{01})>0}<P^{*}_{A_2(x_{02})=\infty}$$

(ii) $m_1<m_2=x_{02}<x_{01}$



$P_{A_1(x_{01})>0}$ と $P_{A_2(x_{02}=m_2)}$ を比べるためには $p_1\int_{x_{02}=m_2}^{x_{01}} f_1(x)dx$ と $p_2\int_{x_{02}=m_2}^{x_{01}} f_2(x)dx$ を比べればよい. 今, 区間 $[x_{02}\ x_{01}]$ において $p_1f_1(x)\geqq p_2f_2(x)$ であると仮定しよう. $m_2=x_{02}$ における $p_1f_1(x)$, $p_2f_2(x)$ は $p_2f_2(x)$ が $p_1f_1(x)$ に対して下から接する状態となり, 両

者は x_{02} において接線を共有すべきである．ところが $x_{02}=m_2 \neq m_1$ であって，$p_2 f_2'(x_{02})=0$，$p_1 f_1'(x_{02}) \neq 0$，故にこの仮定は成り立たない．即ち両曲線は x_{02} において互いに交り，かつ交点は x_{02}，x_{01} しかないから $[x_{02},\ x_{01}]$ においては

$$p_1 f_1(x) < p_2 f_2(x)$$

$$\therefore\quad p_1 \int_{x_{02}=m_2}^{x_{01}} f_1(x)dx < p_2 \int_{x_{02}=m_2}^{x_{01}} f_2(x)dx$$

的中率は

$$P_{A_1(x_{01})>0} < P^*_{A_2(x_{02})=\infty}$$

（2） $A_1=0$ となった場合 $\left(\dfrac{c}{k}<1\right)$

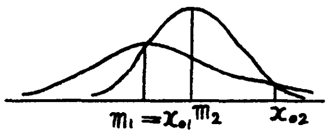
(11) 式より $\alpha \neq 1$ として

$$A_2=\frac{2k^2}{1+k^2}>0$$

$\alpha=1$ とすると (10) 式より $A_1=\dfrac{1+k^2}{2}=0$ であるからこれを満足する k は虚数となる．故に $\alpha=1$，即ち A_2 が ∞ となることはない．ところで，$A_1=0$ ということは，x_{01} が mode と一致していることであって，(1) の x_{02} が mode と一致している場合と同じ型である．m_1，m_2，x_{01}，x_{02} の排列は

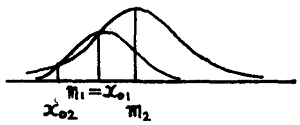
(i)　$m_1=x_{01}<m_2<x_{02}$　　(ii)　$x_{02}<m_1=x_{01}<m_2$

的中率は (1) と全く同じやり方で



（i）　$m_1=x_{01}<m_2<x_{02}$

$$\int_{x_{01}}^{x_{02}} p_1 f_1(x)dx < \int_{x_{01}}^{x_{02}} p_2 f_2(x)dx$$



(ii)　$x_{02}<m_1=x_{01}<m_2$

$$\int_{x_{02}}^{x_{01}} p_1 f_1(x)dx > \int_{x_{02}}^{x_{01}} p_2 f_2(x)dy$$

故に (i) も (ii) も共に

$$P^*_{A_1(x_{01})=0} > P_{A_2(x_{02})>0}$$

以上，A_1，A_2 が共に正なる場合について，両者の選択の詳細を述べた．ところで，これを纏めると非常に簡単な結論となる．即ち

a)　A_1，A_2 が共に正なる場合

$\dfrac{c}{k}>1$　なら　　大なる方

$0<\dfrac{c}{k}<1$　なら　　小なる方

の A をとるべきであることがわかった．

（b）　$A_1<0<A_2$　（A_1，A_2 がそれぞれ正負に別れた場合）

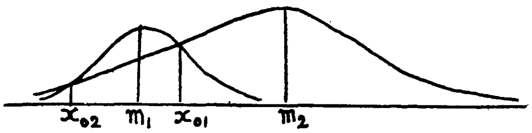
$A_1<0<A_2$ を満足させる x_{01}，x_{02}，m_1，m_2 の排列は

（イ）	$x_{02}<m_1<m_2$
（ロ）	$m_1<x_{01}<m_2$
（ハ）	$m_1<m_2<x_{02}$

より $\begin{cases} \text{(i)} & x_{02}<m_1<x_{01}<m_2 \\ \text{(ii)} & m_1<x_{01}<m_2<x_{02} \end{cases}$

各グループはガウス分布に従っているから

(i)　$x_{02}<m_1<x_{01}<m_2$ の場合



第I群については

$$p_1f_1(x_{02})<p_1f_1(m_1)$$
$$>p_1f_1(x_{01})>p_1f_1(m_2)$$

第II群については

$$p_2f_2(x_{02})<p_2f_2(m_1)$$
$$<p_2f_2(x_{01})<p_2f_2(m_2)$$

x_{01}, x_{02} は分布の交点であるから

$$p_1f_1(x_{02})=p_2f_2(x_{02})$$
$$p_1f_1(x_{01})=p_2f_2(x_{01}),$$

ところが

$$p_1f_1(m_1)>p_1f_1(x_{01})$$
$$p_2f_2(m_1)<p_2f_2(x_{01})=p_1f_1(x_{01})$$
$$\therefore\quad p_1f_1(m_1)>p_2f_2(m_1)$$

而して交点は，x_{02}, x_{01} のみなる故，x_{02}, x_{01} にはさまれる両グループの面積は

$$p_1\int_{x_{02}}^{x_{01}}f_1(x)dx>p_2\int_{x_{02}}^{x_{01}}f_2(x)dx$$

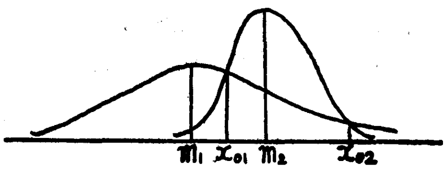
この関係を使って判断的中率は

$$P_{A_1(x_{01})<0}=p_1\int_{-\infty}^{x_{02}}f_1(x)dx+p_1\int_{x_{02}}^{x_{01}}f_1(x)dx+p_2\int_{x_{01}}^{+\infty}f_2(x)dx>P_{A_2(x_{02})>0}$$
$$=p_1\int_{-\infty}^{x_{02}}f_1(x)dx+p_2\int_{x_{02}}^{x_{01}}f_2(x)dx+p_2\int_{x_{01}}^{\infty}f_2(x)dx$$
$$P^*_{A_1(x_{01})<0}>P_{A_2(x_{02})>0}$$

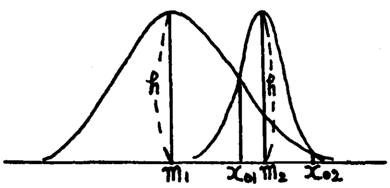
(ii) の場合も全く同様の証明で

$$P^*_{A_1(x_{01})<0}>P_{A_2(x_{02})>0}$$

最後に特殊な場合として



(b′)　$A_1=-k$, $A_2=k$ がある．(10) 式より $\alpha=0$ として得られたもので，$\dfrac{1}{\gamma}\left(\dfrac{1}{\eta^2}-1\right)=\alpha$



から $\dfrac{c}{k}=1$, あるいは $\eta=1$ (k, c 任意) を得る．A_1, A_2 の選択は (b) と同様，負の A, ($A_1=-k$) をとればよい．

以上，b についてまとめると

b)　A_1, A_2 がそれぞれ正負の場合は k, c の大小にかかわりなく，負の A をとるべきである．

§6. 的中率と相関比との関係，並びに図表

まず，以上述べ来ったことをまとめてみよう．計算法としては，k, c, η を与えて

$$\eta^2=\frac{1}{1+\gamma\dfrac{A^2-k^2}{(A-1)^2}},\quad ただし\quad \gamma=\frac{(c+1)(ck^2+1)}{2ck^2\log\dfrac{c}{k}}\tag{9}$$

一般に

$$A=\frac{-\alpha\pm\sqrt{\alpha+k^2(1-\alpha)}}{(1-\alpha)},\quad ただし\quad \alpha=\frac{1}{\gamma}\left(\frac{1}{\eta^2}-1\right)$$

より A_1, A_2 を求め

(a) A_1, A_2 がともに正であるなら

(i) $c>k$ のときは，大きい方

(ii) $c<k$ のときは，小さい方

(b) A_1, A_2 がそれぞれ正負なら負の方をとって，

$$\begin{cases} \dfrac{\bar{X}_1}{\sigma_1}=\dfrac{A}{k}\dfrac{X_2}{\sigma_2} & (5) \\ \dfrac{\bar{X}_2}{\sigma_2}=k\sqrt{2\log\dfrac{c}{k}\dfrac{1}{A^2-k^2}} & (6) \end{cases}$$

より，上下限 $\dfrac{\bar{X}_1}{\sigma_1}$, $\dfrac{\bar{X}_2}{\sigma_2}$ を得る．的中率 P は

(a) の場合

$$P=0.5+\left|\frac{c}{c+1}\int_0^{\left|\frac{\bar{X}_1}{\sigma_1}\right|}\varphi(t)dt-\frac{1}{c+1}\int_0^{\left|\frac{\bar{X}_2}{\sigma_2}\right|}\varphi(t)dt\right| \tag{7-a}$$

(b) の場合

$$P=0.5+\frac{c}{c+1}\int_0^{\left|\frac{\bar{X}_1}{\sigma_1}\right|}\varphi(t)dt+\frac{1}{c+1}\int_0^{\left|\frac{\bar{X}_2}{\sigma_2}\right|}\varphi(t)dt \tag{7-b}$$

となる．

特に (b) において $k=c$ なるときは，$A=-k$ となり

$$\begin{cases} \dfrac{\bar{X}_1}{\sigma_1}=-\dfrac{\bar{X}_2}{\sigma_2} \\ \dfrac{\bar{X}_2}{\sigma_2}=\sqrt{\dfrac{(k^3+1)}{k(k+1)}\cdot\dfrac{\eta^2}{1-\eta^2}} \end{cases}$$

として，(7-b) 式より的中率を求めればよい．$k=c=1$ のときは，max.-min. の方法において $\sigma_1=\sigma_2$ としたときと同じになる．

次に，A が特異点 $(\infty, 0)$ を含む場合については，$A=\infty$ なら，

$$\begin{cases} \dfrac{\bar{X}_1}{\sigma_1}=-\sqrt{2\log\dfrac{c}{k}} \text{ [2)]} \\ \dfrac{\bar{X}_2}{\sigma_2}=0 \end{cases}$$

$$P=0.5+\frac{c}{c+1}\int_0^{\left|\frac{\bar{X}_1}{\sigma_1}\right|}\varphi(t)dt \tag{7-c-1}$$

$A=0$ なら

$$\begin{cases} \dfrac{\bar{X}_1}{\sigma_1}=0 \\ \dfrac{\bar{X}_2}{\sigma_2}=\sqrt{2\log\dfrac{c}{k}} \end{cases}$$

となり的中率は

$$P=0.5+\frac{1}{c+1}\int_0^{\left|\frac{\bar{X}_2}{\sigma_2}\right|}\varphi(t)dt \tag{7-c-2}$$

---

2) $A=\infty$, 即ち $m_2=x_0$ の場合で

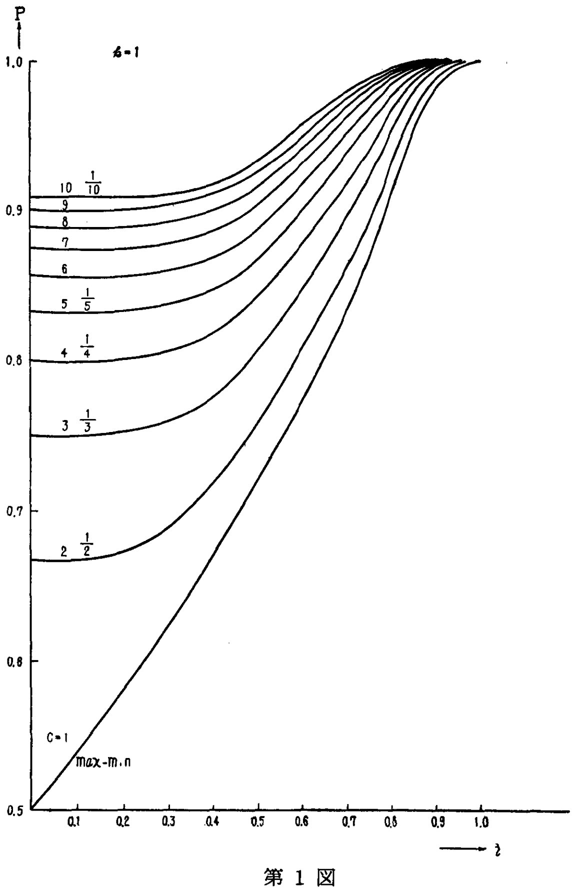
$$\log\frac{p_1}{p_2}\frac{\sigma_2}{\sigma_1}=\frac{(x_0-m_1)^2}{2\sigma_1^2}-\frac{(x_0-m_2)^2}{2\sigma_2^2} \qquad (2) \text{ より}$$

$$\frac{\bar{X}_1}{\sigma_1}=-\sqrt{2\log\frac{c}{k}}\ , \qquad \frac{\bar{X}_2}{\sigma_2}=0$$

最後に $\eta=0$ のときは，$k\neq c$ ならば，$A=1$ として (5)，(6)，(7-a) 式より的中率は求められる．特に $c=k$ あるいは $k=1$ のときは $P=\max.(p_1,\ p_2)$．

いずれの場合にも実際には，max. $(p_1,\ p_2)$ 以上の的中率と相関比との関係が論議される．

かくして，与えられた k, c, η に対して，成功率が定まった．ここでは k, c を $\frac{1}{10}$, $\frac{1}{5}$, $\frac{1}{4}$, $\frac{1}{3}$, $\frac{1}{2}$, 1, 2, 3, … 9, 10 としたときの相関比と成功率の関係を求めたが，以下はそのうちの代表的な図表および数表である．第1図では $k=1$ の下に P, η の関係に対する c の影響をみた．max.-min. の方法の c による変動をみたものともいえよう．なお，パラメーター k は，個体のもつ得点を個々に算出することなく，数量比の計算過程において容易に得られるものであるが，これを未知数として c, η から P を推定する場合を考慮して，k の極限値[3]に対する P の値を第 2，第 3，第4図に点線で画いてみた．実際には k は $\frac{1}{10}$〜10 まで動くとすれば充分である．



第 1 図

3) k が ∞ に近ずいた時の P の考え方は σ_2 の極限を 0 として第2グループを1点分布と考え

$$P=p_1\int_{-\infty}^{\frac{\bar{X}_1}{\sigma_1}}\varphi(x)dx+p_2 \qquad \text{とした．}$$

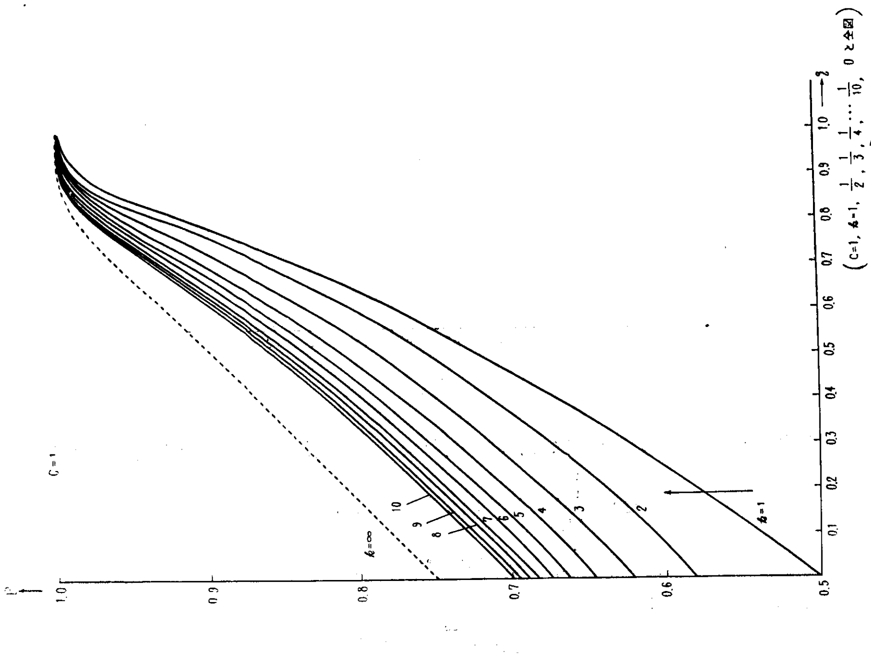
ここに

$$\frac{\bar{X}_1}{\sigma_1}=\sqrt{(c+1)\cdot\left(\frac{\eta^2}{1-\eta^2}\right)}$$

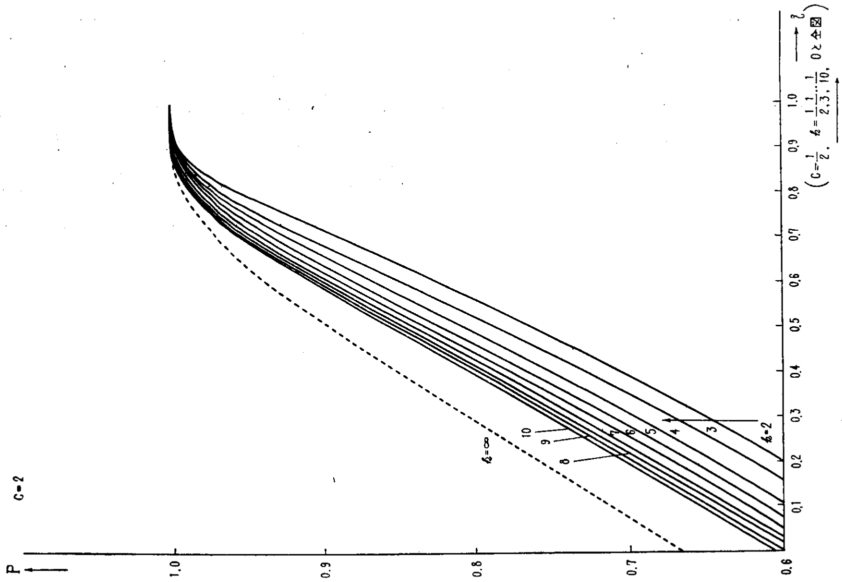
k が 0 のときは $\sigma_1 \longrightarrow 0$ とし

$$P=p_1+p_2\int_{-\frac{\bar{X}_2}{\sigma_2}}^{\infty}\varphi(t)dt$$

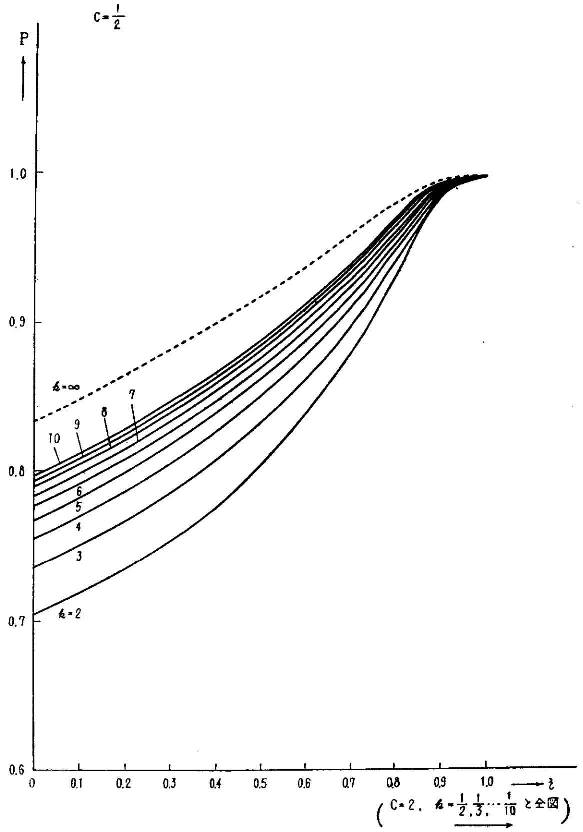
$$\frac{\bar{X}_2}{\sigma_2}=\sqrt{\frac{c+1}{c}\cdot\frac{\eta^2}{1-\eta^2}}$$



第 2 図



第 3 図



第 4 図

第 I 表 $P(k=3)$

η \ c	1	2	3	4	5	6	7	8	9	10
0	.6210	.5267	.5000							
0.1	.6480	.5733	.5611							
0.2	.6775	.6220	.6220							
0.3	.7105	.6733	.6851							
				$\begin{pmatrix}.3009\\.7100\end{pmatrix}$	$\begin{pmatrix}.3591\\.7751\end{pmatrix}$	$\begin{pmatrix}.3839\\.8147\end{pmatrix}$	$\begin{pmatrix}.3954\\.8419\end{pmatrix}$			
0.4	.7465	.7280	.7471	.7742	.8004	.8245	.8443			
								$\begin{pmatrix}.4005\\.8620\end{pmatrix}$	$\begin{pmatrix}.4022\\.8779\end{pmatrix}$	$\begin{pmatrix}.4017\\.8902\end{pmatrix}$
0.5	.7870	.7850	.8111	.8382	.8627	.8840	.9008	.9154	.9266	.9366
0.6	.8325	.8450	.8740	.9000	.9207	.9367	.9498	.9594	.9675	.9737
0.7	.8850	.9060	.9331	.9522	.9667	.9763	.9831	.9877	.9912	.9937
0.8	.9415	.9633	.9791	.9882	.9934	.9964	.9978	.9988	.9989	.9999
0.9	.9910	.9973	.9991	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

(　　)…上段は $\dfrac{\bar{X}_1}{\sigma_1}$, $\dfrac{\bar{X}_2}{\sigma_2}$ を実数として与える η の最小値，下段は成功率

第 II 表 $P\left(k=\frac{1}{3}\right)$

η \ c	1	2	3	4	5	6	7	8	9	10
0	.6210	.7353	.7975	.8358	.8616	.8811	.8955	.9060	.9155	.9225
0.1	.6480	.7507	.8080	.8436	.8687	.8870	.9003	.9112	.9194	.9269
0.2	.6780	.7676	.8202	.8532	.8764	.8935	.9062	.9162	.9248	.9311
0.3	.7105	.7863	.8332	.8646	.8853	.9010	.9147	.9231	.9301	.9368
0.4	.7465	.8087	.8497	.8770	.8963	.9158	.9214	.9300	.9373	.9431
0.5	.7870	.8337	.8685	.8922	.9090	.9220	.9314	.9391	.9457	.9504
0.6	.8325	.8636	.8905	.9100	.9245	.9353	.9434	.9500	.9557	.9605
0.7	.8850	.8997	.9188	.9330	.9435	.9522	.9586	.9638	.9684	.9714
0.8	.9415	.9433	.9530	.9608	.9677	.9727	.9770	.9806	.9829	.9850
0.9	.9610	.9877	.9893	.9914	.9936	.9947	.9956	.9963	.9968	.9980

第 III 表 $P(k=5)$

η \ c	1	2	3	4	5	6	7	8	9	10
0	.6615	.5667	.5258	.5080	.5000					
0.1	.6895	.6124	.5860	.5788	.5820					
0.2	.7190	.6604	.6481	.6514	.6619					
						(.2175 .6910)	(.2726 .7533)			
0.3	.7505	.7107	.7116	.7240	.7400	.7574	.7756			
								(.3012 .7926)	(.3178 .8206)	(.3279 .8414)
0.4	.7850	.7633	.7738	.8118	.8140	.8324	.8507	.8644	.8792	.8918
0.5	.8230	.8186	.8383	.8610	.8810	.8991	.9143	.9275	.9384	.9467
0.6	.8650	.8750	.8983	.9194	.9380	.9516	.9627	.9706	.9768	.9821
0.7	.9100	.9297	.8513	.9668	.9780	.9846	.9896	.9934	.9956	.9967
0.8	.9590	.9757	.9875	.9942	.9970	.9980	.9990	1.0000	1.0000	1.0000
0.9	.9950	.9990	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

第 IV 表 $P\left(k=\frac{1}{5}\right)$

η \ c	1	2	3	4	5	6	7	8	9	10
0	.6615	.7673	.8225	.8594	.8797	.8964	.9090	.9185	.9269	.9331
0.1	.6895	.7833	.8343	.8652	.8872	.9025	.9146	.9235	.9305	.9372
0.2	.7190	.8000	.8460	.8750	.8950	.9094	.9197	.9289	.9353	.9416
0.3	.7605	.8190	.8595	.8856	.9040	.9170	.9264	.9347	.9407	.9457
0.4	.7850	.8394	.8748	.8972	.9127	.9246	.9337	.9411	.9467	.9512
0.5	.8230	.8626	.8910	.9104	.9239	.9348	.9421	.9488	.9535	.9583
0.6	.8650	.8893	.9111	.9260	.9375	.9460	.9520	.9577	.9616	.9657
0.7	.9110	.9210	.9343	.9450	.9538	.9594	.9647	.9684	.9720	.9744
0.8	.9590	.9570	.9631	.9684	.9730	.9770	.9797	.9817	.9913	.9859
0.9	.9950	.9917	.9923	.9930	.9938	.9944	.9951	.9957	99.69	.9973

ここで，2, 3 の例を掲げておこう．例は昭和 33 年から昭和 35 年にわたって，東京都 23 区で行なった市場調査のうち，冷蔵庫購入予測を扱ったものである．図表をつかって，c, k, η より P を読んだところ，データーより算出したものとよくあっている．ガウス分布の近似度も確率紙にの

せてみると，ほぼ直線的である．なお，例2について χ^2 検定を行なった結果

第I群　$0.10<P\{x^2>7.45\}<0.20,\ D.F.=4$

第II群　$0.80<P\{x^2>0.73\}<0.90,\ D.F.=3$

となった．

例 1)

昭和 33 年，氷冷蔵庫所有世帯の昭和 35 年電気冷蔵庫所持状況について，

第I群…電気冷蔵庫を買わなかった，

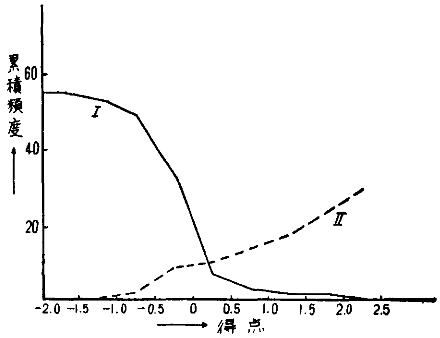
第II群…電気冷蔵庫を買った．

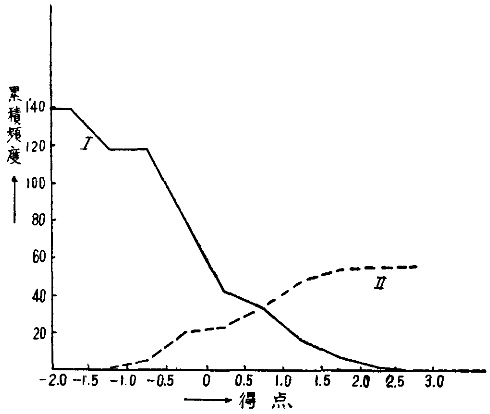
所持変化に対する数量比得点累積分布

$c=1.95,\ k=0.72,\ \eta=0.67,$

$\hat{P}=85.7\%$（テータより算出したもの），

$P=85.6\%$（図表より読んだもの）





例 2)

昭和 33 年，冷蔵庫未所有世帯の昭和 35 年冷蔵庫所持状況について，

第I群…冷蔵庫未所有，

第II群…冷蔵庫（電・氷）所有

所持変化に対する数量比得点累積分布

$c=2.5,\ k=1.0,\ \eta=0.455,$

$\hat{P}=75.0\%,\ P=76.6\%$

例 3)

昭和 33 年冷蔵庫未所有世帯の昭和 35 年冷蔵庫所持状況について，

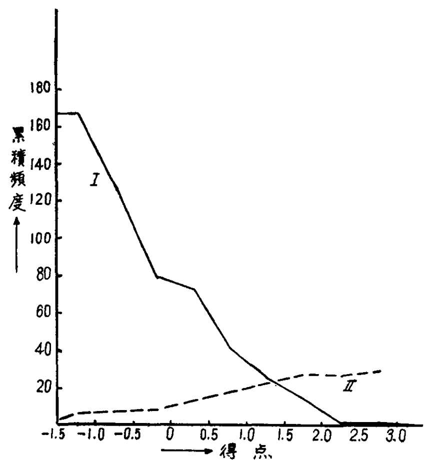
第I群…電気冷蔵庫未所有，

第II群…電気冷蔵庫所有

所持変化に対する数量化得点累積分布

$c=5.72,\ k=0.75,\ \eta=0.405,$

$\hat{P}=86.0\%,\ P=87.8\%$



§7. あ と が き

以上，図表をもって P と η との関係を表わした．これが実際の適用にどの位役立つかは，「各群に属する得点（尺度値）の分布はガウス分布である．」という仮定に対して，$P,\ \eta$ の関係が，如

何ほど，強靱さをもつかという問題にかかっている．今後，多くの資料を集めて確かめてみたい．

終りに，この報告をなすに当って，終始，ご指導いただいた青山博次郎先生に，心から感謝の意を捧げるものである．

（統計数理研究所）