ON THE NULL-DISTRIBUTION OF THE F-STATISTIC IN A RANDOMIZED BALANCED INCOMPLETE BLOCK DESIGN UNDER THE NEYMAN MODEL

By Junjiro Ogawa

Nihon University, Tokyo

Summary and introduction. While the analysis of variance test statistic F in a balanced incomplete block design without randomization is a constant under the Fisher model, i.e., a linear model without technical errors, and it has no distribution at all, under the Neyman model, i.e., a model with technical errors, its null-distribution is a non-central F-distribution whose non-centrality parameter being a quadratic form of unit-errors. This is carried out in Section 1. The mean value and the variance of θ with respect to the permutation distribution due to the randomization are calculated in Section 2, and in Section 3, the null-distribution of the F-statistic after the randomization is shown to be approximated by the familiar central F-distribution under the Neyman model assuming no interaction between treatments and experimental units, if the following two conditions are satisfied:

(i) the variances of unit-errors within blocks are sufficiently uniform from block to block, and

(ii) the number of blocks is sufficiently large.

Since the unit-errors are not directly observable, how one can group the experimental units into blocks in such a way as the above Condition (i) would be satisfied is another problem, which is left open in this paper.

R. A. Fisher [2] initiated the use of the so-called "randomization procedure" in order to control the unit-errors in block designs. Mathematical treatments of the Fisher randomization in randomized block and the Latin-square designs were made by B. L. Welch [12], E. J. G. Pitman [11] and M. B. Wilk [14]. Underlying models in those works may be called the "Fisher models", i.e., containing no technical errors. J. Neyman et al. [7] and M. B. Wilk [13] pointed out that there are instances in which a model with technical errors is more adequate by the very nature of the problem under consideration, and the present author calls this sort of models the "Neyman models" for convenience. M. D. McCarthy [6] investigated the null-distribution of the analysis of variance test statistic in a randomized block design under the Neyman model, and he came out with rather pessimistic results. J. Ogawa [10] treated the same problem, and his result turned out to be supporting the usual approximation by the familiar central F-distribution.

The purpose of this article is the treatment of the same null-distribution problem for a randomized balanced incomplete block design under the Neyman

Received August 15, 1961; revised May 20, 1963.

model. Since a randomized block design is a limiting case of a randomized BIBD, this article should be regarded as a generalization of the earlier work.

1. The null-distribution of the F-statistic of a balanced incomplete block design under the Neyman model before and after the randomization. We are concerned with the analysis of variance of a BIBD with parameters v, b, r, k and λ. About the definition of BIBD and notations being used here, references should be made to R. C. Bose [1] and A. T. James [3].

Let the incidence matrices of treatments and blocks be, respectively, $\mathbf{\Phi} = \|\boldsymbol{\zeta}_1 \cdots \boldsymbol{\zeta}_v\|$ and $\mathbf{\Psi} = \|\boldsymbol{\eta}_1 \cdots \boldsymbol{\eta}_b\|$, where

$$\boldsymbol{\zeta}_\alpha' = (\zeta_{\alpha 1} \cdots \zeta_{\alpha n}),\ \zeta_{\alpha f} = 1, \text{ if the } f\text{th unit receives the } \alpha\text{th treatment,}$$
$$= 0, \text{ otherwise,}$$
$$\alpha = 1, \cdots, v; f = 1, \cdots, n,$$

and

$$\boldsymbol{\eta}_a' = (\eta_{a1} \cdots \eta_{an}),\ \eta_{af} = 1, \text{ if the } f\text{th unit belongs to the } a\text{th block,}$$
$$= 0, \text{ otherwise,}$$
$$a = 1, \cdots, b; f = 1, \cdots, n.$$

We have, of course, $n = vr = bk$. Then a general additive Neyman model can be expressed as

$$(1.1)\quad x_f = \gamma + \sum_{\alpha=1}^{v} \zeta_{\alpha f}\tau_\alpha + \sum_{a=1}^{b} \eta_{af}\beta_a + \sum_{\alpha=1}^{v} \zeta_{\alpha f}\pi_{\alpha f} + e_f, \quad f = 1, \cdots, n.$$

In Equation (1.1), x_f stands for the observation on the fth unit, γ is the general mean, τ_α, $\alpha = 1, \cdots, v$, and β_a, $a = 1, \cdots, b$, are treatment effects and block effects which are subject to the restrictions

$$(1.2)\quad \sum_{\alpha=1}^{v} \tau_\alpha = 0 \quad \text{and} \quad \sum_{a=1}^{b} \beta_a = 0.$$

Also, $\pi_{\alpha f}$ stands for the unit-error of the fth unit when it receives the αth treatment ($\alpha = 1, \cdots, v$ and $f = 1, \cdots, n$) and the unit errors are subjected to the restrictions

$$(1.3)\quad \mathbf{\Psi}'\boldsymbol{\pi}_\alpha = \mathbf{0}, \qquad \alpha = 1, \cdots, v,$$

where $\boldsymbol{\pi}_\alpha' = (\pi_{\alpha 1} \cdots \pi_{\alpha n})$. Finally, in Equation (1.1), e_f is the technical error of the fth unit and $\mathbf{e}' = (e_1 \cdots e_n)$ is assumed to be distributed as $N(\mathbf{0}', \sigma^2\mathbf{I})$, where $\mathbf{I}$ stands for the unit matrix of order n.

If there is no interaction between treatments and units, (1.1) becomes $x_f =$

$\gamma + \sum_{\alpha=1}^{v} \zeta_{\alpha f}\tau_\alpha + \sum_{a=1}^{b} \eta_{af}\beta_a + \pi_f + e_f$, $f = 1, \cdots, n$, or, in vector notation,

$$\mathbf{x} = \gamma\mathbf{1} + \mathbf{\Phi\tau} + \mathbf{\Psi\beta} + \boldsymbol{\pi} + \mathbf{e}, \tag{1.4}$$

where $\mathbf{1}' = (1\cdots1)$, $\boldsymbol{\tau}' = (\tau_1\cdots\tau_v)$, $\boldsymbol{\beta}' = (\beta_1\cdots\beta_b)$, and $\boldsymbol{\pi}' = (\pi_1\cdots\pi_n)$.

Assuming the Neyman model (1.4), the sampling distribution of the analysis of variance test statistic $F = (n - v - b + 1)s_t^2/[(v - 1)s_e^2]$ under the null-hypothesis $H_0 : \boldsymbol{\tau} = \mathbf{0}$, will be considered firstly without randomization and then with randomization. The quantities s_t^2 and s_e^2, the sums of squares due to treatments (adjusted) and due to errors respectively, are given by

$$s_t^2 = (k/vr\lambda)\mathbf{x}'(\mathbf{T} - (1/k)\mathbf{BT})(\mathbf{T} - (1/k)\mathbf{TB})\mathbf{x}, \tag{1.5}$$

and

$$s_e^2 = \mathbf{x}'[\mathbf{I} - (1/k)\mathbf{B} - (k/vr\lambda)(\mathbf{T} - (1/k)\mathbf{BT})(\mathbf{T} - (1/k)\mathbf{TB})]\mathbf{x}, \tag{1.6}$$

where

$$\mathbf{T} = \mathbf{\Phi\Phi}', \qquad \mathbf{B} = \mathbf{\Psi\Psi}'. \tag{1.7}$$

Let the permutation matrix corresponding to the permutation

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

be S_σ, i.e., $(1\ 2\cdots k)S_\sigma' = (\sigma(1)\sigma(2)\cdots\sigma(k))$. Then it is easy to see that $S_\sigma' = S_{\sigma^{-1}} = S_\sigma^{-1}$.

The experimental units are numbered lexicographically with respect to blocks and the order of units within blocks, i.e., if the fth unit is the ith unit in the pth block, then $f = (p - 1)k + i$. Suppose a random assignment of treatments to units is made in the pth block by means of a permutation σ_p, then, since it is equivalent to the random assignment of units to treatments by means of the inverse permutation σ_p^{-1}, we obtain the following expressions.

$$x_f = \gamma + \sum_{\alpha=1}^{v} \zeta_{\alpha,(p-1)k+\sigma_p^{-1}(i)}\tau_\alpha + \sum_{a=1}^{b} \eta_{af}\beta_a + \pi_f + e_f, \tag{1.8}$$

if $f = (p - 1)k + i$,

or, in vector notation,

$$\mathbf{x} = \gamma\mathbf{1} + \mathbf{U}_\delta'\mathbf{\Phi\tau} + \mathbf{\Psi\beta} + \boldsymbol{\pi} + \mathbf{e}, \tag{1.9}$$

where

$$\mathbf{U}_\delta = \begin{Vmatrix} S_{\sigma_1} & & & 0 \\ & S_{\sigma_2} & & \\ & & \ddots & \\ 0 & & & S_{\sigma_b} \end{Vmatrix}. \tag{1.10}$$

Thus, under this model, the randomization causes the incidence matrix of treatments $\mathbf{\Phi}$ to be a discrete random matrix taking on each value $\mathbf{U}_\sigma'\mathbf{\Phi}$ ($\mathbf{\Phi}$ being arbitrarily fixed) with probability $(k!)^{-b}$, and that is independent of the technical error $\mathbf{e}$ in the sense of the probability.

If we denote the null-distribution of F before the randomization by $P_{H_0}(F \leqq x \mid \mathbf{\Phi})$, then the null-distribution of F after the randomization is given by

$$(1.11) \qquad P_{H_0}(F \leqq x) = (k!)^{-b} \sum_{\sigma_1, \cdots, \sigma_b \varepsilon \mathfrak{S}_k} P_{H_0}(F \leqq x \mid \mathbf{U}_\sigma'\mathbf{\Phi}),$$

where $\mathfrak{S}_k$ stands for the symmetric group of all permutations of degree k.

Let the incidence matrix of the design be $\mathbf{N}$, then $\mathbf{N} = \mathbf{\Phi}'\mathbf{\Psi}$. Since

$$(\mathbf{T} - (1/k)\mathbf{BT})(\mathbf{T} - (1/k)\mathbf{TB}) = r(\mathbf{\Phi} - (1/k)\mathbf{\Psi N}')(\mathbf{\Phi}' - (1/k)\mathbf{N\Psi}')$$

and, under H_0,

$$(\mathbf{\Phi}' - (1/k)\mathbf{N\Psi}')\mathbf{x} = (\mathbf{\Phi}' - (1/k)\mathbf{N\Psi}')\mathbf{e} + (\mathbf{\Phi}' - (1/k)\mathbf{N\Psi}')\boldsymbol{\pi},$$

we have

$$(1.12) \qquad \begin{aligned} s_t^2 = {} & (k/vr\lambda)\mathbf{e}'(\mathbf{T} - (1/k)\mathbf{BT})(\mathbf{T} - (1/k)\mathbf{TB})\mathbf{e} \\ & + 2(k/vr\lambda)\mathbf{e}'(\mathbf{T} - (1/k)\mathbf{BT})(\mathbf{T} - (1/k)\mathbf{TB})\boldsymbol{\pi} \\ & + (k/vr\lambda)\boldsymbol{\pi}'(\mathbf{T} - (1/k)\mathbf{BT})(\mathbf{T} - (1/k)\mathbf{TB})\boldsymbol{\pi}, \end{aligned}$$

and

$$(1.13) \qquad \begin{aligned} s_e^2 = {} & \mathbf{e}'[\mathbf{I} - (1/k)\mathbf{B} - (k/vr\lambda)(\mathbf{T} - (1/k)\mathbf{BT})(\mathbf{T} - (1/k)\mathbf{TB})]\mathbf{e} \\ & + 2\mathbf{e}'[\mathbf{I} - (1/k)\mathbf{B} - (k/vr\lambda)(\mathbf{T} - (1/k)\mathbf{BT})(\mathbf{T} - (1/k)\mathbf{TB})]\boldsymbol{\pi} \\ & + \boldsymbol{\pi}'[\mathbf{I} - (1/k)\mathbf{B} - (k/vr\lambda)(\mathbf{T} - (1/k)\mathbf{BT})(\mathbf{T} - (1/k)\mathbf{TB})]\boldsymbol{\pi}. \end{aligned}$$

Since the matrices $(k/vr\lambda)(\mathbf{T} - (1/k)\mathbf{BT})(\mathbf{T} - (1/k)\mathbf{TB})$ and $\mathbf{I} - (1/k)\mathbf{B} - (k/vr\lambda)(\mathbf{T} - (1/k)\mathbf{BT})(\mathbf{T} - (1/k)\mathbf{TB})$ are idempotent and orthogonal to each other [3], the distributions before the randomization of s_t^2 and s_e^2 are mutually independent ([8], [9]).

The conditional distribution of $\chi_1^2 = s_t^2/\sigma^2$ is the non-central chi-square distribution with degrees of freedom $(v - 1)$ and non-centrality parameter $\lambda_1 = (2\sigma^2)^{-1}(k/vr\lambda)\boldsymbol{\pi}'(\mathbf{T} - (1/k)\mathbf{BT})(\mathbf{T} - (1/k)\mathbf{TB})\boldsymbol{\pi} = (k/2\sigma^2 v\lambda)\boldsymbol{\pi}'\mathbf{T}\boldsymbol{\pi}$ [5]. Thus the distribution of χ_1^2 before the randomization is given by

$$\exp(-\lambda_1) \sum_{\mu=0}^{\infty} \frac{\lambda_1^\mu}{\mu!} \frac{(\chi_1^2/2)^{(v-1)/2+\mu-1}}{\Gamma[(v - 1)/2 + \mu]} \exp(-\chi_1^2/2)\, d(\chi_1^2/2).$$

Similarly, the distribution of $\chi_2^2 = s_e^2/\sigma^2$ before the randomization is given by

$$\exp(-\lambda_2) \sum_{\nu=0}^{\infty} \frac{\lambda_2^\nu}{\nu!} \frac{(\chi_2^2/2)^{(n-v-b+1)/2+\nu-1}}{\Gamma[(n - v - b + 1)/2 + \nu]} \exp(-\chi_2^2/2)\, d(\chi_2^2/2),$$

where $\lambda_2 = (2\sigma^2)^{-1}\boldsymbol{\pi}'[\mathbf{1} - (1/k)\mathbf{B} - (k/vr\lambda)(\mathbf{T} - (1/k)\mathbf{BT})(\mathbf{T} - (1/k)\mathbf{TB})]\boldsymbol{\pi}$

$= (2\sigma^2)^{-1}\boldsymbol{\pi}'\boldsymbol{\pi} - \lambda_1$. Hence the conditional distribution of F under H_0 is given by

$$\exp(-\boldsymbol{\pi}'\boldsymbol{\pi}/2\sigma^2) \sum_{\mu,\nu=0}^{\infty} \frac{\lambda_1^\mu \lambda_2^\nu}{\mu!\nu!} \frac{\Gamma[(n-b)/2+\mu+\nu]}{\Gamma[(v-1)/2+\mu]\cdot\Gamma[(n-v-b+1)/2+\nu]}$$
$$\cdot\left(\frac{v-1}{n-v-b+1}F\right)^{(v-1)/2+\mu-1}$$
$$\cdot\left(1+\frac{v-1}{n-v-b+1}F\right)^{-[(n-b)/2+\mu+\nu]} d\left(\frac{v-1}{n-v-b+1}F\right),$$

and this can be rewritten as

$$\frac{\Gamma[(n-b)/2]}{\Gamma[(v-1)/2]\Gamma[(n-v-b+1)/2]}\left(\frac{v-1}{n-v-b+1}F\right)^{(v-1)/2-1}$$
$$\cdot\left(1+\frac{v-1}{n-v-b+1}F\right)^{-(n-b)/2} d\left(\frac{v-1}{n-v-b+1}F\right)$$
$$(1.14) \qquad \cdot\exp(-\Delta/2\sigma^2)\sum_{l=0}^{\infty}\frac{(\Delta/2\sigma^2)^l}{l!}\left(1+\frac{v-1}{n-v-b+1}F\right)^{-l}$$
$$\cdot\sum_{\mu+\nu=l}\frac{l!}{\mu!\nu!}\theta^\mu(1-\theta)^\nu\left(\frac{v-1}{n-v-b+1}F\right)^\mu$$
$$\cdot\frac{\Gamma[(v-1)/2]\Gamma[(n-v-b+1)/2]\Gamma[(n-b)/2+l]}{\Gamma[(n-b)/2]\Gamma[(v-1)/2+\mu]\Gamma[(n-v-b+1)/2+\nu]},$$

where

$$(1.15) \qquad \Delta = \boldsymbol{\pi}'\boldsymbol{\pi} \quad \text{and} \quad \theta = (k/v\lambda)\Delta^{-1}\boldsymbol{\pi}'\mathbf{T}\boldsymbol{\pi}.$$

After the randomization, the quantity θ in Equation (1.14) is a random variable, and therefore the unconditional distribution of F can be obtained by averaging the probability element given by (1.14) with respect to the permutation distribution of θ due to the randomization.

2. The mean value and the variance of θ with respect to the permutation distribution due to the randomization. We calculate the mean and the variance of the quantity θ given by (1.15) with respect to the permutation distribution due to the randomization. We use the special numbering system of the experimental units mentioned in the preceding section.

Let us write $\pi_f = \pi_i^{(p)}$ if $f = (p-1)k + i$, and let

$$\boldsymbol{\pi}^{(p)\prime} = (\pi_1^{(p)}\pi_2^{(p)}\cdots\pi_k^{(p)}) \quad \text{and} \quad \Delta_p = \boldsymbol{\pi}^{(p)\prime}\boldsymbol{\pi}^{(p)}.$$

Then, $\Delta = \sum_{p=1}^b \Delta_p$, and $\sum_{i=1}^k \pi_i^{(p)} = 0$, $p = 1, \cdots, b$.

Let

$$\boldsymbol{\Phi} = \begin{Vmatrix} \boldsymbol{\Phi}_1 \\ \boldsymbol{\Phi}_2 \\ \vdots \\ \boldsymbol{\Phi}_b \end{Vmatrix}, \quad \text{where} \quad \boldsymbol{\Phi}_p = \begin{Vmatrix} \zeta_{1,(p-1)k+1} & \zeta_{2,(p-1)k+1} & \cdots & \zeta_{v,(p-1)k+1} \\ \zeta_{1,(p-1)k+2} & \zeta_{2,(p-1)k+2} & \cdots & \zeta_{v,(p-1)k+2} \\ \vdots & \vdots & & \vdots \\ \zeta_{1,(p-1)k+k} & \zeta_{2,(p-1)k+k} & \cdots & \zeta_{v,(p-1)k+k} \end{Vmatrix},$$

and

$$\mathbf{T}_{pq} = \|t_{ij}^{pq}\| = \mathbf{\Phi}_p\mathbf{\Phi}_q'.$$

Now

$$(2.1)\quad \mathcal{E}(\theta) = (k/v\lambda)\Delta^{-1}\mathcal{E}(\boldsymbol{\pi}'\mathbf{T}\boldsymbol{\pi}) = (k/v\lambda)(\Delta^{-1}(k!)^{-b})\boldsymbol{\pi}' \sum_{\sigma_1,\cdots,\sigma_b\varepsilon\mathfrak{S}_k} \mathbf{U}_\delta'\mathbf{T}\mathbf{U}_\delta\boldsymbol{\pi}$$

Since

$$\boldsymbol{\pi}' \sum_{\sigma_1,\cdots,\sigma_b} \mathbf{U}_\delta'\mathbf{T}\mathbf{U}_\delta\boldsymbol{\pi} = \sum_{p,q}\sum_{\sigma_1,\cdots,\sigma_b} \boldsymbol{\pi}^{(q)'}\mathbf{S}_{\sigma_p}'\mathbf{T}_{pq}\mathbf{S}_{\sigma_q}\boldsymbol{\pi}^{(q)},\quad \mathbf{S}_{\sigma_p}'\mathbf{T}_{pq}\mathbf{S}_{\sigma_q} = \mathbf{I}_k,\quad \text{if } p = q,$$

and

$$\sum_{\sigma_1,\cdots,\sigma_b} \mathbf{S}_{\sigma_p}'\mathbf{T}_{pq}\mathbf{S}_{\sigma_q} = (k!)^{b-2}\{(k-1)!\}^2\mathbf{G}_k\mathbf{T}_{pq}\mathbf{G}_k,\quad \text{if } p \neq q,$$

it follows that

$$\boldsymbol{\pi}' \sum_{\sigma_1,\cdots,\sigma_b} \mathbf{U}_\delta'\mathbf{T}\mathbf{U}_\delta\boldsymbol{\pi} = (k!)^b \sum_{p=1}^{b} \boldsymbol{\pi}^{(p)'}\boldsymbol{\pi}^{(p)} + (k!)^{b-2}\{(k-1)!\}^2 \sum_{p\neq q} \boldsymbol{\pi}^{(p)'}\mathbf{G}_k\mathbf{T}_{pq}\mathbf{G}_k\boldsymbol{\pi}^{(q)} = (k!)^b\Delta.$$

Hence, we have

$$(2.2)\quad \mathcal{E}(\theta) = k/v\lambda.$$

Next, in order to obtain the variance of θ, we have to obtain an expansion for $\mathcal{E}(\boldsymbol{\pi}'\mathbf{U}_\delta'\mathbf{T}\mathbf{U}_\delta\boldsymbol{\pi})^2$.

$$\begin{aligned}
(\boldsymbol{\pi}'\mathbf{U}_\delta'\mathbf{T}\mathbf{U}_\delta\boldsymbol{\pi})^2 &= (\Delta + \sum_{p\neq q} \boldsymbol{\pi}^{(p)'}\mathbf{S}_{\sigma_p}'\mathbf{T}_{pq}\mathbf{S}_{\sigma_q}\boldsymbol{\pi}^{(q)})^2 \\
&= \Delta^2 + 2\Delta\sum_{p\neq q} \boldsymbol{\pi}^{(p)'}\mathbf{S}_{\sigma_p}'\mathbf{T}_{pq}\mathbf{S}_{\sigma_q}\boldsymbol{\pi}^{(q)} + (\sum_{p\neq q} \boldsymbol{\pi}^{(p)'}\mathbf{S}_{\sigma_p}'\mathbf{T}_{pq}\mathbf{S}_{\sigma_q}\boldsymbol{\pi}^{(q)})^2 \\
&= \Delta^2 + 2\Delta\sum_{p\neq q} \boldsymbol{\pi}^{(p)'}\mathbf{S}_{\sigma_p}'\mathbf{T}_{pq}\mathbf{S}_{\sigma_q}\boldsymbol{\pi}^{(q)} \\
&+ \sum_{p\neq q} [\boldsymbol{\pi}^{(p)'}\mathbf{S}_{\sigma_p}'\mathbf{T}_{pq}\mathbf{S}_{\sigma_q}\boldsymbol{\pi}^{(q)}\boldsymbol{\pi}^{(p)'}\mathbf{S}_{\sigma_p}'\mathbf{T}_{pq}\mathbf{S}_{\sigma_q}\boldsymbol{\pi}^{(q)} + \boldsymbol{\pi}^{(p)'}\mathbf{S}_{\sigma_p}'\mathbf{T}_{pq}\mathbf{S}_{\sigma_q}\boldsymbol{\pi}^{(q)}\boldsymbol{\pi}^{(q)'}\mathbf{S}_{\sigma_q}'\mathbf{T}_{qp}\mathbf{S}_{\sigma_p}\boldsymbol{\pi}^{(p)}] \\
&+ \sum_{p\neq q\neq r} [\boldsymbol{\pi}^{(p)'}\mathbf{S}_{\sigma_p}'\mathbf{T}_{pq}\mathbf{S}_{\sigma_q}\boldsymbol{\pi}^{(q)}\boldsymbol{\pi}^{(p)'}\mathbf{S}_{\sigma_p}'\mathbf{T}_{pr}\mathbf{S}_{\sigma_q}\boldsymbol{\pi}^{(r)} + \boldsymbol{\pi}^{(p)'}\mathbf{S}_{\sigma_p}'\mathbf{T}_{pq}\mathbf{S}_{\sigma_q}\boldsymbol{\pi}^{(q)}\boldsymbol{\pi}^{(q)'}\mathbf{S}_{\sigma_q}'\mathbf{T}_{qr}\mathbf{S}_{\sigma_r}\boldsymbol{\pi}^{(r)}] \\
&+ \sum_{p\neq q\neq r} [\boldsymbol{\pi}^{(p)'}\mathbf{S}_{\sigma_p}'\mathbf{T}_{pq}\mathbf{S}_{\sigma_q}\boldsymbol{\pi}^{(q)}\boldsymbol{\pi}^{(r)'}\mathbf{S}_{\sigma_r}'\mathbf{T}_{rp}\mathbf{S}_{\sigma_p}\boldsymbol{\pi}^{(p)} + \boldsymbol{\pi}^{(p)'}\mathbf{S}_{\sigma_p}'\mathbf{T}_{pq}\mathbf{S}_{\sigma_q}\boldsymbol{\pi}^{(q)}\boldsymbol{\pi}^{(r)'}\mathbf{S}_{\sigma_r}'\mathbf{T}_{rq}\mathbf{S}_{\sigma_q}\boldsymbol{\pi}^{(q)}] \\
&+ \sum_{p\neq q\neq r\neq s} \boldsymbol{\pi}^{(p)'}\mathbf{S}_{\sigma_p}'\mathbf{T}_{pq}\mathbf{S}_{\sigma_q}\boldsymbol{\pi}^{(q)}\boldsymbol{\pi}^{(r)'}\mathbf{S}_{\sigma_r}'\mathbf{T}_{rs}\mathbf{S}_{\sigma_s}\boldsymbol{\pi}^{(s)}.
\end{aligned}$$

Terms in the above expansion which are linear with respect to some $\mathbf{S}_{\sigma_p}$ vanish when their expectations are considered. Hence

$$\begin{aligned}
&\mathcal{E}(\boldsymbol{\pi}'\mathbf{U}_\delta'\mathbf{T}\mathbf{U}_\delta\boldsymbol{\pi})^2 = \Delta^2 \\
(2.3)\quad &+ \sum_{p\neq q} [\mathcal{E}(\boldsymbol{\pi}^{(p)'}\mathbf{S}_{\sigma_p}'\mathbf{T}_{pq}\mathbf{S}_{\sigma_q}\boldsymbol{\pi}^{(q)})^2 + \mathcal{E}(\boldsymbol{\pi}^{(p)'}\mathbf{S}_{\sigma_p}'\mathbf{T}_{pq}\mathbf{S}_{\sigma_q}\boldsymbol{\pi}^{(q)}\boldsymbol{\pi}^{(q)'}\mathbf{S}_{\sigma_q}'\mathbf{T}_{qp}\mathbf{S}_{\sigma_p}\boldsymbol{\pi}^{(p)})] \\
&= \Delta^2 + 2\sum_{p\neq q} \mathcal{E}(\boldsymbol{\pi}^{(p)'}\mathbf{S}_{\sigma_p}'\mathbf{T}_{pq}\mathbf{S}_{\sigma_q}\boldsymbol{\pi}^{(q)})^2.
\end{aligned}$$

Since

$$\pi^{(p)'}\mathbf{S}'_\sigma\mathbf{T}_{pq}\mathbf{S}_\tau\pi^{(q)} = \sum_{i=1}^k t_{ii}^{pq}\pi_{\sigma(i)}^{(p)}\pi_{\tau(i)}^{(q)} + \sum_{i\neq j} t_{ij}^{pq}\pi_{\sigma(i)}^{(p)}\pi_{\tau(j)}^{(q)},$$

it follows that

$$\begin{aligned}(\pi^{(p)'}\mathbf{S}'_\sigma\mathbf{T}_{pq}\mathbf{S}_\tau\pi^{(q)})^2 \\ &= \sum_{i=1}^k t_{ii}^{pq}\pi_{\sigma(i)}^{(p)^2}\pi_{\tau(i)}^{(q)^2} + \sum_{i\neq j} t_{ii}^{pq}t_{jj}^{pq}\pi_{\sigma(i)}^{(p)}\pi_{\sigma(j)}^{(p)}\pi_{\tau(i)}^{(q)}\pi_{\tau(j)}^{(q)} \\ &+ 2\cdot\sum_{l\neq i\neq j} t_{ll}^{pq}t_{ij}^{pq}\pi_{\sigma(l)}^{(p)}\pi_{\sigma(i)}^{(p)}\pi_{\tau(l)}^{(q)}\pi_{\tau(j)}^{(q)} \\ &+ 2\cdot\sum_{i\neq j}[t_{ii}^{pq}t_{ij}^{pq}\pi_{\sigma(i)}^{(p)^2}\pi_{\tau(i)}^{(q)}\pi_{\tau(j)}^{(q)} + t_{jj}^{pq}t_{ij}^{pq}\pi_{\sigma(i)}^{(p)}\pi_{\tau(j)}^{(p)}\pi_{\tau(j)}^{(q)^2}] \\ (2.4)\qquad &+ \sum_{i\neq j}[t_{ij}^{pq}\pi_{\sigma(i)}^{(p)^2}\pi_{\tau(j)}^{(q)^2} + t_{ij}^{pq}t_{ji}^{pq}\pi_{\sigma(i)}^{(p)}\pi_{\sigma(j)}^{(p)}\pi_{\tau(i)}^{(q)}\pi_{\tau(j)}^{(q)}] \\ &+ \sum_{i\neq j\neq l}[t_{ij}^{pq}t_{il}^{pq}\pi_{\sigma(i)}^{(p)^2}\pi_{\tau(j)}^{(q)}\pi_{\tau(l)}^{(q)} + t_{ij}^{pq}t_{li}^{pq}\pi_{\sigma(i)}^{(p)}\pi_{\sigma(l)}^{(p)}\pi_{\tau(i)}^{(q)}\pi_{\tau(j)}^{(q)}] \\ &+ \sum_{i\neq j\neq l}[t_{ij}^{pq}t_{jl}^{pq}\pi_{\sigma(i)}^{(p)}\pi_{\sigma(j)}^{(p)}\pi_{\tau(j)}^{(q)}\pi_{\tau(l)}^{(q)} + t_{ij}^{pq}t_{lj}^{pq}\pi_{\sigma(i)}^{(p)}\pi_{\sigma(l)}^{(p)}\pi_{\tau(j)}^{(q)^2}] \\ &+ \sum_{i\neq j\neq l\neq m} t_{ij}^{pq}t_{lm}^{pq}\pi_{\sigma(i)}^{(p)}\pi_{\sigma(l)}^{(p)}\pi_{\tau(j)}^{(q)}\pi_{\tau(m)}^{(q)}.\end{aligned}$$

Here we have used the fact that, since $t_{ij}^{pq} = 0$ or 1, $t_{ij}^{pq^2} = t_{ij}^{pq}$.

Now,

$$\begin{aligned}\mathcal{E}(\pi_{\sigma(i)}^{(p)^2}\pi_{\tau(i)}^{(q)^2}) &= (1/k^2)\Delta_p\Delta_q, \\ \mathcal{E}(\pi_{\sigma(i)}^{(p)}\pi_{\sigma(j)}^{(p)}\pi_{\tau(i)}^{(q)}\pi_{\tau(j)}^{(q)}) &= [1/k^2(k-1)^2]\Delta_p\Delta_q, \\ \mathcal{E}(\pi_{\sigma(l)}^{(p)}\pi_{\sigma(i)}^{(p)}\pi_{\tau(l)}^{(q)}\pi_{\tau(j)}^{(q)}) &= [1/k^2(k-1)^2]\Delta_p\Delta_q, \\ \mathcal{E}(\pi_{\sigma(i)}^{(p)^2}\pi_{\tau(i)}^{(q)}\pi_{\tau(j)}^{(q)}) &= [-1/k^2(k-1)]\Delta_p\Delta_q, \\ \mathcal{E}(\pi_{\sigma(i)}^{(p)}\pi_{\sigma(j)}^{(p)}\pi_{\tau(l)}^{(q)^2}) &= [-1/k^2(k-1)]\Delta_p\Delta_q, \\ (2.5)\qquad \mathcal{E}(\pi_{\sigma(i)}^{(p)^2}\pi_{\tau(j)}^{(q)^2}) &= (1/k^2)\Delta_p\Delta_q, \\ \mathcal{E}(\pi_{\sigma(i)}^{(p)^2}\pi_{\tau(j)}^{(q)}\pi_{\tau(l)}^{(q)}) &= [-1/k^2(k-1)]\Delta_p\Delta_q, \\ \mathcal{E}(\pi_{\sigma(i)}^{(p)}\pi_{\sigma(l)}^{(p)}\pi_{\tau(i)}^{(q)}\pi_{\tau(j)}^{(q)}) &= [1/k^2(k-1)^2]\Delta_p\Delta_q, \\ \mathcal{E}(\pi_{\sigma(i)}^{(p)}\pi_{\sigma(j)}^{(p)}\pi_{\tau(j)}^{(q)}\pi_{\tau(l)}^{(q)}) &= [1/k^2(k-1)^2]\Delta_p\Delta_q, \\ \mathcal{E}(\pi_{\sigma(i)}^{(p)}\pi_{\sigma(l)}^{(p)}\pi_{\tau(j)}^{(q)^2}) &= [-1/k^2(k-1)]\Delta_p\Delta_q, \\ \mathcal{E}(\pi_{\sigma(i)}^{(p)}\pi_{\sigma(l)}^{(p)}\pi_{\tau(j)}^{(q)}\pi_{\tau(m)}^{(q)}) &= [1/k^2(k-1)^2]\Delta_p\Delta_q.\end{aligned}$$

Thus we have

$$(2.6)\qquad \mathcal{E}(\pi^{(p)'}\mathbf{T}_{pq}\pi^{(q)})^2 = V_{pq}\Delta_p\Delta_q,$$

where

$$
\begin{aligned}
V_{pq} = \frac{1}{k^2} \sum_{i,j=1}^{k} t_{ij}^{pq} + \frac{1}{k^2(k-1)^2} [\sum_{i \neq j} t_{ii}^{pq} t_{jj}^{pq} + 2 \sum_{l \neq i \neq j} t_{ll}^{pq} t_{ij}^{pq} \\
+ \sum_{i \neq j} t_{ij}^{pq} t_{ji}^{pq} + \sum_{i \neq j \neq l} (t_{ji}^{pq} t_{li}^{pq} + t_{ij}^{pq} t_{jl}^{pq}) + \sum_{i \neq j \neq l \neq m} t_{ij}^{pq} t_{lm}^{pq}] \\
- \frac{1}{k^2(k-1)} [2 \sum_{i \neq j} t_{ii}^{pq} t_{ij}^{pq} + 2 \sum_{i \neq j} t_{jj}^{pq} t_{ij}^{pq} + \sum_{i \neq j \neq l} (t_{ij}^{pq} t_{il}^{pq} + t_{ij}^{pq} t_{lj}^{pq})].
\end{aligned} \tag{2.7}
$$

It can be assumed without loss of generality, in the fixed configuration, that, if $p \neq q$, the units bearing the same number within the pth block and the qth block do not receive the same treatment. Thus, for $p \neq q$, $\operatorname{tr}\mathbf{T}_{pq} = 0$, $\sum_{i,j=1}^{k} = t_{ij}^{pq} = \sum_{i \neq j} t_{ij}^{pq} = \lambda_{pq}$, the number of treatments common to the pth and qth blocks, and $t_{ii}^{pq} t_{ij}^{pq} = t_{jj}^{pq} t_{ij}^{pq} = 0$, if $i \neq j$. Therefore, we have

(i) For the first term of (2.7): $\sum_{i,j} t_{ij}^{pq} = \lambda_{pq}$.

(ii) For the second term of (2.7): $\sum_{i \neq j} t_{ii}^{pq} t_{jj}^{pq} = 0$, $\sum_{l \neq i \neq j} t_{ll}^{pq} t_{ij}^{pq} = 0$, and

$$
\begin{aligned}
\sum_{i \neq j} t_{ij}^{pq} t_{ji}^{pq} + \sum_{i \neq j \neq l} (t_{ij}^{pq} t_{li}^{pq} + t_{ij}^{pq} t_{jl}^{pq}) + \sum_{i \neq j \neq l \neq m} t_{ij}^{pq} t_{lm}^{pq} \\
= \sum_{i \neq j} t_{ij}^{pq} [t_{ji}^{pq} + \sum_{l \neq (i,j)} (t_{li}^{pq} + t_{jl}^{pq}) + \sum_{l \neq m \neq (i,j)} t_{lm}^{pq}] \\
= \sum_{i \neq j} t_{ij}^{pq} [t_{ji}^{pq} + \sum_{l} (t_{li}^{pq} + t_{jl}^{pq}) - t_{ji}^{pq} - t_{ji}^{pq} + \sum_{l,m} t_{lm}^{pq} \\
- \sum_{l} (t_{li}^{pq} + t_{jl}^{pq}) - \sum_{l} (t_{il}^{pq} + t_{lj}^{pq}) + t_{ij}^{pq} + t_{ji}^{pq}] \\
= \sum_{i \neq j} t_{ij}^{pq} [\sum_{l,m} t_{lm}^{pq} - \sum_{l} (t_{il}^{pq} + t_{lj}^{pq}) + t_{ij}^{pq}] = \lambda_{pq}^2 - \lambda_{pq},
\end{aligned}
$$

because

$\sum_{l} t_{il}^{pq} = \delta_{i*}^{pq} = 1$, if the treatment in the ith unit of the pth block is also contained in the qth block,
$= 0$, otherwise,

and similarly

$\sum_{l} t_{lj}^{pq} = \delta_{*j}^{pq} = 1$, if the treatment in the jth unit of the qth block is also contained in the pth block,
$= 0$, otherwise,

and hence $\sum_{i} \delta_{i*}^{pq} = \sum_{j} \delta_{*j}^{pq} = \lambda_{pq}$.

(iii) For the third term of (2.7): $\sum_{i \neq j} t_{ii}^{pq} t_{ij}^{pq} = \sum_{i \neq j} t_{jj}^{pq} t_{ij}^{pq} = 0$, and

$$
\begin{aligned}
\sum_{i \neq j \neq l} (t_{ij}^{pq} t_{il}^{pq} + t_{ij}^{pq} t_{lj}^{pq}) = \sum_{i \neq j} t_{ij}^{pq} [\sum_{l} (t_{il}^{pq} + t_{lj}^{pq}) - t_{ij}^{pq} - t_{ij}^{pq}] \\
= \sum_{i \neq j} t_{ij}^{pq} [\delta_{i*}^{pq} + \delta_{*j}^{pq} - 2 t_{ij}^{pq}] = 2\lambda_{pq} - 2\lambda_{pq} = 0.
\end{aligned}
$$

Consequently we obtain

$$V_{pq} = \lambda_{pq}/k^2 + (\lambda_{pq}^2 - \lambda_{pq})/k^2(k-1)^2. \tag{2.8}$$

Finally

$$V(\theta) = 2(k/v\lambda)^2 \cdot (1/(k-1))W, \tag{2.9}$$

where

$$W = \Delta^{-2}[\Lambda_2/(k^2(k-1)) + \{(k-2)/(k(k-1))\}\Lambda_1 - \sum_{p=1}^{b} \Delta_p^2], \tag{2.10}$$

with

$$\Lambda_1 = \sum_{p,q} \lambda_{pq}\Delta_p\Delta_q \quad \text{and} \quad \Lambda_2 = \sum_{p,q} \lambda_{pq}^2\Delta_p\Delta_q\,. \tag{2.11}$$

For a randomized (complete) block design, i.e., $k = v$, $r = b$, $\lambda = b$, it is known that $\lambda_{pq} = k$. Hence

$$W = \Delta^{-2}(\Delta^2 - \sum_p \Delta_p^2) = [(b-1)/b](1 - V/b),$$

where

$$V = (\Delta/b)^{-2}(b-1)^{-1} \sum_p (\Delta_p - \Delta/b)^2.$$

Therefore,

$$V(\theta) = 2(b-1)b^{-3}(k-1)^{-1}(1 - V/b), \tag{2.12}$$

as shown in [4], [10], [11], [12] and [14].

3. Approximate null-distribution of the F-statistic of the randomized incomplete block design under the Neyman model.

Since $0 \leqq \theta \leqq 1$, we may fit a beta distribution

$$\{\Gamma[(\nu_1 + \nu_2)/2]/\Gamma(\nu_1/2)\Gamma(\nu_2/2)\}\theta^{\nu_1/2-1}(1-\theta)^{\nu_2/2-1}d\theta \tag{3.1}$$

to the permutation distribution of θ by equating the first two moments, i.e., choose ν_1 and ν_2 such that

$$\frac{\nu_1}{\nu_1+\nu_2} = \frac{k}{v\lambda}, \; \frac{2\nu_1\nu_2}{(\nu_1+\nu_2)^2(\nu_1+\nu_2+2)} = 2\frac{k^2}{v^2\lambda^2}\frac{W}{k-1}\,.$$

This gives us

$$\nu_1 = (v-1)\phi, \qquad \nu_2 = (n - v - b + 1)\phi, \tag{3.2}$$

where

$$\phi = W^{-1}(v\lambda - k)/(vr) - 2k/(vr(k-1)). \tag{3.3}$$

If, in particular, the variances of unit effects within blocks are uniform, i.e.,

$\Delta_p = \Delta_0$, $p = 1, 2, \cdots, b$, and hence $\Delta = b\Delta_0$, then, it turns out that

$$W = \frac{1}{b^2}\left[\frac{1}{k^2(k-1)}\sum_{p,q}\lambda_{pq}^2 + \frac{k-2}{k(k-1)}\sum_{p,q}\lambda_{pq} - b\right]. \tag{3.4}$$

Now, since $\sum_{p,q}\lambda_{pq}^2 = \mathrm{tr}(\mathbf{N}'\mathbf{N})^2 = \mathrm{tr}(\mathbf{N}\mathbf{N}')^2 = vr^2 + v(v-1)\lambda^2$, and $\sum_{p,q}\lambda_{pq} = \mathbf{1}'\mathbf{N}'\mathbf{N}\mathbf{1} = vr^2$, it follows that

$$W = [r + \lambda(k-1) + rk(k-2) - k(k-1)]/(bk(k-1)).$$

Thus,

$$\phi = W^{-1}(v\lambda - k)/(vr) - 2k/(vr(k-1)) = 1 - 2/(b(k-1)). \tag{3.5}$$

Therefore, if the variances of the unit effects within blocks are nearly uniform and the number of blocks is sufficiently large, then $\phi \approx 1$. In other words, in such circumstances we may take the beta-distribution

$$\frac{\Gamma[(n-b)/2]}{\Gamma[(v-1)/2]\Gamma[(n-v-b+1)/2]}\theta^{(v-1)/2-1}(1-\theta)^{(n-v-b+1)/2-1}d\theta \tag{3.6}$$

as an approximation to the permutation distribution of θ due to the randomization.

Taking the expectation of (1.14) with respect to (3.6), we have the approximate unconditional distribution of F-statistic as follows:

$$\begin{aligned}&\frac{\Gamma[(n-b)/2]}{\Gamma[(v-1)/2]\Gamma[(n-v-b+1)/2]}\left(\frac{v-1}{n-v-b+1}F\right)^{(v-1)/2-1} \\ &\qquad\cdot\left(1+\frac{v-1}{n-v-b+1}F\right)^{-(n-b)/2} d\left(\frac{v-1}{n-v-b+1}F\right),\end{aligned} \tag{3.7}$$

which is the central F-distribution with degrees of freedom $(v - 1, n - v - b + 1)$ obtained under the familiar normal theory assumptions.

Acknowledgment. The author expresses his appreciation to Professor J. L. Hodges, Jr., the editor of the Annals of Mathematical Statistics, the referee and the associate editor for their comments and suggestions which were helpful in revising the paper.

REFERENCES

[1] Bose, R. C. (1949). Least square aspects of analysis of variance. Inst. Statist. Mimeo. Series 6. Univ. of North Carolina.

[2] Fisher, R. A. (1926). The arrangement of field experiments. *J. Ministry of Agric.* **33** 503–513.

[3] James, A. T. (1957). The relationship algebra of an experimental design. *Ann. Math. Statist.* **28** 993–1002.

[4] Kempthorne, O. (1952). *The Design and Analysis of Experiments.* Wiley, New York, 142.

[5] Mann, H. B. (1949). *Analysis and Design of Experiments.* Dover, New York.

[6] McCarthy, M. D. (1939). On the application of the z-test to randomized blocks. *Ann. Math. Statist.* **10** 337–359.
[7] Neyman, J. with the cooperation of K. Iwaskiewicz and S. Kolodziejczyk (1935). Statistical problems in agricultural experimentation. *J. Roy. Statist. Soc. Suppl.* **2** 107–154. Discussions pp. 154–180.
[8] Ogawa, J. (1949). On the independence of bilinear and quadratic forms of a random sample from a normal population, *Ann. Inst. Statist. Math.* **1** 83–108.
[9] Ogawa, J. (1950). On the independence of quadratic forms in a non-central normal system. *Osaka Math. J.* **2** 151–159.
[10] Ogawa, J. (1961). The effect of randomization on the analysis of randomized block design. *Ann. Inst. Statist. Math.* **13** 105–117.
[11] Pitman, E. J. G. (1937). Significance tests which can be applied to samples from any populations, III. The analysis of variance test. *Biometrika* **42** 70–79.
[12] Welch, B. L. (1937). On the z-test in randomized blocks and Latin squares. *Biometrika* **29** 21–52.
[13] Wilk, M. B. (1955). Linear models and randomized experiments. Ph.D. thesis, Iowa State College.
[14] Wilk, M. B. (1955). The randomization analysis of a generalized randomized block design. *Biometrika* **42** 70–79.