Note on the Determination of the Replication Number for the Slippage Problem in r-way Layout

Shirafuji, Michie
Kyushu University

https://doi.org/10.5109/12976

出版情報：統計数理研究. 7 (1/2), pp.46-51, 1956-12. Research Association of Statistical Sciences

NOTE ON THE DETERMINATION OF THE REPLICATION NUMBERS FOR THE SLIPPAGE PROBLEM IN r-WAY LAYOUT

By

Michie SHIRAFUJI

(*Received May.* 15, 1956)

§ 1. **Summary.** In an r-way layout of the parametric type, the m-th observation ($m=1, 2, \cdots, n$) corresponding to the combination of the levels t_j ($1 \leqq t_j \leqq l_j$) of the j-th factor ($j=1, 2, \cdots, r$) is given by

$$(1.01)\qquad x_{t_1 t_2 \cdots t_r m} = \mu + \sum_{\alpha=1}^{r} \sum_{1, 2, \cdots, \alpha} \mu(i_1, i_2, \cdots, i_\alpha ; t_{i_1}, t_{i_2}, \cdots, t_{i_\alpha}) + \varepsilon_{t_1 t_2 \cdots t_r m} ,$$

where $\sum_{(1, 2, \cdots, \alpha)} \mu(i_1, i_2, \cdots, i_\alpha ; t_{i_1}, t_{i_2}, \cdots, t_{i_\alpha})$ denote the summation over all combinations of $(i_1, i_2, \cdots, i_\alpha)$ chosen from $1, 2, \cdots, r$ with $i_1 < i_2 < \cdots < i_\alpha$, $\mu(i_1, i_2, \cdots, i_\alpha ; t_{i_1}, t_{i_2}, \cdots, t_{i_\alpha})$ are the constants satisfying the conditions

$$(1.02)\qquad \sum_{t_{i_j}=1}^{l_{i_j}} \mu(i_1, i_2, \cdots, i_\alpha ; t_{i_1}, t_{i_2}, \cdots, t_{i_\alpha}) = 0,$$

and random errors $\varepsilon_{t_1 t_2 \cdots t_r m}$ are normally independently distributed with the mean value 0 and the common unknown variance σ^2. We are concerned with a choice of the combination of the levels which will give the maximum effect for each assigned set of $(i_1, i_2, \cdots, i_\alpha)$, that is, to find out the combination $(t^*_{i_1}, t^*_{i_2}, \cdots, t^*_{i_\alpha})$ which will give us the maximum value of $\mu(i_1, i_2, \cdots, i_\alpha ; t_{i_1}, t_{i_2}, \cdots, t_{i_\alpha})$, where $(t_{i_1}, t_{i_2}, \cdots, t_{i_\alpha})$ run through $1 \leqq t_{i_j} \leqq l_j$, $j=1, 2, \cdots, \alpha$. Let $A(k_1, k_2, \cdots, k_\beta ; a_{k_1}, a_{k_2}, \cdots, a_{k_\beta})$ be the maximum likelihood estimate of $\mu(k_1, k_2, \cdots, k_\beta ; a_{k_1}, a_{k_2}, \cdots, a_{k_\beta})$ and let us define

$$(1.03)\qquad x(i_1, i_2, \cdots, i_\alpha ; t_{i_1}, t_{i_2}, \cdots, t_{i_\alpha}) = \sum_{\beta=0}^{\alpha} \sum_{i_1, i_2, \cdots, i_\alpha} A(k_1, k_2, \cdots, k_\beta ; a_{k_1}, a_{k_2}, \cdots, a_{k_\beta}) ,$$

whose detailed formulation can be seen in H. B. MANN [4].

We postulate the following statistical procedure; if

$$(1.04)\qquad \max_{1 \leqq t_{i_j} \leqq l_{i_j} (j=1, 2, \cdots, \alpha)} x(i_1, i_2, \cdots, i_\alpha ; t_{i_1}, t_{i_2}, \cdots, t_{i_\alpha}) = x(i_1, i_2, \cdots, i_\alpha ; t^*_{i_1}, t^*_{i_2}, \cdots, t^*_{i_\alpha}) ,$$

then we choose $(t^*_{i_1}, t^*_{i_2}, \cdots, t^*_{i_\alpha})$. This procedure corresponds to the one justified by R. R. BAHADURE [1] in his paper concerning a problem of k populations.

Under this postulate, our problem is to find out a suitable number of replication n. In this paper, the risk functions are set up which take into consideration the loss of making a wrong decision and the cost of repetition. We are able to find a maximum of the risk function over all possible parameter values. It is then possible to minimize this maximum risk with respect to the sample size. It is to be noted that P. N. SOMERVILLE [7] takes the similar attitude in determing sample size.

§ 2. **The case where we are concerned merely with main effects.** Here we shall adopt the risk function $R(n, \delta)$ defined as follows;

$$R(n, \delta) = n + \sum_{j=1}^{r} \frac{1}{\lambda_j} \sum_{t_j \neq t_j^*} \delta(j;\, t_j) P(j;\, t_j)\,, \tag{2.01}$$

where

$$\delta(j;\, t_j) = \{\mu(j;\, t_j^*) - \mu(j;\, t_j)\}/\sigma\,, \tag{2.02}$$

$$P(j;\, t_j) = \Pr\{\max_{1 \leqq t_{j_k} \leqq l_j,\ t_{j_k} \neq t_j} x(j;\, t_{j_k}) = x(j;\, t_j)\}\,, \tag{2.03}$$

$$\lambda_j = a/b_j\,, \tag{2.04}$$

where a is the cost of repetition, b_j the loss of making a wrong decision in the j-th state. In what follows the ratios λ_j are assumed to be known to us, but all $\delta(j;\, t_j)$ are unknown to us. Now n is not a continuous variable, but an integral valued one. In what follows, however, we shall treat n as if it were a continuous variable so that we may get approximate results enough for practical applications.

Theorem 1. *For the case when $l_j = 2$, that is, the 2^r design with n repetitions $2^r \times (n)$, n^* minimizing the maximum risk with respect to Ω: $\{\delta(j;\, t_j),\ t_j = 1, 2,\ j = 1, 2, \cdots, r\}$ is given by*

$$n^* = \left(\frac{0.1202}{2^{(r-1)/2}} \sum_{j=1}^{r} \frac{1}{\lambda_j}\right)^{2/3} (1 + 63 \times 10^{-4}\theta_1)^{2/3} \equiv n_0 (1 + 63 \times 10^{-4}\theta_1)^{2/3}\,, \quad |\theta_1| < 1\,, \tag{2.05}$$

and in virtue of this solution we have

$$\min_n \max_{\delta \in \Omega} R(n, \delta) = R(n_0, \delta_0) + 0.0030\,\theta_2 \left(2^{-(r-1)/2} \sum_{j=1}^{r} \frac{1}{\lambda_j}\right)^{2/3}\,, \quad |\theta_2| < 1\,, \tag{2.06}$$

where

$$R(n_0, \delta_0) = 0.7307 \left(2^{-(r-1)/2} \sum_{j=1}^{r} \frac{1}{\lambda_j}\right)^{2/3}. \tag{2.07}$$

Proof. Putting $t_j^* = 1$, $j = 1, 2, \cdots, r$, the risk function $R(n, \delta)$ is given by the following

$$R(n, \delta) = n + \sum_{j=1}^{r} \frac{\delta(j;2)}{\lambda_j} \Phi(-\sqrt{2^{(r-2)}n}\,\delta(j;2)), \tag{2.08}$$

where

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\{-t^2/2\}\, dt \equiv \int_{-\infty}^{x} \varphi(t)\, dt. \tag{2.09}$$

For any fixed n, the maximum point of $R(n, \delta)$ can be found by ordinary calculus. Since

$$\frac{\partial R(n, \delta)}{\partial \delta(j;2)} = 0, \quad j = 1, 2, \cdots, r, \tag{2.10}$$

we have

$$\sqrt{2^{(r-2)}n}\,\delta(j;2) = \Phi(-\sqrt{2^{(r-2)}n}\,\delta(j;2))/\varphi(\sqrt{2^{(r-2)}n}\,\delta(j;2)). \tag{2.11}$$

This equation may be solved approximately with the aid of a normal integral table [5], which gives us

$$\delta^*(j;2) = \delta_0(j;2) + 10^{-4}\theta_3, \quad |\theta_3| < 1, \tag{2.12}$$

where

$$\delta_0(j;2) = 0.7518 \times 2^{-(r-2)/2} n^{-1/2}. \tag{2.13}$$

Since

$$\frac{\partial^2 R(n, \delta)}{\partial \delta(j;2)\partial \delta(i;2)} = \begin{cases} 0 & (j \neq i), \\ \lambda_j^{-1}(2^{(r-2)}n\delta(j;2)^2 - 2\sqrt{n})\varphi(\sqrt{n}\,\delta(j;2)) & (j = i), \end{cases} \tag{2.14}$$

and

$$\left[\frac{\partial^2 R(n, \delta)}{\partial \delta(j;2)^2}\right]_{\delta(j;2)=\delta_0(j;2)} < 0, \quad n \neq 0, \tag{2.15}$$

(2.08) is maximized by (2.12). Substituting (2.12) in (2.08), we have

$$R(n, \delta^*) = R(n, \delta_0) + \frac{15 \times 10^{-4}\theta_4}{\sqrt{2^{(r-1)}n}} \sum_{j=1}^{r} \frac{1}{\lambda_j}, \quad |\theta_4| < 1, \tag{2.16}$$

where

$$R(n, \delta_0) = n + \frac{0.2404}{\sqrt{2^{(r-1)}n}} \sum_{j=1}^{r} \frac{1}{\lambda_j}. \tag{2.17}$$

Our problem is to find out n^* which minimizes the risk given by (2.16). This can be done by solving the equation

(2.18) $$\frac{\partial R(n, \delta^*)}{\partial n} = 0,$$

and we are lead to (2.05).

From (2.12), (2.13) and (2.05) we have (2.06) and (2.07).

Remark 1. For the case when $r=1$ our result is coincident with that obtained by I. BROSS [2] and G. TAGUCHI [9].

Theorem 2. *For the case when* $\delta(j\,;\,t_j) \equiv \delta_j$ $(1 \leq t_j \leq l_j,\ j=1, 2, \cdots, r)$, n^* *which minimizes the maximum risk respect to* Ω: $(\delta_1, \delta_2, \cdots, \delta_r)$ *is given by*

(2.19) $$n^* = \left(\frac{1}{\sqrt{2}\sqrt{l_1 l_2 \cdots l_r}} \sum_{j=1}^{r} \frac{\sqrt{l_j} M_{l_j-1}}{\lambda_j}\right)^{2/3},$$

and for this solution we have

(2.20) $$\min_n \max_{\delta \in \Omega} R(n, \delta) = 3 \times 2^{-1/3} \left(\frac{1}{\sqrt{l_1 l_2 \cdots l_r}} \sum_{j=1}^{r} \frac{\sqrt{l_j} M_{l_j-1}}{\lambda_j}\right)^{2/3},$$

where

(2.21) $$M_{l_j-1} \equiv \max_{0 \leq \delta_j < \infty} \delta_j \{1 - P(j\,;\,t_j^*)\} \sqrt{l_1 l_2 \cdots l_r n} / \sqrt{2 l_j},$$

(2.22) $$P(j\,;\,t_j^*) = \sum_{k=0}^{\infty} \frac{(2k)!}{k!} \left(\frac{V_j^2}{4}\right)^k \left(\frac{V_j^2}{2}\right)^{k+l_j-l} B_{l_j-1, 2k}\left(\frac{\sqrt{2}\,\delta_j}{V_j^2}, \cdots, \frac{\sqrt{2}\,\delta_j}{V_j^2}\right),$$

(2.23) $$B_{n,k}(x_1, x_2, \cdots, x_n) = \sum_{i=0}^{k} b_i(x_n) B_{n-1, k-i}(x_1, x_2, \cdots, x_{n-1}),$$

(2.24) $$b_i(x) = \frac{1}{i!} \Phi^{(i)}(x),$$

and

(2.25) $$V_j^2 = \sigma^2(2 l_j) / (l_1 l_2 \cdots l_r n).$$

Proof. We can give the proof of (2.19), (2.20) and (2.21) in the same way as to that of P. N. SOMERVILLE [7]. The values of M_{l_j-1} for $l_j = 2, 3,$

Table 1. Value of n^* in (2.19) for $l^r \times (n)$-design
($l_j = l$, $\lambda_j = \lambda$, $j = 1, 2, 3, 4, 5$)

	$\lambda = 0.01$				
l \ r	1	2	3	4	5
2	5.3	6.6	6.9	6.6	6.1
3	7.0	7.7	7.0	5.9	4.7
4	8.1	8.1	6.7	5.1	3.7
5	8.9	8.3	6.3	4.5	3.1
6	9.5	8.3	6.0	4.0	2.6

	$\lambda = 0.05$				
l \ r	1	2	3	4	5
2	1.8	2.3	2.4	2.3	2.1
3	2.4	2.6	2.4	2.9	1.6
4	2.8	2.8	2.3	1.8	1.3
5	3.1	2.8	2.2	1.5	1.0
6	3.3	2.8	2.1	1.5	0.9

4, 5, 6 were calculated by him. On the other hand, we have (2.22), (2.23), (2.24) and (2.25) by the expansion method of A. KUDÔ [3].

In order to calculate the (2.19), we made use of the value of M_{l_j-1} for $l_j=2, 3, 4, 5, 6$ calculated by P.N. SOMERVILLE [7].

§3. The case where we are concerned with a interaction. By the similar way to § 2, we have

Theorem 3. *For the case when* $\delta(j, k; t_j, t_k)\equiv\delta_{jk}$ $(1\leqq t_j\leqq l_j,\ 1\leqq t_k\leqq l_k,\ j\neq k)$, *the risk function*

$$R(n, \delta)=n+\lambda_{jk}^{-1}\delta_{jk}\{1-P(j, k;\ t_j^*, t_k^*)\}, \tag{3.01}$$

where

$$\delta(j, k;\ t_j, t_k)=\{\mu(j, k;\ t_j^*, t_k^*)-\mu(j, k;\ t_j, t_k)\}/\sigma, \tag{3.02}$$

$$\lambda_{jk}=a/b_{jk}, \tag{3.03}$$

where a, b_{jk} *are the ones defined by* (2.04) *with* j *replaced by* (j, k) *and*

$$\begin{aligned} P(j, k;\ t_j^*, t_k^*)=\int_{-\infty}^{\infty}\Bigg\{&\sum_{n=0}^{\infty}\sum_{r=0}^{n}\frac{{}_nC_r}{n!}\left(\frac{1}{l_k-1}\right)^{r}\left(\frac{1}{l_j-1}\right)^{n-r}\sum_{\Sigma_{(t_j, t_k)\neq(t_j^*, t_k^*)}\nu_{t_jt_k}=2r}\\ &\sum_{\Sigma_{(t_k\neq t_k^*)}\mu_{t_k}=n-r}\ \sum_{\Sigma_{(t_j\neq t_j^*)}\xi_{t_j1}=2\mu_1}\cdots\cdots\sum_{\Sigma_{(t_j\neq t_j^*)}\xi_{t_jt_k^*-1}=2\mu_{t_k^*-1}}\\ &\sum_{\Sigma_{(t_j\neq t_j^*)}\xi_{t_jt_k^*+1}=2\mu_{t_k^*+1}}\cdots\cdots\sum_{\Sigma_{(t_j\neq t_j^*)}\xi_{t_jt_k}=2\mu_{t_k}}\left(1+\frac{1}{l_k-1}+\frac{1}{l_j-1}\right)^{-n+\{(l_jl_k-1)/2\}}\\ &\times\prod_{(t_j, t_k)\neq(t_j^*, t_k^*)}\Phi^{(\nu_{t_jt_k}+\xi_{t_jt_k})}\left(y-\frac{\delta_{jk}}{V_{jk}}\right)\Bigg\}\varphi(y)\,dy, \end{aligned} \tag{3.04}$$

$$V_{jk}^2=\sigma^2(l_j-1)(l_k-1)/(l_1l_2\cdots l_rn). \tag{3.05}$$

It should be noticed here that the evaluation of the integral is possible by making use of the formula

$$\int_{-\infty}^{\infty}f(x)\exp\{-x^2\}\,dx=\sum_{i=1}^{n}\alpha_i^{(n)}f(x_i^{(n)})+R_n, \tag{3.06}$$

where

$$R_n=\frac{n!}{2^n(2n)!}\pi^{1/2}f^{(2n)}(\xi) \tag{3.07}$$

for some ξ, $-\infty<\xi<\infty$ [8], and $x_i^{(n)}$ and $\alpha_i^{(n)}$ are the zero points and the weight factors of the Hermitian Polynomials respctively, and these values are tabulated by H. E. SALZAR and others [6].

In the case where we are concerned with a k-th interaction, we have the similar theorem. But we shall omit the statement as it is very complicated.

The author would like to express her thanks to Professor T. KITAGAWA for the suggestions and criticism given by him.

KYUSHU UNIVERSITY

References

[1] R. R. BAHADUR; *On a problem in the theory of k populations.* Ann. Math. Stat., **21** (1950), 362-375.

[2] I. BROSS; *Two-choice selection.* Jur. Amer. Statist. Ass., **45** (1950), 530-540.

[3] A. KUDÔ; *On the distribution of the maximum value of an equally correlated sample from a normal population.* Submitted to Sankhyā.

[4] H. B. MANN; Analysis and design of experiments. New York, 1949.

[5] K. PEARSON; *Tables for Statisticians and Biometricians.* Vol. II. Table II. Cambridge. 1931.

[6] H. E. SALZAR, R. ZUCKER and R. CAPUANO; *Table of the zeros and weight factors of the first twenty Hermite Polinamials.* Jour. Res. Nat. Bur. Standard., **48** (1952), 111-116.

[7] P. N. SOMERVILLE; *Some problems of optimum sampling.* Biometrika, **41** (1941), 420-429.

[8] G. SZEGÖ; Orthogonal polynomials. Amer. Math. Soc. Colloq. Pub., **23** (1939).

[9] G. TAGUCHI; *Note on design of experiments* (7), (in Japanese) Statistical Quality Control, **4** (1953), 231-236.