

ブロック配置における無作為化に就て

小川潤次郎

(1970年7月受付)

On The Randomization of Block Designs

Junjiro Ogawa

The author believes that he has been able to present a fairly satisfactory treatment of the randomization of block designs in this article. In order to get to the heart of the problem as quickly as possible avoiding the mathematical complexities, the simplest case i.e., the null-distribution of the F-statistic appearing in the analysis of variance of the observations obtainable by a randomized block design or a complete block design is considered here. The conclusion is that under certain uniformity conditions imposed on the unit-errors, the null-distribution of the F-statistic even under the Neyman model is asymptotically equivalent to the central F-distribution in the sense of the type $(M)_d$ as the number of blocks b goes to infinity, provided that the allocation of the treatments to the experimental units within each block is randomized.

The most general situation in this frame of research is the derivation of the non-central F-distribution which is asymptotically equivalent in the sense of the type $(M)_d$ as b goes to infinity to the power function of the F-statistic for testing a partial null-hypothesis in the analysis of variance of the observations obtainable from a randomized PBIB design with m associate classes under the Neyman model and this has been worked out in other articles.

University of Calgary, Canada

はしがき

実験によって処理効果を比較しようという場合、その基礎におく設定又はその設定を数学的に表現すべき模型（モデル）について考えよう。

或薬剤の効果を知りたい為に、この薬を取るであろう患者の母集団から $m+n$ 人の患者が無作為に抽出され、その内 n 人に此薬剤が投与され、残り m 人は対照として用いられるとする。ところでこの薬剤によって影響される性質について、特性質を観察して X_1, \dots, X_m ; Y_1, \dots, Y_n を得るものとする。すなわち、 X_1, \dots, X_m は対照につかわれた患者についての観察値であり、 Y_1, \dots, Y_n は薬剤を授与された患者についての観察値である。ところで、この観察値は二つの互に独立な成分 U_i と V_i の和と考えられる。 U_i は i 番目の患者のもつ個体値で全母集団では $N(\mu, \sigma_1^2)$ に従うものとし、 V_i は処理による効果であって、薬剤を投与されたものについては $N(\xi, \sigma^2)$ 、対照については $N(\eta, \sigma^2)$ に従うものと考える。このときは X 's 及び Y 's は互に独立な正規分布に従い、共通分散 $\sigma^2 + \sigma_1^2$ で平均は夫々

$$E(X) = \mu + \xi, E(Y) = \mu + \eta$$

であるから、薬剤の効果なしという仮説 $H: \xi = \eta$ を検定するには、

$$t(X, Y) = \frac{(\bar{Y} - \bar{X}) / \sqrt{\frac{1}{m} + \frac{1}{n}}}{\sqrt{[\sum (X_i - \bar{X})^2 + \sum (Y_i - \bar{Y})^2] / (m + n - 2)}}$$

によれば良いのである。

ところで実際の実験の場合の実験単位は或る病院に入院中の患者であるとか、囚人中からの志願者であるとかであって、決して患者の母集団からの無作為標本ではない。このときには上述の U_i は確率変数ではなく、定数又は母数であるとするのが至当なのである。この場合には母分布は

$$\frac{1}{(\sqrt{2\pi}\sigma)^n} \exp \left[-\frac{1}{2\sigma^2} \left(\sum_{i=1}^m (x_i - u_i - \xi)^2 + \sum_{j=1}^n (y_j - u_{m+j} - \eta)^2 \right) \right]$$

となる。若し無作為化をしなければ u_{m+j} が大きい為に $\xi = \eta$ であっても、 $\bar{Y} > \bar{X}$ ということが起り得る。

そこで無作為化をするのだが、それには $N=m+n$ のアンブルを $N!$ 通りに順列をつくって、各順列が等確率 $1/N!$ で N 人の患者に施されるようにする。

$$\sigma = \begin{pmatrix} 1 & 2 \cdots & N \\ \sigma(1) & \sigma(2) \cdots & \sigma(N) \end{pmatrix}$$

を一つの順列、その全体のなす対称群を S_N とかくことにすれば、この無作為化に対応する母分布は

$$\frac{1}{N!} \sum_{\sigma \in S_N} \frac{1}{(\sqrt{2\pi}\sigma)^N} \exp \left[-\frac{1}{2\sigma^2} \left(\sum_{i=1}^m (x_i - u_{\sigma(i)} - \xi)^2 + \sum_{j=1}^n (y_j - u_{\sigma(m+j)} - \eta)^2 \right) \right]$$

となる。ところでここで吾々は仮説 $\eta = \xi$ を対立仮説 $\eta > \xi$ に対して検定したいのであるが、若し u_i の間の変動が大きいならば X_i, Y_i にもとづく検定はその検定力が小さくなってしまう。そこで成る可く個体差の小さいものを集めてブロックを作るのが常識である。例えは、動物実験でなら、同腹のものを 1 ブロックにとるというのが常識である。

ところでブロック分けしても、ブロック内部での個体差というものは、依然として無視出来ない程残る場合がしばしばある。それだからこそ R. A. Fisher によって提唱された無作為化 (1926) が重大なのである。

ブロック配置における無作為化

Fisher による無作為化の数学的取扱いは B. L. Welch (1937), E. J. Pitman (1937) 及び M. B. Wilk (1955) によってなされている。彼等の研究で採用している模型は単位誤差（個体効果のこと）のみ考慮して、技術誤差は全く欠いた線形模型であって、確率は無作為化の操作を通じて持込まれる訳である。吾々はこれを“Fisher 模型”と呼んでおく。J. Neyman et al (1935) は考察する問題そのものの本性上単位誤差のみならず、更に技術誤差をも同時に考慮した方がより適切である場合のあることを指摘した。我々はこのようなものを“Neyman 模型”と呼ぶことにする。1939 年に M. D. McCarthy は Neyman 模型にもとづいて乱塊法の場合の分散分析の F 検定について調べた。

今から約 25 年程前に著者は近代統計学の勉強を始めたのであったが、その当初から、このブロック配置の無作為化ということは、何となく後味の悪いものであった。というのは、大抵の統計学の書物は実験計画の部の初頭に無作為化の大切なことを強調し、説得力があるのだが、かくして得られた標本を分析する分散分析の段になると、無作為化はどこかに行ってしまって、単純な正規回帰理論になってしまった。そして、無作為化という操作によって、単位誤差は無視しても良いという証明はない。M. D. McCarthy の前述の仕事はこの問題に対する最初の試みではあったが、本人も認めているように、無作為化の数学的取扱いが不充分であった。H. B. Mann はその著書 Analysis and Design of Experiments, Dover Publication Inc 1949 の 76 頁で次のように述べている。“此配置一乱塊法のことへの数学的に厳密な取扱いは未だ出来ていない。処理効果に対する近似的検定法としては、(ブロックを行に、処理を列にとって)これを二元配置と考えて、各行内での単位誤差を無視するやり方がある”。しかし、ブロック内の単位誤差を無視して良いなら、無作為化をやかましく云う必要はない。単位誤差と技術誤差を共に含む線形 Neyman 模型に立脚して、分散分析検定の標本分布導出の過程に無作為化ということが本質的に勘考された上で、何等かの意味で、通常の正規回帰模型からの分布が近似的に有効であると云えることが望ましいのである。

兎に角も、このようなやり方が過去何 10 年も相当広範に実験計画の分析に用いられてきて、

それ程致命的な不都合がなかったらしい以上、これはきっと正当化できるであろうという見込みをもった著者は、1959年頃から此問題に取組み、1965年以降は若い有能な統計学者池田貞雄博士の協力を得て、今や大略所期の目標に到達し得たように思われる。

問題の要点を明確にする為に、最も簡単な乱塊法の場合について述べることにする。

大きさ k のブロックが b 個与えられたものとする。従って全体で $n=bk$ 個の実験単位がある訳だが、これに通し番号を次のようにつける。 p 番目ブロックの i 番目の実験単位の番号は $f = (p-1)k+i$ とする。実験によって比較したい処理の数は $v=k$ である。ここで処理のインシデンスペクトルを次のように定義する。

$$\zeta_a = \begin{vmatrix} \zeta_{a1} \\ \zeta_{a2} \\ \vdots \\ \zeta_{an} \end{vmatrix}, \text{ ここに } \zeta_{aj} = \begin{cases} a \text{ 番目の処理が } f \text{ 番目の実験単位に施されれば } 1, \\ \text{然らざれば } 0, \end{cases}$$

$n \times k$ 行列

$$\Phi = \|\zeta_1 \ \zeta_2 \ \cdots \ \zeta_k\|$$

を処理のインシデンス行列という。そうなれば

$$\Phi' \Phi = \begin{vmatrix} b & & & \\ & 0 & & \\ & & b & \\ & & & b \end{vmatrix}$$

今 $n \times n$ 行列 T を

$$T \equiv \|t_{fg}\| = \Phi \Phi'$$

で定義すれば

$$t_{fg} = \begin{cases} f \text{ 番目と } g \text{ 番目の実験単位が同一処理を受けければ } 1, \\ \text{然らざれば } 0, \end{cases}$$

$$T^2 = b T \text{ 従って } \left(\frac{1}{b} T \right)^2 = \frac{1}{b} T \text{ (幂等)}$$

同様にしてブロックのインシデンスペクトルは

$$\eta_p = \begin{vmatrix} \eta_{p1} \\ \eta_{p2} \\ \vdots \\ \eta_{pn} \end{vmatrix}, \text{ ここで } \eta_{pf} = \begin{cases} f \text{ 番目の実験単位が } p \text{ 番目のブロックに属せば } 1, \\ \text{然らざれば } 0, \end{cases}$$

$n \times b$ 行列 $\Psi = \|\eta_1 \ \eta_2 \ \cdots \ \eta_b\|$ はブロックのインシデンス行列であって、

$$\Psi' \Psi = \begin{vmatrix} k & & & \\ & 0 & & \\ & & k & \\ & & & k \end{vmatrix}$$

で又

$$B = \Psi \Psi' = \begin{vmatrix} G_k & & & \\ & G_k & & \\ & & \ddots & \\ 0 & & & G_k \end{vmatrix}$$

但し G_k はその要素が凡て 1 である $k \times k$ 行列である。

$$N \equiv \|n_{ap}\|_{\substack{a=1 \cdots k \\ p=1 \cdots b}} = \Phi' \Psi$$

はこの配置のインシデンス行列であって、今の場合は

$$N = \begin{vmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \cdots & \cdots & \cdots & \cdots \\ 1 & 1 & \cdots & 1 \end{vmatrix}, \quad NN' = b G_k.$$

扱い、 $TB=BT=G_n$ であるから、 $I=I_n$, $G=G_n$, B , T は可換な一次代数をなし、これは乱塊法の関係代数といわれる、この代数の主算等元は

$$\frac{1}{n} G, \quad \frac{1}{k} B - \frac{1}{n} G, \quad \frac{1}{b} T - \frac{1}{n} G, \quad I - \frac{1}{b} T - \frac{1}{k} B + \frac{1}{n} G$$

であって、その階数は夫々

$$\alpha_0 = 1, \quad \alpha_1 = b - 1, \quad \alpha_2 = k - 1, \quad \alpha_3 = (b - 1)(k - 1)$$

である。

さて、正規回帰模型は

$$x_f = \gamma + \sum_{\alpha=1}^k \zeta_{\alpha f} \tau_\alpha + \sum_{p=1}^b \eta_{pf} \beta_p + e_f, \quad f = 1, \dots, n$$

ベクトル記号では

$$(Normal Regression) \quad \mathbf{x} = \gamma \mathbf{1} + \Phi \boldsymbol{\tau} + \Psi \boldsymbol{\beta} + \mathbf{e}$$

$\mathbf{x} = (x_1, \dots, x_n)'$ は観察値ベクトル、 γ は一般平均、 $\mathbf{1} = (1, \dots, 1)'$ 、 $\boldsymbol{\tau} = (\tau_1, \dots, \tau_b)'$ は処理効果、 $\boldsymbol{\beta} = (\beta_1 \cdots \beta_b)'$ はブロック効果で

$$\tau_1 + \cdots + \tau_b = 0, \quad \beta_1 + \cdots + \beta_b = 0$$

とする。 $\mathbf{e} = (e_1, \dots, e_n)'$ は誤差ベクトル（技術誤差）で $N(0, \sigma^2 I)$ に従う。

Fisher 模型は

$$(Fisher) \quad \mathbf{x} = \gamma \mathbf{1} + \Phi \boldsymbol{\tau} + \Psi \boldsymbol{\beta} + \pi,$$

$\pi = (\pi_0, \dots, \pi_n)'$ は単位誤差ベクトルであって、 $\pi_{(p-1)k+i} = \pi_i^{(p)}$ とかくと

$$\sum_{i=1}^k \pi_i^{(p)} = 0, \quad p = 1, \dots, b$$

又はベクトル記号では $\Psi \pi = 0$,

線形 Neyman 模型は

$$(Neyman) \quad \mathbf{x} = \gamma \mathbf{1} + \Phi \boldsymbol{\tau} + \Psi \boldsymbol{\beta} + \pi + \mathbf{e}$$

一般 Neyman 模型は

$$x_f = \gamma + \sum_{\alpha=1}^k \zeta_{\alpha f} \tau_\alpha + \sum_{p=1}^b \eta_{pf} \beta_p + \sum_{\alpha=1}^k \zeta_{\alpha f} \pi_{\alpha f} + e_f$$

$\pi_\alpha = (\pi_{\alpha 1}, \dots, \pi_{\alpha b})'$ は α 番目の処理に対する単位誤差ベクトル、これは直ちにベクトル記号ではかけない。

線形 Neyman 模型は正規回帰模型と Fisher 模型をその特別の場合として含んでいる。

さて、以下吾々は分散分析で現われる F -統計量の標本分布を線形 Neyman 模型の下で考えよう。

処理和を $T_\alpha = \zeta_\alpha' \mathbf{x}$, $\alpha = 1, \dots, k$ として $\mathbf{t} = (T_1, \dots, T_k)'$

ブロック和を $B_p = \eta_p' \mathbf{x}$, $p = 1, \dots, b$ として $\mathbf{b} = (B_1, \dots, B_b)'$

さて

$$\left(\frac{1}{b} T - \frac{1}{n} G \right) \mathbf{x} = \frac{1}{b} \Phi \mathbf{t} - \bar{x} \mathbf{1}$$

だから、

$$\mathbf{x}' \left(\frac{1}{b} T - \frac{1}{n} G \right) \mathbf{x} = \frac{1}{b} (T_1^2 + \cdots + T_k^2) - n \bar{x}^2: \text{ 处理平方和}$$

同様にして

$$\begin{aligned} \mathbf{x}' \left(\frac{1}{k} \mathbf{B} - \frac{1}{n} \mathbf{G} \right) \mathbf{x} &= \frac{1}{k} (B_1^2 + \dots + B_b^2) - n\bar{x}^2: \text{ブロック平方和} \\ \mathbf{x}' \left(\mathbf{I} - \frac{1}{k} \mathbf{B} - \frac{1}{b} \mathbf{T} + \frac{1}{n} \mathbf{G} \right) \mathbf{x} \\ &= \mathbf{x}' \left(\mathbf{I} - \frac{1}{n} \mathbf{G} \right) \mathbf{x} - \mathbf{x}' \left(\frac{1}{k} \mathbf{B} - \frac{1}{n} \mathbf{G} \right) \mathbf{x} - \mathbf{x}' \left(\frac{1}{b} \mathbf{T} - \frac{1}{n} \mathbf{G} \right) \mathbf{x} \\ &= \sum_{f=1}^n x_f^2 - \frac{1}{k} \sum_{p=1}^b B_p^2 - \frac{1}{b} \sum_{a=1}^k T_a^2 + n\bar{x}^2: \text{誤差平方和}. \end{aligned}$$

である。よって

$$\frac{F}{b-1} = \frac{\mathbf{x}' \left(\frac{1}{b} \mathbf{T} - \frac{1}{n} \mathbf{G} \right) \mathbf{x}}{\mathbf{x}' \left(\mathbf{I} - \frac{1}{k} \mathbf{B} - \frac{1}{b} \mathbf{T} + \frac{1}{n} \mathbf{G} \right) \mathbf{x}}$$

線形 Neyman 模型 $\mathbf{x} = \gamma \mathbf{1} + \Phi \boldsymbol{\tau} + \Psi \boldsymbol{\beta} + \boldsymbol{\pi} + \boldsymbol{e}$ の下では

$$\left(\frac{1}{b} \mathbf{T} - \frac{1}{n} \mathbf{G} \right) \mathbf{x} = \Phi \left(\boldsymbol{\tau} + \frac{1}{b} \boldsymbol{\Pi} \right) + \left(\frac{1}{b} \mathbf{T} - \frac{1}{n} \mathbf{G} \right) \boldsymbol{e}$$

但し $\boldsymbol{\Pi}_a = \zeta_a' \boldsymbol{\pi}$, $\boldsymbol{\Pi} = (\boldsymbol{\Pi}_1, \dots, \boldsymbol{\Pi}_k)'$ だから, 無作為化前には

$$\chi_1^2 = \mathbf{x}' \left(\frac{1}{b} \mathbf{T} - \frac{1}{n} \mathbf{G} \right) \mathbf{x} / \sigma^2$$

は自由度 $(k-1)$ の非心カイ自乗分布で, 非心定数は $\delta_1/2\sigma^2$ で

$$\delta_1 = b \left(\boldsymbol{\tau} + \frac{1}{b} \boldsymbol{\Pi} \right)' \left(\boldsymbol{\tau} + \frac{1}{b} \boldsymbol{\Pi} \right).$$

従って, その確率要素は

$$\exp(-\delta_1/2\sigma^2) \sum_{\mu=0}^{\infty} \frac{(\delta_1/2\sigma^2)^\mu}{\mu!} \frac{\left(\frac{\chi_1^2}{2}\right)^{\frac{k-1}{2}+\mu-1}}{\Gamma\left(\frac{k-1}{2}+\mu\right)} e^{-\frac{\chi_1^2}{2}} d\left(\frac{\chi_1^2}{2}\right).$$

同様にして, 無作為化前には

$$\chi_2^2 = \mathbf{x}' \left(\mathbf{I} - \frac{1}{k} \mathbf{B} - \frac{1}{b} \mathbf{T} + \frac{1}{n} \mathbf{G} \right) \mathbf{x} / \sigma^2$$

は自由度 $(b-1)(k-1)$ の非心カイ自乗分布で, 非心定数は $\delta_2/2\sigma^2$

$$\delta_2 = \boldsymbol{\pi}' \boldsymbol{\pi} - \frac{1}{b} \boldsymbol{\Pi}' \boldsymbol{\Pi} = \mathbf{A} - \frac{1}{b} \boldsymbol{\Pi}' \boldsymbol{\Pi}.$$

従って, その確率要素は

$$\exp(-\delta_2/2\sigma^2) \sum_{\nu=0}^{\infty} \frac{(\delta_2/2\sigma^2)^\nu}{\nu!} \frac{\left(\frac{\chi_2^2}{2}\right)^{\frac{(b-1)(k-1)}{2}+\nu-1}}{\Gamma\left(\frac{(b-1)(k-1)}{2}+\nu\right)} e^{-\frac{\chi_2^2}{2}} d\left(\frac{\chi_2^2}{2}\right).$$

これらは互に独立だから, $F/(b-1)$ の無作為化前の確率要素は次のようになる.

$$\exp\left(-\frac{\delta_1 + \delta_2}{2\sigma^2}\right) \sum_{\mu=0}^{\infty} \sum_{\nu=0}^{\infty} \frac{(\delta_1/2\sigma^2)^\mu}{\mu!} \frac{(\delta_2/2\sigma^2)^\nu}{\nu!}$$

$$\begin{aligned}
& \times \frac{\Gamma\left(\frac{b(k-1)}{2} + \mu + \nu\right)}{\Gamma\left(\frac{k-1}{2} + \mu\right)\Gamma\left(\frac{(b-1)(k-1)}{2} + \nu\right)} \left(\frac{F}{b-1}\right)^{\frac{k-1}{2} + \mu - 1} \\
& \times \left(1 + \frac{F}{b-1}\right)^{-\frac{b(k-1)}{2} - \mu - \nu} d\left(\frac{F}{b-1}\right) \\
= & \exp(-\xi/2\sigma^2) \sum_{l=0}^{\infty} \frac{(\xi/2\sigma^2)^l}{l!} \sum_{\mu+\nu=l} \frac{l!}{\mu!\nu!} \eta^\mu (1-\eta)^\nu \\
& \times \frac{\Gamma\left(\frac{b(k-1)}{2} + \mu + \nu\right)}{\Gamma\left(\frac{k-1}{2} + \mu\right)\Gamma\left(\frac{(b-1)(k-1)}{2} + \nu\right)} \left(\frac{F}{b-1}\right)^{\frac{k-1}{2} + \mu - 1} \\
& \times \left(1 + \frac{F}{b-1}\right)^{-\frac{b(k-1)}{2} - \mu - \nu} d\left(\frac{F}{b-1}\right) \tag{1}
\end{aligned}$$

但しここで

$$\Xi = \tau' \Pi, \quad H = \frac{1}{b} \Pi' \Pi$$

$$\xi = A + \tau' \tau + 2\Xi, \quad \eta = (\tau' \tau + 2\Xi + H)/(A + \tau' \tau + 2\Xi)$$

又は

$$\eta = 1 - \frac{A - H}{A + \tau' \tau + 2\Xi}$$

ところで単位誤差ベクトル $\pi=0$ とした正規回帰模型の下では $F/(b-1)$ の確率要素は

$$\begin{aligned}
& \exp(-\tau' \tau / 2\sigma^2) \sum_{l=0}^{\infty} \frac{(\tau' \tau / 2\sigma^2)^l}{l!} \frac{\Gamma\left(\frac{b(k-1)}{2} + l\right)}{\Gamma\left(\frac{k-1}{2} + l\right)\Gamma\left(\frac{(b-1)(k-1)}{2}\right)} \\
& \times \left(\frac{F}{b-1}\right)^{\frac{k-1}{2} + l - 1} \left(1 + \frac{F}{b-1}\right)^{-\frac{b(k-1)}{2} - l} d\left(\frac{F}{b-1}\right) \tag{2}
\end{aligned}$$

で与えられる。

無作為化を考慮したときの $F/(b-1)$ 確率要素は (1) の平均を取って

$$\begin{aligned}
& \mathcal{E} \left[\exp\left(-\frac{\xi}{2\sigma^2}\right) \sum_{l=0}^{\infty} \frac{(\xi/2\sigma^2)^l}{l!} \sum_{\mu+\nu=l} \frac{l!}{\mu!\nu!} \eta^\mu (1-\eta)^\nu \right. \\
& \times \frac{\Gamma\left(\frac{b(k-1)}{2} + \mu + \nu\right)}{\Gamma\left(\frac{k-1}{2} + \mu\right)\Gamma\left(\frac{(b-1)(k-1)}{2} + \nu\right)} \left(\frac{F}{b-1}\right)^{\frac{k-1}{2} + \mu - 1} \\
& \times \left. \left(1 + \frac{F}{b-1}\right)^{-\frac{b(k-1)}{2} - \mu - \nu} \right] d\left(\frac{F}{b-1}\right) \tag{3}
\end{aligned}$$

となる。インシデンス行列 Φ が無作為化の操作を通じて確率変数となり、従って、 Π が、更に ξ, η が確率変数になる訳である。吾々は k を止めておいて、 $b \rightarrow \infty$ とするとき、後述する意味で (2), (3) に対応する確率分布は $(M)_d$ 型で漸近的に対等であることが証明出来る。その為にはブロック内の単位誤差について或条件が必要である。

検定力関数については少しく面倒なので、仮説 $H: \tau=0$ の下での分布について上記のこと

を示そう。

$H: \tau=0$ の下では $F/(b-1)$ の無作為化前の確率要素は

$$\exp(-A/2\sigma^2) \sum_{l=0}^{\infty} \frac{(A/2\sigma^2)^l}{l!} \sum_{\mu+\nu=l} \frac{l!}{\mu!\nu!} \theta^\mu (1-\theta)^\nu h_{\mu,\nu}(F) dF \quad (4)$$

但し

$$h_{\mu,\nu}(F) = \frac{\Gamma\left(\frac{b(k-1)}{2} + \mu + \nu\right)}{\Gamma\left(\frac{k-1}{2} + \mu\right)\Gamma\left(\frac{(b-1)(k-1)}{2} + \nu\right)} \left(\frac{F}{b-1}\right)^{\frac{k-1}{2} + \nu - 1} \times \left(1 + \frac{F}{b-1}\right)^{-\frac{b(k-1)}{2} - \mu - \nu} \frac{1}{b-1} \quad (5)$$

又

$$\theta = \frac{1}{b} \Pi' \Pi / A \quad (6)$$

b 個の順列

$$\sigma_p = \begin{pmatrix} 1 & 2 & \cdots & k \\ \sigma_p(1) & \sigma_p(2) & \cdots & \sigma_p(k) \end{pmatrix}, \quad p=1, \dots, b$$

に対応する $k \times k$ の順列行列を S_{σ_p} とする。即ち

$$(1 \ 2 \ \cdots \ k) \ S_{\sigma_p} = (\sigma_p(1) \ \sigma_p(2) \ \cdots \ \sigma_p(k))$$

$$U_\sigma = \begin{vmatrix} S_{\sigma_1} & & & & 0 \\ & S_{\sigma_2} & & & \\ & & \ddots & & \\ 0 & & & \ddots & S_{\sigma_b} \end{vmatrix},$$

ある 1 つの Φ_0 を定めて、無作為化によって インシデンス行列は 確率変数となって $P\{\Phi = U_\sigma \Phi_0\} = 1/(k!)^b$ となる。

$$\Pi_{ap}^\sigma = \sum_{i=1}^k \zeta_{a(p-1)k+i} \pi^{(p)}_{\sigma_p(i)}, \quad a=1, \dots, k; \quad p=1, \dots, b$$

$$\Pi_p^\sigma = (\Pi_{1p}^\sigma \ \cdots \ \Pi_{kp}^\sigma)', \quad p=1, \dots, b$$

とすれば

$$\Pi^\sigma \equiv (U_\sigma \Phi_0)' \pi = \sum_{p=1}^b \Pi_p^\sigma.$$

ここで

$$\mathcal{E}(\Pi_p^\sigma) = 0, \quad \mathcal{E}(\Pi_p^\sigma \Pi_p^{\sigma'}) = \frac{A_p}{k(b-1)} A_p$$

但し

$$A_1 = \begin{vmatrix} (k-1)n_{1p} & -n_{1p} n_{2p} \cdots & -n_{1p} n_{kp} \\ -n_{2p} n_{1p} & (k-1)n_{2p} \cdots & -n_{2p} n_{kp} \\ \cdots & \cdots & \cdots \\ -n_{kp} n_{1p} & -n_{kp} n_{2p} \cdots & (k-1)n_{kp} \end{vmatrix}$$

従って吾々の場合には

$$A_p = \begin{vmatrix} k-1 & -1 & \cdots & -1 \\ -1 & k-1 & \cdots & -1 \\ \cdots & \cdots & \cdots & \cdots \\ -1 & -1 & \cdots & k-1 \end{vmatrix} = k I_k - G_k$$

従つて

$$\sum_{p=1}^b A_p = rk I_v - NN' = b(kI_k - G_k)$$

但しここで $A_p = \sum_{i=1}^k \pi_i^{(p)^2}$, 又

$$\mathcal{E}(\Pi_{\alpha p}^{\sigma}) = \frac{A_p}{k}, \quad \mathcal{E}(\Pi_p^{\sigma} \Pi_p^{\sigma}) = A_p$$

も計算出来る。ここで次の条件をおく

$$(U) \quad \begin{aligned} \bar{A} &\equiv \frac{1}{b} \sum_{p=1}^b A_p \rightarrow A_0 \\ & \quad , \quad b \rightarrow \infty \\ & \quad \frac{1}{b} \sum_{p=1}^b (A_p - \bar{A})^2 \rightarrow 0 \end{aligned}$$

Feller の中心極限定理: $X_1^{(n)}, \dots, X_n^{(n)}$ は互に独立な k 次元の確率ベクトルとし。

$$\mathcal{E}(X_i^{(n)}) = \mathbf{o}, \quad \mathcal{E}(X_i^{(n)} X_i^{(n)\prime}) = A_i^{(n)}, \quad i=1, \dots, n$$

として、次の条件が成立つものとする。

$$\sum_{i=1}^n A_i^{(n)} \rightarrow A (+0), \quad \sum_{i=1}^n \int_{\|\mathbf{x}\| > \varepsilon} \|\mathbf{x}\|^2 dF_i^{(n)} \rightarrow 0, \quad n \rightarrow \infty.$$

但し $F_i^{(n)}$ は $X_i^{(n)}$ の c. d. f. で、 $\|\mathbf{x}\|$ は k 次元ユークリッド空間のベクトル \mathbf{x} のユークリッド・ノルムである。そのときは

$$\mathcal{L} \left(\sum_{i=1}^n X_i^{(n)} \right) \rightarrow N(\mathbf{o}, A), \quad n \rightarrow \infty$$

i.e. $t = (t_1, \dots, t_k)'$ として、 A がノン・シュンギュラーなら

$$\lim_{n \rightarrow \infty} P \left(\sum_{i=1}^n X_i^{(n)} < t \right) = \frac{1}{\sqrt{(2\pi)^k |A|^{1/2}}} \int_{-\infty}^{t_1} \int_{-\infty}^{t_k} e^{-(1/2)u' A^{-1} u} du_1 \cdots du_k.$$

この Feller の中心極限定理を用いる為に、 $n=b$, $i=p$ とし

$$X_i^{(n)} = \frac{1}{\sqrt{b}} \Pi_p^{\sigma}, \quad p = 1, 2, \dots, b$$

とすれば

$$\mathcal{E}(X_i^{(n)}) = \frac{1}{\sqrt{b}} \mathcal{E}(\Pi_p^{\sigma}) = \mathbf{0}, \quad \mathcal{E}(X_i^{(n)} X_i^{(n)\prime}) = \frac{1}{b} \mathcal{E}(\Pi_p^{\sigma} \Pi_p^{\sigma}) = \frac{A_p}{k(k-1)b} A_p$$

従つて

$$\begin{aligned} \sum_{i=1}^n \mathcal{E}(X_i^{(n)} X_i^{(n)\prime}) &= \frac{1}{b k (k-1)} \sum_{p=1}^b A_p A_p = \frac{1}{k(k-1)} \frac{1}{b} \left[\bar{A} \sum_{p=1}^b A_p + \sum_{p=1}^b (A_p - \bar{A}) A_p \right] \\ &= \frac{\bar{A}}{k(k-1)} (kI_k - G_k) + \frac{1}{k(k-1)} \frac{1}{b} \sum_{p=1}^b ((A_p - \bar{A}) A_p) \rightarrow \frac{A_0}{k(k-1)} (kI_k - G_k) \equiv A \end{aligned}$$

何となれば、

$$\begin{aligned} \left| \left(\frac{1}{b} \sum_{p=1}^b (A_p - \bar{A}) A_p \right)_{ij} \right|^2 &= \left| -\frac{1}{b} \sum_{p=1}^b (A_p - \bar{A}) n_{ip} n_{jp} \right|^2 \\ &\leq \left[\frac{1}{b} \sum_{p=1}^b |A_p - \bar{A}| \right]^2 \leq \frac{1}{b} \sum_{p=1}^b (A_p - \bar{A})^2 \rightarrow 0, \quad b \rightarrow \infty. \end{aligned}$$

更に

$$\sum_{p=1}^b \int_{\|\mathbf{x}\| > \varepsilon} \|\mathbf{x}\|^2 dF_p^{(b)} \leq \sum_{p=1}^b \mathcal{E} \left(\frac{1}{b} \Pi_p^{\sigma} \Pi_p^{\sigma} \right) \cdot P_F \left(\frac{1}{b} \Pi_p^{\sigma} \Pi_p^{\sigma} > \varepsilon^2 \right)$$

$$\begin{aligned} &\leq \frac{1}{\varepsilon^2} \sum_{p=1}^b \left[\mathcal{E} \left(\frac{1}{b} \Pi_p^{\sigma'} \Pi_p^{\sigma} \right) \right]^2 = \frac{1}{\varepsilon^2 b^2} \sum_{p=1}^b A_p^2 \\ &= \frac{1}{\varepsilon^2 b} \left[\bar{A}^2 + \frac{1}{b} \sum_{p=1}^b (A_p - \bar{A})^2 \right] \rightarrow 0, \quad b \rightarrow \infty. \end{aligned}$$

これで Feller の中心極限定理の条件は満たされていることが確かめられたから

$$\mathcal{L} \left(\frac{1}{\sqrt{b}} \Pi^{\sigma} \right) \rightarrow N \left(\mathbf{0}, \frac{A_0}{k(k-1)} (kI_k - G_k) \right), \quad b \rightarrow \infty \quad (7)$$

なることが判る。

そして

$$\begin{aligned} \frac{k-1}{A_0} \frac{1}{b} \bar{\Pi}^{\sigma'} \bar{\Pi}^{\sigma} &= \frac{1}{\sqrt{b}} \bar{\Pi}^{\sigma'} \left[\frac{A_0}{k(k-1)} (kI_{k-1} - G_{k-1}) \right]^{-1} \frac{1}{\sqrt{b}} \bar{\Pi}^{\sigma} \\ &= \frac{1}{\sqrt{b}} \bar{\Pi}^{\sigma'} \left[\frac{k-1}{A_0} (I_{k-1} + G_{k-1}) \right] \frac{1}{\sqrt{b}} \bar{\Pi}^{\sigma} \end{aligned}$$

但しここで

$$\bar{\Pi}^{\sigma'} = (\Pi_1^{\sigma'} \Pi_2^{\sigma'} \cdots \Pi_{k-1}^{\sigma'})$$

よって

$$\mathcal{L} \left(\frac{k-1}{A_0} \frac{1}{b} \bar{\Pi}^{\sigma'} \bar{\Pi}^{\sigma} \right) \rightarrow \chi^2_{k-1}, \quad b \rightarrow \infty$$

又は $\theta = \frac{1}{b} \bar{\Pi}' \bar{\Pi} / A$ であるから

$$\mathcal{L}(b(k-1)\theta) \rightarrow \chi^2_{k-1}, \quad b \rightarrow \infty \quad (8)$$

b 確率変数の漸近的同等 池田貞雄博士によって展開された確率変数の漸近的同等に関する理論を引用する。

n_s 次元の確率変数列 $\{X_s^{(n_s)}\}$ と $\{Y_s^{(n_s)}\}$ について

$$d(X_s^{(n_s)}, Y_s^{(n_s)}; C) \equiv \sup_{E \in C} |P X_s^{(n_s)}(E) - P Y_s^{(n_s)}(E)| \rightarrow 0, \quad s \rightarrow \infty$$

但し C は n_s 次元ユークリッド空間 R_{n_s} のボレルフィールド B の空でないある部分族とする。このとき $\{X_s^{(n_s)}\}$ と $\{Y_s^{(n_s)}\}$ は $(C)_d$ 型の意味で漸近的に同等であるという。

$$X_s^{(n_s)} \sim Y_s^{(n_s)} \quad (C)_d, \quad s \rightarrow \infty$$

とかくことにする。

族 M は次のような集合 M の全体から成るものとする。

$$M = \{x = (x_1, \dots, x_{n_s}); -\infty \leq x_i < a_i, i = 1, \dots, n_s\}$$

但し a_i は $\pm\infty$ を含んで実数値をとるものとする。若し或 i に対して $a_i = -\infty$ なら $M = \emptyset$ (空集合) で、又すべての i に対して $a_i = \infty$ なら $M = R_{n_s}$ である。次に族 S は次のような集合 S の全体から成るものとする。

$$S = \{x = (x_1, \dots, x_{n_s}); b_i \leq x_i < a_i, i = 1, \dots, n_s\}$$

但し a_i, b_i は $\pm\infty$ を含んで実数値をとる。若しある i に対して $a_i = b_i$ なら $S = \emptyset$ で、すべての i に対して $b_i = -\infty, a_i = \infty$ なら $S = R_{n_s}$ である。明らかに

$$M \subseteq S \subseteq B.$$

任意に与えられた $\varepsilon > 0$ に対して族 S に属する有界な部分集合 S とある正整数 s_0 があって、すべての $s \geq s_0$ に対して

$$P X_s(S) > 1 - \varepsilon$$

となるとき、 $\{X_s\}$ は性質 $B(S)$ をもつという。更に μ をルベーグ測度として、与えられた

$\varepsilon > 0$ に対して $\delta(\varepsilon)$ と $s_0(\varepsilon)$ があって, $\mu(E) < \delta(\varepsilon)$, $E \in S$ なる限り, すべての $s \geq s_0(\varepsilon)$ に対して

$$P^{X_s}(E) < \varepsilon$$

となるならば, $\{X_s\}$ は性質 $B(S)$ をもつといふ.

池田の定理 $\{X_s^{(n_s)}\}$ は性質 $C(S)$ と $C(S)$ をもつとする. 然ばば, $c_i^s \rightarrow 1$, $s \rightarrow \infty$ なる $\{c_1^s, \dots, c_{n_s}^s\}$ に対して

$$X_s^{(n_s)} = (X_1^s, \dots, X_{n_s}^s) \sim Y_s^{(n_s)} = (c_1^s X_1^s, \dots, c_{n_s}^s X_{n_s}^s), (M)_d, s \rightarrow \infty.$$

が成立つ.

ところで

$$(k-1) \Pi^{\sigma'} \Pi^{\sigma} / (b A_0) = \frac{\bar{A}}{A_0} \cdot (k-1) \Pi^{\sigma'} \Pi^{\sigma} / (b \bar{A}) \text{ であつて, } \frac{\bar{A}}{A_0} \rightarrow 1, b \rightarrow \infty$$

であるから

$$b(k-1)\theta = (k-1) \Pi^{\sigma'} \Pi^{\sigma} / (b \bar{A}) \sim (k-1) \Pi^{\sigma'} \Pi^{\sigma} / (b A_0), (M)_d, b \rightarrow \infty \quad (9)$$

更に吾々は次の池田の定理を引用する.

池田の定理 $X_s^{(n_s)}$, $Y_s^{(n_s)}$ の p. d. f. を夫々 $p_s(x)$, $p_s(y)$ とするとき Kullback-Leibler の平均情報

$$I(p_s(x) : p_s(y)) = \mathcal{E}_{p_s(x)} \left(\log \frac{p_s(x)}{p_s(y)} \right) \rightarrow 0, s \rightarrow \infty$$

ならば

$$X_s^{(n_s)} \sim Y_s^{(n_s)}, (M)_d, s \rightarrow \infty.$$

今 $\{X_b\}$ はベータ型 b. d. f.

$$f_b(x) = \frac{\Gamma\left(\frac{b(k-1)}{2}\right)}{\Gamma\left(\frac{k-1}{2}\right)\Gamma\left(\frac{(b-1)(k-1)}{2}\right)} x^{\frac{k-1}{2}-1} (1-x)^{\frac{(b-1)(k-1)}{2}-1}, 0 < x < 1$$

をもつとし, $\{\theta\}$ は p. d. f.

$$g_b(x) = \frac{\left\{ \frac{b(k-1)}{2} \right\}^{\frac{k-1}{2}}}{\Gamma\left(\frac{k-1}{2}\right)} x^{\frac{k-1}{2}-1} e^{-\frac{b(k-1)}{2}x}, 0 < x < \infty$$

をもつ. Kullback-Leibler の平均情報は

$$\begin{aligned} I(f : g) &= \log \Gamma\left(\frac{b(k-1)}{2}\right) - \log \Gamma\left(\frac{(b-1)(k-1)}{2}\right) - \frac{k-1}{2} \log \frac{b(k-1)}{2} \\ &\quad + \left[\frac{(b-1)(k-1)}{2} - 1 \right] \mathcal{E}_f \log (1 - X_b) + \frac{b(k-1)}{2} \mathcal{E}_f(X_b) \\ &\sim \left[\frac{b(k-1)}{2} + \frac{1}{2} \right] \log \frac{b(k-1)}{2} - \frac{b(k-1)}{2} - \left[\frac{(b-1)(k-1)}{2} + \frac{1}{2} \right] \\ &\quad \times \log \frac{(b-1)(k-1)}{2} + \frac{(b-1)(k-1)}{2} - \frac{k-1}{2} \log \frac{b(k-1)}{2} \\ &\quad + \left[\frac{(b-1)(k-1)}{2} - 1 \right] \left[\frac{\Gamma'\left(\frac{(b-1)(k-1)}{2}\right)}{\Gamma\left(\frac{(b-1)(k-1)}{2}\right)} - \frac{\Gamma'\left(\frac{b(k-1)}{2}\right)}{\Gamma\left(\frac{b(k-1)}{2}\right)} \right] + \frac{k-1}{2} \end{aligned}$$

ここで

$$\frac{\Gamma'(p)}{\Gamma(p)} = \log(p) \cdot \frac{1}{2p} + O\left(\frac{1}{p^2}\right), \quad p \rightarrow \infty$$

を用いると

$$I(f:g) \rightarrow 0, \quad b \rightarrow \infty$$

よって

$$\theta \sim X_b(M)_d, \quad b \rightarrow \infty \quad (10)$$

周辺分布の漸近的同等について次の定理がある。

池田の定理 (X_s, Y_s) を2次元確率変数の列として, X_s の分布関数を $F_s(x)$, $X_s=x$ のときの Y_s の条件付 p.d.f. を $p_s(y|x)$ とする. $\{X_s\}$ の分布関数を $F_s^*(x)$ とする. Y_s の p.d.f. は

$$h_s(y) = \int p_s(y|x) dF_s(x)$$

であるが, p.d.f.

$$h_s^*(y) = \int p_s(y|x) dF_s^*(x)$$

をもつ確率変数を Y_s^* とする。

- (i) $\{(X_s - d_s)/c_s\}$ が性質 $B(S)$ をもつような $\{(c_s, d_s)\}$ がある。
- (ii) $q_s(y|x) = p_s(y|c_s x + d_s)$ は x について微分可能。
- (iii) S に属する有界部分集合 E について, 充分大きな s のすべての値に対して

$$\sup_{x \in E} \int_{R_1} \left| \frac{\partial q_s(y|x)}{\partial x} \right| dy$$

は有界。

以上三条件が成立すれば

$$X_s \sim X^*(M)_d \Rightarrow Y_s \sim Y_s^*(M)_d, \quad s \rightarrow \infty$$

ここで吾々の場合に適用するには (iii) をチェックすれば良い,

$$p_b(F|\theta) = \exp(-A/2\sigma^2) \sum_{l=0}^{\infty} \frac{(A/2\sigma^2)^l}{l!} \sum_{\mu+\nu=l} \frac{l!}{\mu! \nu!} \theta^\mu (1-\theta)^\nu h_{\mu,\nu}(F)$$

より

$$\begin{aligned} \frac{\partial q_b(F|\theta)}{\partial \theta} &= \frac{1}{b(k-1)} \frac{\partial}{\partial \theta} p_b(F|\theta) \Big|_{\theta=\frac{\theta}{b(k-1)}} \\ &= \frac{1}{b(k-1)} \exp(-A/2\sigma^2) \sum_{l=1}^{\infty} \frac{(A/2\sigma^2)^l}{l!} \sum_{\mu+\nu=l} \frac{l!}{\mu! \nu!} \left[\mu \left(\frac{\theta}{b(k-1)} \right)^{\mu-1} \right. \\ &\quad \times \left. \left(1 - \frac{\theta}{b(k-1)} \right)^\nu - \nu \left(\frac{\theta}{b(k-1)} \right)^\mu \left(1 - \frac{\theta}{b(k-1)} \right)^{\nu-1} \right] h_{\mu,\nu}(F) \end{aligned}$$

よって

$$\begin{aligned} \int_{R_1} \left| \frac{\partial q_b(F|\theta)}{\partial \theta} \right| dF &= \frac{1}{b(k-1)} \exp(-A/2\sigma^2) \sum_{l=1}^{\infty} \frac{(A/2\sigma^2)^l}{l!} \sum_{\mu+\nu=l} \frac{l!}{\mu! \nu!} \\ &\quad \times \left[\mu \left(\frac{\theta}{b(k-1)} \right)^{\mu-1} \left(1 - \frac{\theta}{b(k-1)} \right)^\nu \right. \\ &\quad \left. + \nu \left(\frac{\theta}{b(k-1)} \right)^\mu \left(1 - \frac{\theta}{b(k-1)} \right)^{\nu-1} \right] \\ &= \frac{A}{b(k-1)\sigma^2} \rightarrow \frac{A_0}{(k-1)\sigma^2}, \quad b \rightarrow \infty \end{aligned}$$

従つて

$$\sup_{\theta} \int_R \left| \frac{\partial g_b(F|\theta)}{\partial \theta} \right| dF$$

は充分大きな b の値に対して有界である。

結局は吾々は次の定理を得た。

定理 条件 $b \rightarrow \infty$ のとき

$$\frac{1}{b} A = \bar{A} \rightarrow A_0 \quad \frac{1}{b} \sum_{p=1}^b (A_p - \bar{A})^2 \rightarrow 0$$

があれば、乱塊法の分散分析における F 統計量の帰無仮説の下での分布は、Neyman 模型で考えても自由度 $(k-1, (b-1)(k-1))$ の中心 F 分布と $(M)_d$ 型の意味で漸近的に同等である

カルガリィ大学

参 考 文 献

- [1] Fisher, R.A. (1926) The arrangement of field experiment, J. Ministry of Agric. 23, 503–513.
- [2] Ikeda, S. (1963) Asymptotic equivalence of probability distributions with applications to some problems of asymptotic independence. Ann. Inst. Stat. Math. Tokyo, 15, 87–116.
- [3] Ikeda, S., J. Ogawa and M. Ogasawara (1965) On the asymptotic distribution of the F -statistic under the null-hypothesis in a randomized PBIB design with m associate classes under the Neyman model. UNC Inst. Stat. Mimeo Series 454.
- [4] Ikeda, S. and J. Ogawa (1966) On the asymptotic non-null distribution of the F -statistic for testing a partial null-hypothesis in a randomized PBIB design with m associate classes under the Neyman model, UNC Inst. Stat. Mimeo Series 466.
- [5] Ikeda, S. (1968) Asymptotic equivalence of real probability distributions, Ann. Inst. Stat. Math. Tokyo, 20, 339–362.
- [6] Kempthorne, O. (1952) The Design and Analysis of Experiments. Wiley, New York.
- [7] Mann, H.B. (1949) Analysis and Designs of Experiments. Dover, New York.
- [8] McCarthy, M.D. (1939) On the application of the z -test to randomized blocks, Ann. Math. Stat. 10, 337–359.
- [9] Neyman, J., K. Iwaskiewicz and S. Kolodziejczyk (1935) Statistical problems in agricultural experimentation, J. Roy. Statist. Soc. Seyol. 2, 107–154, Discussion 154–180.
- [10] Ogawa, J. (1961) The effect of randomization on the analysis of randomized block design, Ann. Inst. Stat. Math. Tokyo 13, 105–117.
- [11] Ogawa, J.C. (1962) On the randomization in Latin-square design under the Neyman model, Proc. Inst. Stat. Math. Tokyo, 10, 1–16.
- [12] Ogawa, J. (1963) On the null-distribution of the F -statistic in a randomized BIB design under the Neyman model, Ann. Math. Stat. 34, 1558–1568.
- [13] Ogawa, J., S. Ikeda and M. Ogasawara (1964) On the null-distribution of the F -statistic in a randomized PBIB design with two associate classes under the Neyman model, Contributions to Statistics and Probability, Essays in Memory of S.N. Roy. Chapt. 26
- [14] Ogawa, J. and G. Ishii (1965) The relationship algebra and the analysis of a PBIB design, Ann. Math. Statist., 36, 1815–1828.
- [15] Ogawa, J., S. Ikeda and M. Ogasawara (1967) On the null-distribution of the F -statistic for testing a partial null-hypothesis in a randomized PBIB design with m associate classes under the Neyman model, Ann. Inst. Stat. Math. Tokyo, 19, 313–330.
- [16] Ogawa, J. and S. Ikeda (1970) The asymptotic null distribution of the F -statistic for testing a partial null-hypothesis in a randomized PBIB design with m associate classes under the Neyman model, Presented at the 37th Session of the ISI meeting in September 1969.
- [17] Ogawa, J. and S. Ikeda (1970) The asymptotic non-null distribution of the F -statistic for testing a partial null-hypothesis in a randomized PBIB design with m associate classes under the Neyman model, unpublished yet.
- [18] Pitman, E.J.G. (1937) significance tests which can be applied to samples from any populations. II The analysis of variance tests. Biometrika 29, 70–90
- [19] Welch, B.L. (1937) On the z -test in randomized blocks and Latin squares. Biometrika 29, 21–52.

- [20] Wilk, M.B. (1955), Linear models and randomized experiments, Ph. D. thesis, Iowa State College.
- [21] Wilk, M.B. (1955) The randomization analysis of a generalized randomized block design, *Biometrika* 42, 70–79.