九州大学学術情報リポジトリ
Kyushu University Institutional Repository

数個の母集団の比較法について

野町, 幸男
九州大学

https://doi.org/10.15017/12717

出版情報：統計科学研究. 1 (1), pp.21-39, 1956-01. Research Association of Statistical Sciences
バージョン：
権利関係：

綜　合　報　告

数個の母集団の比較法について

野　町　幸　男　（九大）

§1　序論　　40年代の終り頃から"*Ranking*"と名を冠せられる種類の論文が *Ann. Math. Stat.* や *Biometrika* を中心に数多く現われ始めた．これらの論文は，いづれも $k(\geqq 3)$ 個の母集団を対象としていることによって，なしうる統計的判定が多様であるという点とそれに伴う定式化が複雑であるという点で特徴づけられている．これらの定式化の複雑さは，従来の帰無仮説と対立仮説，云いかえると特殊な母集団の規定と普偏な母集団の規定という対立関係の上でなされてきた2つの母集団を中心とする仮説検定の理論では，必ずしも処理出来るものとは限らない．例えば，母集団 π_i の母平均 μ_i についての仮説 $H_0: \mu_i = \mu$ $(i=1, 2, \cdots, k)$ がある有意水準で棄却されたとする当面の問題が母平均にのみ依存するものであったにしろ，このことはどの母平均をもつ母集団を選ぶべきかの解答を意味するものではない．このとき何を目標に何を推測しようとするのかによって，極めて多くの判定が起りうるのである．k 個の母集団 π_i を母平均 μ_i の大きさの順に並べるという課題を，それぞれの標本平均 $\bar{x}_i$ の大きさによって順序づけるという判定方式によって判断するとしよう．そうすると $k!$ の場合が起りうるではないか．こうなると，従来の帰無仮説と対立仮説とゆう枠の中では，必しも取扱えないことが判ろう．こうゆうところに今から取あげる一連の論文の興味と困難があるのである．

われわれは又"*Ranking*"と名を冠せられるにふさわしく，如何なる形であるにしても，判定が母集団の順序づけを目的としているものに制限しよう．北川［1］はこの種の問題を更に一般的に『$[k_1, k_2, \cdots, k_k]$ 分割に関する推測』とゆう形で特徴づけている．こゝで考えられていることは，k 個 $\left(=\sum_1^k k_j\right)$ の母集団を種々の推測目標，云いかえると順序の定義如何によって k 個の組に分ける分割であり，同一組内の k_j $(j=1, 2, \cdots, k)$ 個のものに関しては，この推測目標については順序づけることが必要でないか，又は順序づけを行わないことを意味している．簡単な例を挙げると，k 個の母集団の中から最大の母平均 $\mu_k = \underset{1 \leqq i \leqq k}{Max}\{\mu_i\}$ をもつ π_k を選ぶ推測を考える．このとき，それぞれの母集団 π_i からの標本平均 $\bar{x}_i$

の最大値が，$\overline{x}_{M(j)}=\underset{1\leqq i\leqq k}{Max}\{\overline{x}_i\}$ であるとするときこれに対応する $\pi_{M(j)}$ を選ぶという判定は $[k-1, 1]$ の形式に属することが判る．$[1, k-1]$ は $Max.$ に対して $min.$ をもってくればよい．こゝで注意しなければないことは，”最大の母平均をもつ母集団を選ぶ“ という目標があることによって，既に $[k-s, s]$ という形式に於て，$s=1$ であるという前提をしている点である．実際問題では，$[k-s, s]$ の母集団分割は，必ずしも s を予め定めている場合が多くはなく，むしろ，上記の ”最大“ と ”最小“ の場合を除けば例外的であろう．云いかえると，ある推測目標をもって，実験データをみて，s を決めることが多いのではなかろうか．簡単な例を挙げるとしよう．前の例で，母平均の中で最大なものに対応している母集団が唯一つであるときは，何の疑問も生じないのであるが，もし，最大母平均に対応している母集団がいくつかあるか又はすべてがそうであるときは，単に一つの母集団のみを選ぶことに，若干の疑問を感ずる．このようなとき，$\overline{x}_{M(j)}$ に近い実現値をもつ $\overline{x}_i$ は，その対応している母集団 π_i が最大母平均をもつ母集団群の中の一つである可能性も強いと考えられるわけである．従って，ある適当な距離 $d(>0)$ を定め，$\overline{x}_{M(j)}-\overline{x}_i \leqq d$ ならば，対応する π_i を S-群（*superior group*），$\overline{x}_{M(j)}-\overline{x}_{i'}>d$ ならば，それに対応する $\pi_{i'}$ を I-群（*inferior group*）に判定する方式が考え得られる．このようなとき，S-群に属する π_i の個数 s は判定方式に依存する確率変数であることが判る．北川［1］はこのような場合を，推測過程論の立場からと同時推測論の立場からとに分類して取扱つている．併し乍ら，われわれは，後者の立場が *H. Scheffé*［1］や *J. W. Tukey*［1］の方法に代表されるように，従来の推測理論に沿って展開されているという理由によって，当面の興味から遠れるので割愛することにしよう．

以上のようにして，一般に $[k_1, k_2, \cdots, k_h]$ 分割に関する推測が，こゝで提起される課題となるのであるが，そこには複雑な概念と計算を伴うことが予想される．従って，当然のことではあるが，大部分は未解決のまゝ残されているし，特殊な場合のみが取扱われている．そして又，そこで用いられる解決の方法も，従来の仮説検定論の行きかたに強く影響され，未だ洗練されたものといゝ難いもののように感じられる．併し乍ら，前に述べたように，同時推測論の場合を除けば，何かそこに新しい特徴をもった概念と手法が芽生えているように思えるのである．

以下，§2に於て，共通する仮定と記号，§3に $[k-s, s]$ 分割に於て s が確定値であるもの，§4で，§3の形式で s が確率変数である場合を取扱い，§5はこれらの問題を理論的に解明し，その際採用されている判定方式を，どのよう

な條件の下で最適であるかを究明するものを，§6で，主として§5についてではあるが結言らしきものを述べることにする．

§2．仮定と記号　特別に断らない限り，われわれは，k個の母集団Π_i $(i=1,2,\cdots,k)$を考え，そこからの標本O_{n_i}を$(x_{i1},\cdots,x_{in_i})$とし，x_{ij}はそれぞれiとjに関し独立に$N(\mu_i,\sigma_i^2)$に従っているとする．この場合，σ_i^2が既知であるか，そのiについての比が既知である場合は，適当な変換によって$N(\mu_i,\sigma^2)$に帰せられるし，μ_iが既知ならば同様にして$N(\mu,\sigma_i^2)$に帰せられるので，問題によって前者と後者を使いわけることにする．更に$\mu_1\leqq\mu_2\leqq\cdots\leqq\mu_k$，又は，$\sigma_1\leqq\sigma_2\leqq\cdots\leqq\sigma_k$を仮定しても一般性を失わないので，それぞれの母集団Π_iからの標本平均$\bar{x}_i$は$E(\bar{x}_i)=\mu_i$，標本不偏分散s_i^2は$E(s_i^2)=\sigma_i^2$と仮定する．併し乍ら，標本平均$\bar{x}_i$は，実際的には，前記の仮定通りのiをもつ訳ではないのは云うまでもない．この$\bar{x}_i$ $(i=1,2,\cdots,k)$を順序によって並べたj番目のものを$\bar{x}_{(j)}$とする．即ち$\bar{x}_{1(i_1)}\leqq\bar{x}_{2(i_2)}\leqq\cdots\cdots\leqq\bar{x}_{k(i_k)}$．こゝで，$\bar{x}_{k(i_k)}$は特別に$\bar{x}_{M(j)}$として表わすことが多い．$\bar{x}_{M(j)}$の$j$は，標本平均が最大値となった母集団$\Pi_j$の番号である．一般には，$j$は$1\leqq j\leqq k$を変りうるのであることに注意を要する．又分散についても同様に$s^2_{(i)}, s^2_{M(j)}$を定義する。

§3　sが確定値であるときの$[k-s, s]$分割に関連するもの

[I] 平均値による順序づけ．σ_iがΠ_iによって異なるとき，母平均μ_iの大きさによってのみ順序づけることは，必ずしも意味付けが，実際的であるとは限らない。そのことは，推測目標によって決まるし，それに従って種々の順序づけを定義しなければならないのは前に述べたことである．併し乍ら，$\sigma_i=\sigma$ $(i=1,2,\cdots,k)$のときは，一応自然な考え方として受入れられるし，計算上の困難も幾分和らげられる．*R.E. Bechhofer* [1]は，この問題を実験計画への適用を念頭に手がけている．一元分類に関しては，$(n=n_i,\ i=1,\cdots,k)$判定方式として

$\bar{x}_{1(i_1)}\leqq\bar{x}_{2(i_2)}\leqq\cdots\leqq\bar{x}_{k(i_k)}$ なるとき，

Π_jを$\bar{x}_{i(j)}$ $(i=1,2,\cdots,k-s)$に対応するものを I-群へ

$\bar{x}_{i(j)}$ $(i=k-s+1,\cdots,k)$ に対応するものを，S-群へ分割する．

をとる問題を扱った。このとき正しい判定をする確率は，母平均が$\mu_1\leqq\mu_2\leqq\cdots\leqq\mu_{k-s}\leqq\mu_{k-s+1}\leqq\cdots\cdots\leqq\mu_k$であるから

$$P\{\max_{1\leqq i\leqq k-s}\{\bar{x}_i\}<\min_{k-s+1\leqq i\leqq k}\{\bar{x}_i\}\}$$

である．彼はこの近似的な見透しをうるために，先づ一番簡単な場合，即ち，今母集団が母平均に関して，二つの群にのみ区別されている特別な場合 H_δ について考えてみる．

$$H_\delta : \mu_1 = \mu_2 = \cdots\cdots = \mu_{k-s} = \mu$$
$$\mu_{k-s+1} = \cdots\cdots\cdots\cdots = \mu_k = \mu+\delta \qquad (\delta > 0)$$

すると"正しい判定をする事象" E_0 は，仮説 H_δ の下で次の確率をとる．

$$P\{E_0 \mid H_\delta\}$$
$$= P\{\underset{1\leqq i\leqq k-s}{Max}\{\bar{X}_i\} < \underset{k-s+1\leqq j\leqq k}{min}\{\bar{X}_j\} \mid H_\delta\}$$
$$= s\int_{-\infty}^{\infty}[\Phi(y+d)]^{k-s}[1-\Phi(y)]^{s-1}\Phi'(y)dy$$

但し，$\Phi(y) = \int_{-\infty}^{y}\Phi'(x)dx = \int_{-\infty}^{y}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx$

且つ，$d = \frac{\sqrt{n}\delta}{\sigma} = \sqrt{n}\,\lambda$.

次に一般の場合へ考察を深めて行く．一般の場合 $H_{\{\delta_i\}}$ を

$$H_{\{\delta_i\}} : \mu_{i+1} - \mu_i = \delta_i \qquad (i = 1, 2, \cdots, k-1)$$

とすると，$P(E_0 \mid H_{\{\delta_i\}})$は，(i) $\{\delta_i\}$に関して，単調増加函数であり，(ii) nを大きくすると1に近づく函数であることが判る．従って，$\{\delta_i\}$の下限を δ^* として，次の仮説

$$H_{\{\delta^*\}} : \mu_{i+1} - \mu_i = \delta^* \qquad (i = 1, 2, \cdots, k-1)$$

を考える．そうするとき，

$$P(E_0 \mid H_{\{\delta_i\}}) \geqq P(E_0 \mid H_{\{\delta^*\}})$$

なる関係が成立するので，今 $1-\alpha$ $(0<\alpha<1)$ より大なる信頼度で分割するためには，

$$P(E_0 \mid H_{\{\delta^*\}}) \geqq 1-\alpha$$

なるnを定めればよいのである．

このことは，若干の記号上の注意と適当な変換に注意すれば，二元分類の場合についても，ある程度類似をたどることが出来る．即ち，k個の母集団が $k = r \times l$ とし，仮定 $N(\mu_{ij}, \sigma^2)$を σ^2 が既知で $\mu_{ij} = \mu + \alpha_i + \beta_j$ $(i = 1, 2, \cdots, r;\ j = 1, 2, \cdots, l)$ とすると，この種の問題では，最大の平均効果を与えるiに関する要因若しくはjに関する要因の水準を$[r-s, s]$ 若しくは$[l-s, s]$に分割することになる．更には，全平均効果を最大にする水準の組合せ(i_M, j_M)を求める問題とも

なる．前者の場合は一元分類に帰着出来るので，後者について述べてみよう．この場合 π_{ij} の標本は $0_{n_{ij}}: x_{ij1}, x_{ij2}, \cdots, x_{ijn_{ij}}$ とし，$n_{ij}=n$ と考える．判定方式は

$$x_{1(i_1)\cdot} \leqq x_{2(i_2)\cdot} \leqq \cdots\cdots \leqq x_{r(i_r)\cdot}$$

$$x_{\cdot 1(j_1)} \leqq x_{\cdot 2(j_2)} \leqq \cdots\cdots \leqq x_{\cdot \ell(j_\ell)}$$

ならば π_{i_r, j_ℓ} を選ぶ，即ち (i_r, j_ℓ) 又は (i_M, j_M) を，全平均効果を最大ならしめる水準の組合せと判定する．

をとる．今特別の場合 H_{δ_1,δ_2} を，

$$H_{\delta_1,\delta_2}: \alpha_1=\alpha_2=\cdots=\alpha_{r-1}=\alpha,\quad \alpha_r=\alpha+\delta_1$$

$$\beta_1=\beta_2=\cdots=\beta_{\ell-1}=\beta,\quad \beta_\ell=\beta+\delta_2$$

とすると，この仮説の下で正しい判断をする確率は

$$P(E_0 \mid H_{\delta_1,\delta_2})$$

$$=P(\bar{X}_{r(i_r)\cdot}=\bar{X}_{r\cdot} \text{ 且つ } \bar{X}_{\cdot\ell(j_\ell)}=\bar{X}_{\cdot\ell} \mid H_{\delta_1,\delta_2})$$

こゝで $y_i=\bar{x}_{r\cdot}-\bar{x}_{i\cdot}$ $(i=1,2,\cdots,r-1)$

$z_j=\bar{x}_{\cdot\ell}-\bar{x}_{\cdot j}$ $(j=1,2,\cdots,\ell-1)$

とすると，$y_i \perp\!\!\!\perp z_j$ $(i=1,2,\cdots,r-1;\ j=1,2,\cdots,\ell-1)$ なる故

$$P(E_0 \mid H_{\delta_1,\delta_2})=P(y_i \geqq 0;\ i=1,2,\cdots,r-1)\cdot P(z_j \geqq 0;\ j=1,2,\cdots,\ell-1)$$

又 $d_1=\frac{\sqrt{n\ell}}{\sigma}\delta_1=\sqrt{n\ell}\,\lambda_1$，$d_2=\frac{\sqrt{nr}}{\sigma}\delta_2=\sqrt{nr}\,\lambda_2$

とすると

$$P(E_0 \mid H_{\delta_1,\delta_2})=\int_{-\infty}^{\infty}[\Phi(t+d_1)]^{r-1}\Phi'(t)dt\int_{-\infty}^{\infty}[\Phi(u+d_2)]^{\ell-1}\Phi'(u)du$$

である．

このようにして，交互作用を無視しうるとすれば，二元分類によって判定の効率がよくなることがあるのは云うまでもない．このことは，r 元分類（ラテン方格等も含めて）についても言えるのであるが，交互作用を問題にすれば，一元分類としてしか取扱うことは出来ない．更に，以上のことは $[k-1,\ 1]$ であるわけだが，これを一般に $[k-s,\ s]$ とすると，二元分類として扱う限り，複雑な形となってくる．i-要因の水準と j-要因の水準の適当な組合せをどのようにして，s 個選ぶかに困ってくるからである．

次に当然の道筋として，以上のような判定方式を逐次的方式として考えること

がある．*R.E. Behhofer and M. Sobel* [1] は　これらを同称な仮定の下で，母平均が最大なるものを求める問題についてとり上げた．このとき，母集団 $\{\Pi_i\}$ からの標本 $\{x_{i1}, \cdots, x_{im}\}$ の系列で m は次の判定方式に依存する確率変数となる点と，それぞれの m 回目の段階に於て，同一母集団 Π_i からのデータの和 $T_i^m = \sum_{j=1}^{m} x_{ij}$ の順序 $T_{(i)}^m$，即ち，$T_{1(i_1)}^n \leqq T_{2(i_2)}^m \leqq \cdots \leqq T_{k(i_k)}^m$ が m によって異った母集団 $\Pi_{(i_k)}$（こゝで i_k は $1 \leqq i_k \leqq k$ で m に依存する）に対応していることに注意すればよい．今 y_m を $y_m = T_{k(i_k)}^m - T_{k-1(i_{k-1})}^m$ $(i=1,2,\cdots)$ とおくと，判定方式は，

$y_m \geqq D$　ならば

$T_{k(i_k)}^m$ に対応する母集団 Π_{i_k} を選ぶ．

$y_m < D$　ならば

観測を続ける．

とする．今，$\delta_{k,i} = \mu_k - \mu_i$ $(i=1,2,\cdots,k-1)$ とする．このとき $\delta_{k,k-1}$ の可能な最小の値 $\delta^* > 0$ を仮定し，$\delta_{k,k-1} \geqq \delta^*$ の時に，正しい選択をする確率が $1-\beta$ $(\frac{1}{k} < 1-\beta < 1)$ となるように要求する．$C = (k-1)(1-\beta)/\beta$，$D = (\sigma^2/\delta^*) \log_e C$ とおいて前述の判定方式を用いると，$\delta_{k,k-1} \geqq \delta^*$ の時に正しい選択をする確率は $\geqq 1-\beta$ で，等号は $\delta_{k,i} = \delta^*$ $(i=1,\cdots,k-1)$ の時に成立つ．この判定方式を用いる時の正しい選択をする確率の $g.l.b.$ は，$\delta = \min \delta_{k,i}$ $(i=1,2,\cdots,k-1)$ であるような凡ての $\delta_{k,i}$ に対して，

$$L(\delta) = C^{\delta/\delta^*} \div \{C^{\delta/\delta^*} + (k-1)\}$$

によって与えられる．

今までの議論はすべて，共通分散 σ^2 が既知である場合を取りあげてきたのであるが，この σ^2 が未知である場合は，同じ判定方式は使えない．このとき丁度従来の仮説検定論に於ける *Student* の t と，類似の方向で解決しようとするのが，*R.E. Bechhofer, C.W. Dunnett, M. Sobel* [1]の方法である．併しながら，与えられた信頼度に於て問題を処理するためには，今まで述べてきた一回標本の手続では，取扱えないので，*C. Stein* [1]の二標本の手続を採用し，近似的に解こうというのである．問題は $[k-1, 1]$ の場合に制限してあることに注意する．今各 Π_i から n_0[註1]の大いさの才一回標本をとる．s_i^2 をそれぞれの Π_i の不偏分散推定とする．そうするとき $S_0^2 = \frac{1}{k}\sum_{i=1}^{k} s_i^2$ は $S_0^2(n_0-1)k/\sigma^2$ が自由度 $\nu = (n_0-1)k$ なる χ^2-分布に従うものである．このとき与えられた $k_\nu(>0)$ 及び $\delta^*(>0)$ に対し

（註1）一般に n_0 が n_i であっても，自由度に変化あるのみ．

て，$n=Max\{n_0, [\![2S_0^2(\frac{h_\nu}{\delta})^2]\!]\}$(註2) をとり，更に各$\pi_i$から大きさ$(n-n_0)$の第二回標本をとり，この二つの標本の合計から得られたπ_iの標本平均を$\bar{x}_i$とする．このときの判定方式は，σ^2が既知であるときと全く同様に，$\bar{x}_i$の順序づけによる最大標本平均$\bar{x}_{M(j)}$に対応するπ_jを選ぶのである．このとき，前述の仮説H'_δの下で正しい判定をなす確率は，

$$P(E_0 \mid H'_\delta)$$
$$= P\{\bar{X}_i < \bar{X}_k \ ; \ i=1,2,\cdots,k-1\}$$
$$= P\left\{\sqrt{\frac{nk}{2}}\ \frac{\bar{X}_i-\bar{X}_k+\delta}{S_0} < \sqrt{\frac{nk}{2}}\ \frac{\delta}{S_0} \ ; \ i=1,2,\cdots,k-1\right\}$$

で表わされる．こゝで，$\sqrt{\frac{nk}{2}}\frac{\delta}{S_0}=h$及び，$\sqrt{\frac{nk}{2}}\frac{\bar{X}_i-\bar{X}_k+\delta}{S_0}=t_i$ $(i=1,2,\cdots,k-1)$ とおいて，S_0について，$0<S_0<\infty$ で積分すると

$$P(E_0 \mid H'_\delta)$$
$$= K\int_{-\infty}^{h}\cdots\int_{-\infty}^{h}\left\{1+\frac{1}{n}\sum_{i=1}^{k-1}\sum_{j=1}^{k-1} b_{ij}\, t_i t_j\right\}^{-\frac{1}{2}(n+k-1)} dt_1\cdots dt_{k-1}$$

但し

$$K=\frac{\Gamma[\frac{1}{2}(n+k-1)]}{\sqrt{k}\,(\frac{1}{2}n\pi)^{\frac{1}{2}(k-1)}\Gamma(\frac{1}{2}n)}$$

$$b_{ij}=\begin{cases}2(k-1)/k, & i=j\\ -2/k, & i\neq j\end{cases}$$

となる．従って，この確率を指定された$1-\alpha$ $(\frac{1}{k}<\alpha<1)$ に等しくなるようにhを定めればよいわけであるが，そのためには，一般k次元t分布の表が必要である．特殊の場合には，*M. Merrington*[1], *C. Dunnett, M. Sobel*[1]が利用出来るのであるが，聞くところによると，*H. O. Hartley*が近き将来2次元F-分布及びn次元t-分布の製表計画を米国学会に提案する由であることをつけ加えておこう．更に，*C. Stein*[1]によると，二標本による平均全標本の大きさは，$kE(n)$で表わせば，

$$E(n)=n_0\, P\left\{\chi_\nu^2 < \frac{\nu n_0}{2k_\nu^2}\left(\frac{\delta}{\sigma}\right)^2\right\}$$
$$+2k_\nu^2\left(\frac{\sigma}{\delta}\right)^2 P\left\{\chi_{\nu+2}^2 > \frac{\nu n_0}{2k_\nu}\left(\frac{\delta}{\sigma}\right)^2\right\}$$

（註２）[\![ ]\!]は*Gauss*の記号を表わす．

$$+\theta P\left\{\chi_\nu^2 > \frac{\nu n_0}{2h_\nu^2}\left(\frac{\delta}{\sigma}\right)^2\right\}$$

但し　$0 \leqq \theta \leqq 1$

で表わされる．

［II］ 分散による順序づけ．　このときは平均値による場合と異なり，平均値μ_iが未知であることは，自由度に変化を生ずるのみで特別視する必要はない．だが$\mu_i=\mu$ $(i=1,2,\cdots,k)$ であるかどうかは，順序づけの内容の意味づけを左右するのに注意する必要がある．*R.E. Bechhofer, M. Sobel*[2]は分散に関する［$k-s$, s］の場合を扱った．その判定方式は

$$s^2_{1(i_1)} \leqq s^2_{2(i_2)} \leqq \cdots\cdots \leqq s^2_{k(i_k)}$$

なるとき，

Π_j の中 $s^2_{i(j)}$ $(i=1,2,\cdots,k-s)$ に対応するものをI-群へ

Π_j の中 $s^2_{i'(j)}$ $(i'=k-s+1,\cdots,k)$ に対応するものをS-群へ分割する．

である．このとき正しい判定を下す確率は，

$$P\left\{\underset{1\leqq i\leqq k-s}{Max}\{s_i^2\} < \underset{k-s+1\leqq j\leqq k}{min}\{s_j^2\}\right\}$$

であるが，平均値による場合と同様に特別な場合 H_θ を，

$$H_\theta : \sigma_1 = \sigma_2 = \cdots\cdots = \sigma_{k-s} = \sigma_1$$
$$\sigma_{k-s+1} = \cdots\cdots = \sigma_k = \sigma_2$$
$$\sigma_2^2/\sigma_1^2 = \theta \quad (1<\theta<\infty)$$

とすると，この仮定 H_θ の下で正しい判定をなす確率 $P(E_0|H_\theta)$ は，$n_i=n$とすると

$$P(E_0|H_\theta)$$
$$=s\int_0^\infty [K_\nu(\theta y)]^{k-1}[1-K_\nu(y)]^{k-s}K_\nu'(y)dy$$

但し $K_\nu(y)=\int_0^y K'(y)dy$

$$=\frac{1}{\Gamma(\frac{\nu}{2})}\left(\int_0^y\left(\frac{\chi^2}{2}\right)^{\frac{\nu}{2}-1}exp\left\{-\frac{\chi^2}{2}\right\}d\left\{\frac{\chi^2}{2}\right\}\right.$$

且つνは s_i^2 の共通なる自由度とする．

である．標本の大きさを一般に n_i とし，且つ一般の仮説 $H_{\{\theta_{ij}\}}$

$$H_{\{\theta_{ij}\}} : \theta_{ij} = \sigma_i^2/\sigma_j^2 \quad (i=1,2,\cdots,k;\ j\neq i \text{ 且つ } j=1,2,\cdots,k)$$

の下では，

$$P(E_0 \mid H_{\{\theta_{ij}\}}$$
$$= \sum_{j=1}^{s} \int_0^{\infty} \Big[\prod_{\substack{\alpha=1 \\ \alpha \neq j}}^{s} K_{\nu_\alpha}\Big(\frac{\nu_\alpha}{\nu_j}\theta_{j\alpha} y\Big)\Big]\Big[\prod_{\beta=s+1}^{k}\Big(1-K_{\nu_\beta}\Big(\frac{\nu_\beta}{\nu_j}\theta_{j\beta} y\Big)\Big)\Big]$$
$$\times K'_{\nu_j}(y)\,dy$$

こゝで ν_i は s_i の自由度とする．

である．この確率を得るためには，多次元 χ^2-分布を知ることが必要である．

以上，不充分ながら，s が確定値であるとき $[k-s, s]$ 分割に関連するものを述べてみた．次には序論にも述べたように，$[k-s, s]$ の場合，s が確定値であることが，実際上，理論上不合理である場合，云いかえると s が判定方式に依存する確率変数である場合に触れることにする．

§4. s が確率変数であるときの $[k-s, s]$ 分割に関連するもの　この種の特別な場合として，*E. Paulson* [3] の k-標本卓越問題（k *sample slippage problem*）がある．共通分散 σ^2 が既知である場合は，未知である場合の不偏分散推定 s_0^2 に対して，その自由度 ν が $\nu=\infty$ であるときを考えればよいので，以下述べるすべての場合について，σ^2 は未知であると仮定しよう．さて，問題の k 標本卓越問題は，確率変数 s がその実現値に 0 又は 1 をとる特別の場合である．

$[k, 0]$ のとき，$\mu_i = \mu$ $(i=1,2,\cdots,k)$ であるという判定 D_0 であり，$[k-1, 1]$ のとき，D_0 は成立たず，且つ $\bar{x}_{M(j)}$ に対応する π_j が卓越した母平均 $\mu_{M(j)}$ をもつという判定 D_j $(j=1,2,\cdots,k)$ をとるとする．従ってわれわれの判定は $\{D_0, D_1, \cdots, D_k\}$ の $k+1$ 通りの場合が起りうるわけだ．今 $\bar{x} = \frac{1}{k}\sum_1^k \bar{x}_i$，$S = \sum_1^k \sum_1^{n_i} (x_{ij}-\bar{x})^2$，$s_0^2 = \sum_1^k s_i^2/k$ とすると，判定方式は，適当な常数 λ_α に対して

(1°) $n(\bar{x}_{M(j)} - \bar{x})/S^{\frac{1}{2}} > \lambda_\alpha$ ならば D_j を，

(2°) $n(\bar{x}_{M(j)} - \bar{x})/S^{\frac{1}{2}} \leq \lambda_\alpha$ ならば D_0 を

判定する．

である．この場合，*E. Paulson* [3] は，従来の仮説検定に於ける仮説と対立仮説のような H_0 と $H_1(\Delta)$ とを考える．こゝで

$H_0 : \mu_1 = \mu_2 = \cdots\cdots = \mu_k$

$H_k(\Delta) : \mu_1 = \mu_2 = \cdots\cdots = \mu_{k-1} = \mu,\ \mu_k = \mu + \Delta$

今，仮説 H_0 の下で，間違った判定 (1°) の起る確率は，事象 A_i

$A_i : n(\bar{x}_i - \bar{x})/S^{\frac{1}{2}} > \lambda_\alpha \quad (i=1,2,\cdots,k)$

の中，少くとも一つが起る確率であるので，H_0 が真なるとき，誤って $D_j(j=1, 2, \cdots, k)$ なる判定をする確率を α 以下にする λ_α を近似的に $P(A_i \mid H_0)=\alpha/k$ によって求める．このとき λ_α は近似的に

$$\lambda_\alpha = \left(\frac{n(k-1)F_\nu^{-1}(\frac{2\alpha}{k})}{k(\nu+F_\nu^{-1}(\frac{2\alpha}{k}))} \right)^{\frac{1}{2}}, \quad \nu = kn-2$$

となる．又，

$$P(A_k \mid H_k) = \frac{1}{2} P\left\{ F(1, \nu; \tfrac{1}{2}\delta^2 > \frac{\nu\lambda_\alpha^2 k}{n(k-1)-k\lambda^2} \right\}$$

但し，$F(1, \nu; \delta)$は自由度$[1, \nu]$，偏心率δなる非心F-分布に従う確率変数であり，δは，$\delta = -\frac{n^{\frac{1}{2}}\Delta}{\sigma}\left(\frac{k-1}{k}\right)^{\frac{1}{2}}$であるとする．

従って，H_0 が真のとき (i) の起る確率をα以下，H_1 が真のとき (ii) の起る確率をβ以下にするためには，近似的に

$$\alpha = kP(A_k \mid H_0) = \frac{k}{2} P\left\{ F(1, \nu; 0) > \frac{\nu\lambda_\alpha^2 k}{n(k-1)-k\lambda^2} \right\}$$

$$\beta = 1 - P(A_k \mid H_0) = \frac{1}{2} P\left\{ F(1, \nu; \frac{\delta^2}{2}) > \frac{\nu\lambda_\alpha^2 k}{n(k-1)-k\lambda^2} \right\}$$

なる二條件から

$$n = \frac{k}{k-1}\left(\frac{\sigma}{\Delta}\right)^2 \left(\zeta_{kn-2}(\beta; \frac{\alpha}{k})\right)^2$$

が得られる．（北川[1]参照）

但し，$\zeta_\nu(\beta; \alpha)$ は $P\{t(\nu; 0) > t_\nu(2\alpha)\} = \alpha$，$P\{t(\nu; \delta) > t_\nu(2\alpha)\} = \beta$ を満す δ を表わし，$t(\nu; \delta)$ は自由度ν，偏位δの非心t-分布に従うとする。

更に $E.\ Paulson$[1]は，$[k-s, s]$ なる形でsが 0,1 の実現値をとる場合のみでなく，$0 \leqq s \leqq k$ なる場合について述べている。この場合の判定方式は，

$\sqrt{n}(\bar{x}_{M(j)} - \bar{x}_{j(ij)})/s_0 > \lambda_\alpha$ ならば

$\bar{x}_{j(ij)}$ に対応する π_{ij} をS-群へ

$\sqrt{n}(\bar{x}_{M(j)} - \bar{x}_{j'(ij')})/s_0 \leqq \lambda_\alpha$ ならば

$\bar{x}_{j'(ij')}$ に対応する $\pi_{ij'}$ をI-群へ分割する。

をとる．このとき，H_0 の下で誤った分割をする事象 E_1 の起る確率は，

$$P(E_1 \mid H_0) = 1 - P\left(w < \frac{\lambda_\alpha s_0}{\sqrt{n}} \mid H_0\right),$$

但し　$w=\bar{x}max.-\bar{x}min.$

H_1の下で誤った分割をする事象 E_2 の起る確率は，

$$P(E_2|H_k)=1-C\int_0^{\infty}\int_{-\infty}^{\infty}w^{\nu-1}e^{-\frac{(\nu w^2+y^2)}{2}}\left[\Phi\left(y+\frac{\Delta}{\sigma}\sqrt{n}-\lambda_\alpha w\right)\right]^{k-1}\cdot dy\,dw$$

である．但し，ν は s_0^2 の自由度であり，

$$C=\frac{(\nu)^{\frac{\nu}{2}}}{\sqrt{2\pi}\,2^{\frac{\nu-2}{2}}\Gamma(\frac{\nu}{2})}$$

である．$P(w<\frac{\lambda_\alpha s_0}{\sqrt{n}})$は *Pearson, Hartly* [1], [2] の *studentized range* の表から $P(E_1|H_0)=\alpha$ なる λ_α を求めることが出来るが，$P(E_2|H_k)=\beta$ なるnを同時に定めることは，表がないので実際には求めることができない(註3)．尚，この場合は，H_0 と H_k の二つの場合についてのみ考えることは必ずしも妥当ではない。というのは，前記k標本卓越問題と違って，この判定は(2^n-1)個の場合がありうるからである．従って H_k の代りに，多くの対立仮説に類するものを考えねばならぬわけである．*E. Paulson* [1] は，特別に σ^2 が既知の場合，S-群が(k)番目の集団（標本平均が *Max.* なる母集団）と，t個より少い(k)番目の母集団以外の集団からなる確率，即ち$[k-s,s]$が$s-1\leqq t$（t: 確定値）判定となるすべての場合が起る確率 $1-P_t^*$ を

$$1-P_t^*=\sum_{\alpha=0}^{t}\binom{k-1}{t}\left[T_{1\alpha}+\alpha T_{2\alpha}\right]$$

但し

$$T_{1\alpha}=\int_{-\infty}^{\infty}\left[\Phi\left(y+\frac{\Delta}{\sigma}\sqrt{n}-\lambda\right)\right]^{k-\alpha-1}\left[\Phi\left(y+\frac{\Delta}{\sigma}\sqrt{n}\right)-\Phi\left(y+\frac{\Delta}{\sigma}\sqrt{n}-\lambda\right)\right]^{\alpha}\Phi'(y)\,dy$$

$$T_{2\alpha}=\int_{-\infty}^{\infty}\left[\Phi(y-\lambda)\right]^{k-\alpha-1}\left[\Phi(y)-\Phi(y-\lambda)\right]^{\alpha-1}\cdot\left[\Phi\left(y-\frac{\Delta}{\sigma}\sqrt{n}\right)-\Phi\left(y-\frac{\Delta}{\sigma}\sqrt{n}-\lambda\right)\right]\cdot\Phi'(y)\,dy$$

（註3）　近似的には以下の方式の場合と同様によれば求めうるのであるが，λ_α は正確に求めうる．

を求めた．

今までこの節で述べたことは，k個の母集団の中から母平均が卓越していると思われる母集団を一個又は若干個とるか，或いは，すべてを同一とみるかの問題であったわけであるが，次はこのk個の母集団の中に，一つの対照母集団，例えばπ_1があるとしよう．このことはπ_1が従来から使用されている処理に対応するものであり，他の$(k-1)$個は新に採用すべきかどうかが，問題になっている処理に対応するものと考えればよい．こういうとき，対照母集団を基準にして，これより優れたものがあるときはそれを採用し，さもないときは，従来のもの即ち，対照集団を採用するという種類の判定を行う場合である．*E. Paulson* [2] は，この場合についても，平均値による順序づけをもって$[k-1, 1]$の問題(註4) を手がけている．彼の判定方式は，今 $\bar{x}^* = Max.\{\bar{x}_2, \bar{x}_3, \cdots, \bar{x}_k\}$ とし，$\bar{x}_1$ を対照集団π_1に対応する標本平均とすると

(i)　$\bar{x}^* - \bar{x}_1 > \lambda_\alpha s_0 \left(\frac{1}{n} + \frac{1}{n_1}\right)^{\frac{1}{2}}$　ならば

　　$\bar{x}^*$ に対応する π^* を

(ii)　$\bar{x}^* - \bar{x}_1 \leqq \lambda_\alpha s_0 \left(\frac{1}{n} + \frac{1}{n_1}\right)^{\frac{1}{2}}$ ならば

　　$\bar{x}_1$ に対応する π_1 を選ぶ．

である．但し $n_2 = n_3 = \cdots = n_k = n$，且つ $n_1 = n/k$ とし，$\nu s_0^2/\sigma^2$ は自由度νの χ^2-分布に従うものである．これも全く同様にして，H_0 のとき(ii)の起る確率は，s_0/σ が近似的に$N(1, \frac{1}{\nu})$ であることと，*Bonferroni* の不等式を用いた次の関係

$$1-(k-1)P\{\bar{x}_2 - \bar{x}_1 > \lambda_\alpha s_0 (n^{-1} + n_1^{-1})^{\frac{1}{2}}\}$$
$$\fallingdotseq P\{\bar{x}^* - \bar{x}_1 \leqq \lambda_\alpha s_0 (n^{-1} + n_1^{-1})^{\frac{1}{2}}\} \fallingdotseq 1-(k-1)P(t(\nu; 0) > \lambda)$$

から近似的に求めることができる．従って，この確率をαにするλ_αは近似的に

$$\lambda_\alpha = t_\nu\left(\frac{2\alpha}{k-1}\right)$$

であり，同様にして，H_k の下での(i)の起る確率は，近似的に

$$P\{\bar{x}_k - \bar{x}_1 \geqq \lambda s_0 (n^{-1} + n_1^{-1})^{\frac{1}{2}} \mid H_k\}$$

であるので，自由度ν，偏差$\delta = (\Delta/\sigma) n^{\frac{1}{2}} (1+k)^{-\frac{1}{2}}$ なる非心t-分布から，この確率

(註4)　この場合は，§3に入れるべきであるとも考えられるのであるが，方法が類似していることと，対照母集団を考慮しなければ，$[k-1-s, s]$となり，やはり，sが0, 1をとる確率変数であるといえることを注意しておく．

をβ以下にする　$\frac{1}{n^*}=\frac{1}{n_1}+\frac{1}{n}$ なる n^* を

$$n^*=\left(\frac{\sigma}{\Delta}\right)^2\left(\zeta_\nu\left(\beta\,;\,\frac{\alpha}{k-1}\right)\right)^2$$

から求めることができる.

彼は更に，二項分布の場合についても同じ問題を取扱ったが，これは$\bar{p}$が$y=\sin^{-1}\sqrt{\bar{p}}$ なる変換をすると近似的に $N(\sin^{-1}\sqrt{p},\ 1/4n)$ に従うことを用いているので，本質的に前記のものと変りはない．又前記の場合，$k=2$ とすると，H_0 と H_k 及び α，β は全く *Neyman-Pearson* の仮説検定論の仮説と対立仮説及び第１種と第２種の過誤を犯す確率に対応するものであることに注意する必要がある．σ^2 が既知である場合については，s_0 の代りにσを代入すると共に，s_0^2 の自由度νを $\nu=+\infty$ として考えればよいことは前の断り書きの通りである．従って，この場合はこの節の全頁を通じて特別の場合として出て来るものである．

以上 §3 と §4 で，具体的な判定方式について述べたが，これらの判定方式が，最適なものであるかどうかとか，同一問題に於ける種々の判定方式の優劣については触れてはいなかった。次の節でこの点に触れた論文を挙げることにしよう．

§5．最適な判定方式をめぐって．

今まで列挙してきた種々の問題に対する判定方式は，単に常識的で，もっともらしいものの試みを定式化して来たものであった。この節では，これらの方式の中或るものは，理論的見地によって裏付けうるものである．このような理論的裏付けをなそうという試みが一，二の論文によってなされているのだが，先づk-標本卓越問題を取上げて来た *E. Paulson* [3] の場合を挙げよう．彼は結果として既に述べたk-標本卓越問題に於ける判定方式が，ある種の條件の下で，H_0 のとき正しい判定 D_0 が選ばれる確率を $1-\alpha\ (\alpha>0)$ に固定するとき，$H_j(\Delta)\ (j=1,2,\cdots,k)$のとき正しい判定をする確率を最大にするものであることを証明している．彼の條件とゆうのは，判定方式は (a) 各観測値にある常数を加えても，正の常数をかけても変らないこと．(*invariant*). (b) どの母集団 π_i の母平均が右にすべっていても，そのずれΔが同じであれば判定方式は，すべてのiに対して正確な判定をする確率が等しいとゆう意味で対称であることである．この條件 (a) と $(\bar{x}_1,\bar{x}_2,\cdots,\bar{x}_k\,;\,s_0^2)$ が $(\mu_1,\mu_2,\cdots,\mu_k\,;\,\sigma^2)$ の充足統計量であることから，問題の判定方式が，$w_\alpha=\frac{(\bar{x}_\alpha-\bar{x}_k)}{s_0}$ $(\alpha=1,2,\cdots,k-1)$に依存していると考えても一般性を失わない．従って今，判定の組 $\{\bar{D}_0,\bar{D}_1,\cdots,\bar{D}_k\}$

$\bar{D}_0 : H_0' : a_1 = a_2 = \cdots\cdots = a_{k-1} = 0$ をとる判定

$\bar{D}_j : H_j'(\frac{\Delta}{\sigma}) : a_1 = \cdots\cdots = a_{j-1} = a_{j+1} = \cdots\cdots = a_{k-1},\ a_j = \frac{\Delta}{\sigma}$

$j = 1, 2, \cdots, k-1$

$\bar{D}_k : H_k'(-\frac{\Delta}{\sigma}) : a_1 = \cdots\cdots = a_{k-1} = -\frac{\Delta}{\sigma}$

但し $a_\alpha = (\mu_\alpha - \mu_k)/\sigma \quad (\alpha = 1, 2, \cdots, k-1)$

の中から一つの正しい判定を選ぶ方式を考える．これらは判定の組$\{D_0, D_1, \cdots, D_k\}$に対応するものであるので，前提によって，$H_0'$ のとき $\bar{D}_0$ を選ぶ確率は $1-\alpha$ でなければならず，條件 (b) によって，$H_j'(\Delta/\sigma)$ のとき，$\bar{D}_j$ を選ぶ確率はすべての j に接して等しくなければならない．今こゝで $H_j'(\frac{\Delta}{\sigma})$ の先験確率——これはΔとσの函数であるのだが——を $g_j(>0)\ (j=0, 1, \cdots, k)$ としよう．そして又先に述べた *Paulson* [3] の k-標本卓越問題に於ける判定方式を今のように，$\{\bar{D}_0, \bar{D}_1, \cdots, \bar{D}_k\}$ の中から一つを選ぶ方式に移したとする。このときこの判定方式が，組$\{\bar{D}_0, \bar{D}_1, \cdots, \bar{D}_k\}$の中で正しい判定を下す確率を最大にするものであるそのような先験確率 $\{g_j\}$ が存在することを *Paulson* [3] は証明したのである．このことが示されれば，先に述べられた判定方式

$n(\bar{x}_{M(j)} - \bar{x})/S^{\frac{1}{2}} > \lambda_\alpha$ ならば D_j

$n(\bar{x}_{M(j)} - \bar{x})/S^{\frac{1}{2}} \leqq \lambda_\alpha$ ならば D_0

をとる．

が，H_0 のとき，D_0 をとる確率を $1-\alpha$ としたとき，最適なものであることが判る．

これと少し違った観点から問題の究明にあたつたものに，*Bahadur* [1] がある．今k個の母集団 Π_i が正規分布$G(x, \theta_i)$によって特徴づけられているものとする．但し，θ_i は未知母平均 μ_i と未知共通分散 σ^2 を表わすものとしよう．彼はk個の母集団から，最大平均をもつ母集団を選ぶという問題を次のように定式化している．”選定“ Δを$\Delta = (p_1, p_2, \cdots, p_k)$ 但し，$p_i \geqq 0, \sum_{i=1}^{k} p_i = 1$ と定義し，$G(x|\Delta) = \sum_{i=1}^{k} p_i G(x, \theta_i)$ とおく．このことは大ざっぱに説明すると，母集団 Π_i に p_i とゆう重みを附けた混合母集団を構成することになる．最大の母平均をもつ母集団を選ぶある種の判定方式を採用して，k個の母集団を固定して考え，これを繰返して，卓越していると判定した母集団を重ね合して行くとこの混合母集団に到達するわけである．われわれが判定方式をうまく選べば選ぶ程，この混合母集団の平均はk個の母集団の最大母平均に近づくわけであるから，このような判定方式を探すことは，この”選定“ Δを求めることになるわけである．今，

統計的判定函数或いは方式dが，標本の函数であるところの"選定"sを値域とする函数であるとしよう．逆に，dに対応する判定s, $s(d)=(p_1(d),\cdots,p_k(d))$は一般には確率ベクトルであるから，任意に固定されたxに対して，$G(x|s(d))$は又確率変数である．従って，分布函数

$$H(x|d)=E[G(x|s(d))]=\sum_1^k G(x,\theta_i)E[p_i(d)]$$

は，dを用いたときの平均効果を意味している．今$\alpha(d)$を

$$\alpha(d)\equiv\int_{-\infty}^{\infty}x\,dH(x|d)$$

とおくと，$\alpha(d)$は分布函数$H(x|d)$の平均値を表わしているので，$\alpha(d)$を最大ならしめるdを組立てるのが，当面の問題に外ならない．このように問題が定式化されると，次は，この問題を解くために，損失函数として

$$W(\omega,s)=\underset{1\leq i\leq k}{Max.}\{\mu_i\}-\sum_1^k p_i\mu_i$$

を採用する．但し　$\omega=(\theta_1,\theta_2,\cdots,\theta_k)$である．

更に，採用する判定方式を公平判定方式(*impartial decision rule*)の組に制限する。今$X_i=\phi(X_{i1},\cdots,X_{in})$ $(i=1,2,\cdots,k)$; $Y=\psi(\{X_{ij}\})$とおき，$\{X_i\}\perp\!\!\!\perp Y$としよう．又，これらの統計量に基く判定方式$d=d(\{X_i\};Y)$が，$\sum_{j=1}^k\lambda_j=1$なる$\lambda_j(X_{(1)},\cdots,X_{(k)})$を定義し，$\lambda_j$は$d$によって，$X_{(j)}$に対応する母集団$\Pi_i$に割当てられる割合$p_i(d)$を表わすとする．このとき，このような判定函数を公平判定方式と呼ぶのである．現在の問題では，

$$X_i\equiv\overline{X}_i=\frac{1}{n}\sum_{j=1}^n X_{ij}$$

$$Y\equiv S^2=\sum_1^k\sum_1^n(X_{ij}-\overline{X}_i)^2/k(n-1)$$

即ち，$\{\mu_i\}$とσ^2の不偏推定がこれに当る．この性質が"公平"(*impartial*)と呼ばれるのは，どの$X_{(i)}$がどの母集団に属するかに拘りなく，$[\lambda_1,\cdots,\lambda_k]$を決定するからである．この概念は先に述べた*E. Paulson*[3]の対称の概念と同様なものである．結果として，彼は，上述のような損失函数をとるとき『判定方式d_k("常に最大標本平均をもつ母集団のみを選ぶ")をとるのか，$\{X_i\}$; Yに基く，公平判定方式の組(D)の中で一様に最適である』ことを証明した。このことは云いかえると，$d\in D$に対して$r(d|\omega)=E[W(\omega,s(d))|\omega]$とおくとき，すべての$\omega\in\Omega$に対して，$r(d_k|\omega)=\underset{d\in D}{inf}\,r(d|\omega)$であることを意味している．

この特別の場合 $k=2$ では，同じ観点からの *Bahadur-Robins*[1]の面白い研究がある．

§6. 結　言

以上述べたいくつかの問題は，それぞれに相当な実際上の重要性をもつものであるが，それらの解決のために提案されている手法はいずれも常識の域を出ていない．また取上げられた問題も，思いつくまゝにあれこれと列挙された感があり，それらを統一的に扱う一般論は，未だ殆ど与えられていないと云えよう。

冒頭にも述べたように，これらの問題はいずれも，古典的な推測理論——仮説検定と推定の理論——の枠の中では処理できない新しい分野として，"多重"推測乃至は決定（*multiple inference, multiple decision*）の問題として最近注目されて居り，現代の統計推測理論の発展にとって，試金石的な重要性をもつものと考えられる．

例へば，*A. Wald* は，40年代に統計的判定函数論を提唱したが，その目的の一つは，こうした方面をも，推測理論の対象に組み入れることであった。（*A. Wald*[1]序文参照）。この観点から，k-母集団の比較の問題を取扱ったのが *Bahadur* である．

Wald の理論の中では，こゝに取上げた問題の殆どは，「なしうる決定の数が有限個である」という，特別な場合に含まれる．すなはち，それらは，仮説検定の問題と著しい共通性をもち，その拡張であると考へられるのである．このように考へれば，上記の問題や手法の大部分が，初期のものにおいて特に，仮説検定論との類似点を手がかりとして構成されていることも，当然といえるであらう。

例へば，k 標本卓越問題の定式化においては，すべての母平均が等しい場合を帰無仮説の如く，また何れか一つの母平均だけが他より大きいという場合を対立仮説の如く扱い，定評ある t-検定に模した手法によって問題を処理している．それならば，こうして作られる手法は，この新しい問題に対してもまた，同様な有効性を保つであろうかというと，それは必ずしもたやすく証明できない。*Paulson* によると，彼の手法は，

$$\frac{\bar{x}_i - \bar{x}_k}{\sqrt{\sum_{i=1}^{k}\sum_{\alpha=1}^{n}(x_{i\alpha}-\bar{x}_i)^2}} \qquad (i=1,\cdots,k-1) \qquad \cdots\cdots\cdots\cdots \quad (1)$$

という統計量にのみ依存する，i について対称な手法の中では，最適なものであることが証明できる．（こゝでいう最適とは，上記の帰無仮説に対する，一定の

有意水準のもとでの最有力検定に似たものである.)。*Paulson*[3]はさらに，彼の方法が，この問題を不変にする標本空間の変換について不変性をもつことをあげている．こゝでもしもこうした変換全体に対する最大不変統計量（*Maximal Invariant Statistic*）を考えると，それは

$$\bar{x}_i-\bar{x} : \sqrt{\sum_{\alpha=1}^{n}(x_{i\alpha}-\bar{x}_i)^2} \ : \sqrt{\sum_{i=1}^{k}\sum_{\alpha=1}^{n}(x_{i\alpha}-\bar{x})} \qquad i=1,\cdots,n. \cdots\cdots(2)$$

となり，(1)よりは複雑になる．しかし(1)と(2)との分布状態を比較すれば，*Paulson* の手法は不変性をもつ手法全体のうちでも，相当に良いものであることがわかる．

Paulson が不変性に言及したことは，今後のこの方面への理論の発展に，相当に有力な示唆を与えるものと考えられる．この性質は，すでに仮説検定論においても，検定法に望ましい基準として取上げられ，理論的にも非常に重要な役割を演じている．同様に仮説検定論における他の諸々の概念が，こうした場面にいかに拡張されうるかということは，一つの大きな問題であらう．例へば，k－標本卓越問題は，仮説検定論との類似性を最も強く留めた構造をもっているが，すでにこゝでは第一種及び第二種の過誤といった基本的な概念も，全く変ったものになっている．すなわち，*Paulson* の記号に従へば，$\bar{D}_0$ をとるべきときに $\bar{D}_j$ をとる誤り，$\bar{D}_j$ をとるべきときに $\bar{D}_0$ をとる誤りの他に，$\bar{D}_j$ をとるべきときに他の $\bar{D}_i$ をとるという，第三種の過誤のようなものがあらわれるのである．（*Mosteller*[1]参照）．

このような概念が，新しい問題にどこまで適用しうるか，またいかなる新しい概念が必要であるかを検討することは，当面の基礎的な課題であって，それによって常識的な思いつきによって生れるあれこれの手法を理論的に吟味されることにならう．

こうした観点から"*Ranking*" という名を冠せられた最近のいくつかの労作を検討したと思われるのが，*Seal*[1]である．

ここでは，共通の未知分散をもつ数個の正規母集団から，同じ大きさの標本を各一回抽出し，母平均の大であると思われるいくつかのものを一つのグループとして取出すことが問題である．ここでグループに入る母集団の個数は，信頼度と標本から定まる確率変数であることが特徴である。

このように定式化したのち，*Seal* は最適な判定法のそなえるべき性質として，次の三つをあげている．先づ仮説検定論との類似性を辿って

「平均値最大なる母集団を，このグループに入れないこと」「その他の母集団を，このグループに入れること」を第一種，第二種の過誤と同様に扱い，

① 前者の確率を一定におさえ（有意水準）

② 後者の確率が前者のそれ以上となる（不偏性）

ことを求める．さらに"*Property of Gradation*"という基準を新たに設け，

③ 母集団平均値が大きいほど，上のグループに入りやすい．

という様にする．

Seal は，順序統計量を巧みに用いて，上の三條件を充す判定法の一つのクラスを見出し，そのクラスの中で，「最適」なものとして

④ *Paulson* の問題と同じ状況のもとでは，最良の母集団を上のグループに取出す確率を最大とするもの

を求めている（但し，標本数が相当に大きいことが必要とされる）。こうしてできた手法は，勿論非常に強い條件のもとではじめて「最適」なのであるが，そこに用いる統計量は，t－統計量とよく似たものであることは，注目に値するといえよう。

[引　用　文　献]

R. R. Bahadur [1] *On a problem in the theory of k-populations, Ann. Math. Stat.*, 21 (1950).

R. R. Bahadur and H. Robins [1] *The problem of the greater mean, Ann. Math. Stat.* 21 (1950).

R. E. Bechhofer [1] *A single-sample multiple decision procedure for ranking means of normal populations with known variance, Ann. Math. Stat.*, 25 (1954).

R. E. Bechhofer and M. Sobel [1] *A sequential multiple decision procedure for ranking means of normal populations with known variances (Preliminary Report), abstract, Ann. Math. Stat.*, 24 (1953).

R. E. Bechhofer and M. Sobel [2] *A single-sample multiple decision procedure for ranking variances of normal population, Ann. Math. Stat.*, 25 (1954).

R. E. Bechhofer, C. W. Dunnett and M. Sobel [1] *A two-sample mul-*

tiple decision procedure for ranking means of normal populations with a common unknown variances, Biometrika, 41 (1954).

C. Dunnett and M. Sobel [1] A bivariate generalization of student's t-distribution with tables for certain special cases, Biometrika, 41 (1954).

北川敏男 [1] 実験計画法講義, 才I巻, 培風館, 1955.

M. Merrington [1] Tables of percentage points of the t-distribution, Biometrika, 32 (1942).

F. Mosteller [1] A k-sample slippage test for an extreme population, Ann. Math. Stat., 19 (1948).

E. Paulson [1] A multiple decision procedure for certain problems in the analysis of variance, Ann. Math. Stat., 20 (1949).

E. Paulson [2] On the comparison of several experimental categories with a control, Ann. Math. Stat., 23 (1952).

E. Paulson [3] An optimum solution to the k-sample slippage problem for the normal distribution, Ann. Math. Stat., 23 (1952).

E. S. Pearson and H. O. Hartley [1] Tables of the probability integral of the range in samples from a normal population, Biometrika, 32 (1941-42).

E. S. Pearson and H. O. Hartley [2] Tables of the probability integral of the studentized range, Biometrika, 33 (1943).

H. Scheffé [1] A method for judging all contrasts in the analysis of variance, Biometrika, 40 (1953).

K. C. Seal [1] On a class of decision procedures for ranking meàns of normal populations, Ann. Math. Stat., 26 (1955).

C. Stein [1] A two sample test for a linear hypothesis whose power is independent of the variance, Ann. Math. Stat., 16 (1945).

J. W. Tukey [1] Quick and dirty methods in statistics, Part II, Simple analysis for standard designs, Proc. Fifth Annual Convention, Amer. Soc. for Quality Control.

A. Wald [1] Statistical decision functions, Jhon Wiley and Sons, 1950.