Listy Biometryczne - Biometrical Letters
Vol. 35(1998), No. 2, 117-125

Determination of the size of an experiment in the analysis of variance for fixed and mixed models*

Dieter Rasch and Minghui Wang

Sub-department of Mathematics, Department AMST,
Wageningen Agricultural University,
Dreijenlaan 4, 6703HA Wageningen, The Netherlands
e-mail: Dieter.Rasch@wts.wk.wau.nl, Minghui.Wang@wts.wk.wau.nl

SUMMARY

An algorithm and a program for the determination of the minimal size of experiments to compare levels of a fixed factor for balanced classifications in the analysis of variance is described and demonstrated. The minimal size is determined in dependence on a lower bound for the difference between the maximum and the minimum of the effects to be tested for equality by an F-test, on the risks of the test and on a presumable value of the common residual variance. We determine the minimal size for the least favorable case (the maxi-min size) as well as for the most favorable case (the mini-min size). The classification and the model chosen influence the degrees of freedom for the F-test and the relationship between the size of the experiment and the non-centrality parameter.

KEY WORDS: analysis of variance, minimal size, non-centrality parameter, sub-class number.

1. Introduction

Let us assume that an experimenter decided that his experimental question leads to an F-test in an ANOVA model with error variance σ^2. After fixing the number of factors, their properties (fixed or random) and the type of their combination (crossed or nested) he usually has to answer the question "What is a suitable subclass number n?". To answer this question, knowledge of the error variance σ^2 is required, although the experiment has not yet been performed. From a previous work on this or from

*The paper was submitted on the occasion of 70-th birthday of Professor Tadeusz Caliński.

the literature, the experimenter guesses a value of σ^2. He also must state how precise the F-test should be. He thus has to fix the significance level and the minimal value $1-\beta$ of the power function if the maximum difference between effects to be tested for equality equals d. We assume that at least one of the factors is fixed which means that we have either

– a model I of ANOVA (all factors are fixed)

– or a mixed model.

The procedures are demonstrated for some examples with up to three factors in total.

2. Size of an experiment for the F-test in ANOVA models

The problem of the determination of the size of an experiment for the analysis of variance has amongst others been investigated by Tang (1938), Thompson (1941), Lehmer (1944), Pearson and Hartley (1951, 1972), Fox (1956), Tiku (1967, 1962), Das Gupta (1968), Bratcher et al. (1970), Kastenbaum et al. (1970a, b), Bowman (1972, 1975), Rasch et al. (1996b), Herrendörfer et al. (1997) and Rasch (1998).

The solution of the following equation plays a crucial role:

$$F(f_1, f_2, 0, 1-\alpha) = F(f_1, f_2, \lambda, \beta), \tag{1}$$

where $F(f_1, f_2, 0, 1-\alpha)$ is the $(1-\alpha)$-quantile of the (central) F-distribution with degrees of freedom f_1 and f_2 and non-centrality parameter 0 and where $F(f_1, f_2, \lambda, \beta)$ is the β quantile of the F-distribution with degrees of freedom f_1 and f_2 and non-centrality parameter λ. Below we determine the minimum size of an experiment for testing a special group of effects (main effects, interaction effects) for equality for the least favourable case which we refer to maxi-min size and for the most favourable case that hereafter we refer to mini-min size.

Let $E_{\min}$ be the minimum and $E_{\max}$ be the maximum of a set of q effects $E_1, \ldots, E_q$. The risk of the first kind for the F-test of the null hypothesis H_0: $E_1 = E_2 = \ldots = E_q$ is fixed by α and the power function must have at least the value $1-\beta$ if $E_{\max} - E_{\min} \geq d$, with d as part of our precision requirements. The non-centrality parameter depends not only on $E_{\max}$ and $E_{\min}$ but also on all other E_i. If $E_{\max} - E_{\min} \geq d$ the non-centrality parameter $w\sum_{i=1}^{q}(E_i - \bar{E})^2/\sigma^2$ satisfies

$$w\sum_{i=1}^{q}(E_i - \bar{E})^2/\sigma^2 \geq w(E_{\max} - E_{\min})^2/(2\sigma^2) \geq wd^2/(2\sigma^2). \tag{2}$$

The least favorable case (leading to the minimal non-centrality parameter and the maxi-min size) is the case when the $q-2$ remaining E_i are equal to $(E_{\max} + E_{\min})/2$.

The most favorable case leading to the mini-min size for even $q = 2m$ is given if m of the E_i are equal to $E_{\min}$ and the other m of the E_i are equal to $E_{\max}$. For odd $q = 2m + 1$, where m is a positive integer, the most favorable case is again given if m of the E_i are equal to $E_{\min}$, another m of the E_i are equal to $E_{\max}$ and the remaining E_i equals to either $E_{\min}$ or $E_{\max}$. Our objective is to determine the minimal size $N = v \cdot w$ of the experiment in such a way that a given number w depending on the classification satisfies

$$\frac{wd^2}{2\sigma^2} \geq g(\alpha, \beta, f_1, f_2), \tag{3}$$

where g is the solution λ of (2). The values of w are shown in Table 1.

Below we demonstrate, for even $q = 2m$, the least and most favorable case and the inequality

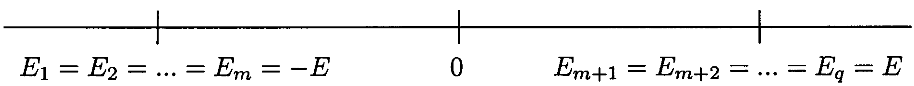
$$\sum_{i=1}^{q} (E_i - \bar{E})^2 \geq \frac{1}{2}(E_{\max} - E_{\min})^2 \tag{4}$$

derived from (2). Without loss of generality we assume that

$$\sum_{i=1}^{q} E_i = 0, \quad E_1 \leq E_2 \leq \ldots \leq E_q, \quad E_{\min} = -E, \quad \text{and} \quad E_{\max} = E.$$

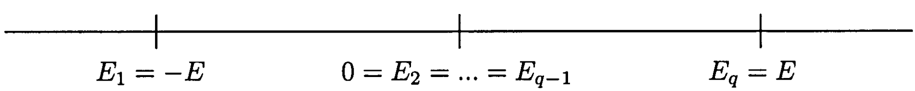
For the most favorable case (as shown in the following figure), we have

$$\bar{E} = 0, \quad \sum_{i=1}^{q} (E_i - \bar{E})^2 = qE^2.$$



For the least favorable case (as shown in the next figure), we have

$$\bar{E} = 0, \quad \sum_{i=1}^{q} (E_i - \bar{E})^2 = 2E^2.$$



In both cases $(E_{\max} - E_{\min})^2 = 4E^2$, and inequality (4) holds. Equality is true only in the least favorable case as expected.

We denote factors by A, B and C and their effects by α_i $(i = 1, 2, \ldots, a)$, β_j $(j = 1, 2, \ldots, b)$ and γ_k $(k = 1, \ldots, c)$, respectively (for the cross classification). We assume that at least one of the factors has fixed effects.

The following notation is used. If two factors A and B are cross classified, we write A × B. If B is nested in A we write A ≻ B. Random factors are underlined. For instance, ($\underline{A}$ ≻ $\underline{B}$) × C denotes a mixed classification for which B is nested within A and the whole combination of A and B is crossed with C. Here $\underline{A}$ and $\underline{B}$ are random and C is fixed.

Table 1 shows examples for some classifications and the value w determining the size of the experiment.

Table 1. Parameters w, f_1 and f_2 for some of the classifications in ANOVA models for testing H_0: $E_1 = E_2 = ... = E_q$; k denotes the block size in BIB and PBIB designs.

Classification	Effects		w	f_1	f_2
	Fixed	Random			
One way classification	A		n	$a-1$	$a(n-1)$
Two way cross A×B	A, B		n	$a-1$	$ab(n-1)$
BIB, PBIB (Two way cross)	A, B		b	$a-1$	$bk-a-b+1$
Two-way cross A×$\underline{B}$	A	B	b	$a-1$	$(a-1)(b-1)$
Two-way nested A≻B	A, B		n	$a-1$	$ab(n-1)$
Two-way nested A≻$\underline{B}$	A	B	b	$a-1$	$a(b-1)$
Two-way nested $\underline{A}$≻B	B	A	n	$a(b-1)$	$ab(n-1)$
Three-way cross A×B×C	A, B, C		n	$a-1$	$abc(n-1)$
Three-way cross A×B×$\underline{C}$	A, B	C	c	$a-1$	$(a-1)(c-1)$
Three-way nested A≻B≻C	A, B, C		n	$a-1$	$abc(n-1)$
Three-way nested A≻$\underline{B}$≻$\underline{C}$	A	B, C	b	$a-1$	$a(b-1)$
Three-way nested $\underline{A}$≻B≻$\underline{C}$	B	A, C	c	$a(b-1)$	$ab(c-1)$
Three-way nested $\underline{A}$≻$\underline{B}$≻C	C	A, B	n	$ab(c-1)$	$abc(n-1)$
Three-way nested A≻B≻$\underline{C}$	A, B	C	c	$a-1$	$ab(c-1)$
Three-way nested A≻$\underline{B}$≻C	A, C	B	b	$a-1$	$a(b-1)$
Three-way nested $\underline{A}$≻B≻C	B, C	A	n	$a(b-1)$	$abc(n-1)$
Three-way mixed (A×B)≻C	A, B, C		n	$a-1$	$abc(n-1)$
Three-way mixed (A×$\underline{B}$)≻C	A, C	B	b	$a-1$	$(a-1)(b-1)$
Three-way mixed (A×B)≻$\underline{C}$	A, B	C	c	$a-1$	$ab(c-1)$
Three-way mixed (A×$\underline{B}$)≻$\underline{C}$	A	B, C	b	$a-1$	$(a-1)(b-1)$
Three-way mixed (A≻B)×C	A, B, C		n	$a-1$	$abc(n-1)$
Three-way mixed ($\underline{A}$≻B)×$\underline{C}$	B	A, C	c	$a(b-1)$	$a(b-1)(c-1)$
Three-way mixed ($\underline{A}$≻$\underline{B}$)×C	C	A, B	a	$c-1$	$(a-1)(c-1)$
Three-way mixed (A≻$\underline{B}$)×C	A, C	B	b	$a-1$	$a(b-1)$
Three-way mixed (A≻B)×$\underline{C}$	A, B	C	c	$a-1$	$(a-1)(c-1)$

A description of the different classifications and the corresponding models can be found in Rasch (1995) and Rasch et. al. (1996a, Chapter 1/61).

3. The algorithm and the program

The solution λ of (3) is a known function $g(\alpha, \beta, f_1, f_2)$ of the risk of the first kind α, the risk of the second kind β, and the degrees of freedom f_1 and f_2 of the numerator and denominator, respectively. Because both sides of (3) depend on the size of the experiment, we have to find the minimal w that satisfies (3) for the subclass number n or for the number b of blocks if $n \equiv 1$ by iteration. We assume that α, β, f_1 and σ^2 (or a guess of σ^2) are given in advance, f_2 is a function of w. During the iteration the values λ_l are calculated from w_l by (3). The algorithm for this iteration is given by Rasch et al. (1996b).

We know that λ is monotonously decreasing with increasing f_2 (therefore, with the sample size, since f_2 is proportional to the sample size) and often the iteration converges.

In cases when the iteration does not converge, we use a systematic search. An example for the t-test (analogue to the F-test for the one-way ANOVA with $a = 2$), where the result can be tested by a pocket calculator, is given in Rasch et al. (1997) and Herrendörfer et al. (1997).

4. Examples

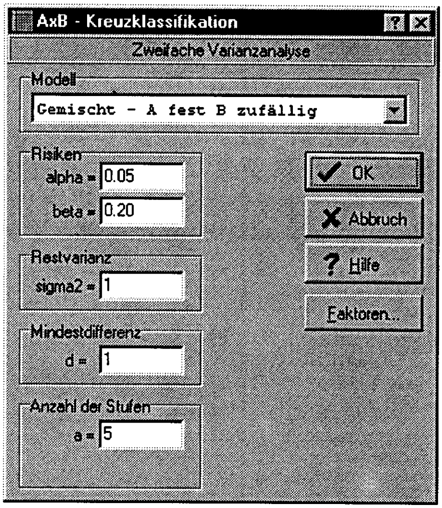
In this chapter we demonstrate how the program[1] works. We present examples of $\alpha = 0.05$, $\beta = 0.2$ and $d = \sigma$ for the cases shown in Table 2.

Table 2. Overview of the examples

Example	Number of factors	Classification and model	Specification
1	2	$A \times \underline{B}$	$a = 5$
2	2	$\underline{A} \succ B$	$b = 6$
3	2	$A \succ \underline{B}$	$a = 6$
4	3	$A \times B \times C$	$a = 4$, $b = 6$, $c = 3$
			$a = 6$, $b = 4$, $c = 3$
5	3	$(A \succ B) \times C$	$c = 2$

0 [1] Remark: The program is now commercially distributed as a part of the module ANOVA which runs under Windows 95 (distributor: BIOMATH GmbH Rostock. Tel. 0049-381-4059610, Fax 0049-381-4059200).

Example 1. Starting the program, we are prompted to input the following parameters



Clicking OK, we get the result:

```
Two-way Cross Classification AxB:
Factor A fixed, B random, testing equality of Factor A
Level of Factor A = 5
Alpha = 0.05, Beta = 0.2, Variance = 1, Precision = 1
Answer:
Levels of Factor B = 26, 12 for the least and most favorable cases respectively
```

Thus, we have $12 \leq b \leq 26$ and the experimenter may make a choice in this range. In the following examples we only give the output of the program and the range.

Example 2.

```
Two-way Nested Classification A>B:
Factor B fixed, A random, testing equality of Factor B
Level of Factor A = 8, Level of Factor B = 6,
Alpha = 0.05, Beta = 0.2, Variance = 1, Precision = 1
Answer:  Number of Replication = 56, 19 for the least and most favorable cases re-
spectively
```

Thus, $19 \leq n \leq 56$.

Example 3.

```
Two-way Nested Classification A>B:
Factor A fixed, B random, testing equality of Factor A
Level of Factor A = 6
Alpha = 0.05, Beta = 0.2, Variance = 1, Precision = 1
Answer:
Levels of Factor B = 27, 10 for the least and most favorable cases respectively
```

Thus, $10 \leq b \leq 27$.

Example 4.

```
Three-way cross classification (A×B×C):
Testing effects of Factor A
Level of Factor A = 4, Level of Factor B = 6, Level of Factor C = 3
Alpha = 0.05, Beta = 0.2, Variance = 1, Precision = 1
Answer:
Number of replication = 2, 1 for the least and most favorable cases respectively
```

Thus, $n = 1$ or 2.

If we like to test the effects of factor B in this example we have to rename the factors. The factor under test must always be factor A. Repeating the calculation with 6 levels of factor A, 4 levels of factor B and factor C unchanged gives the following output:

```
Three-way cross classification (A×B×C):
Testing effects of Factor A
Level of Factor A = 6, Level of Factor B = 4, Level of Factor C = 3
Alpha = 0.05, Beta = 0.2, Variance = 1, Precision = 1
Answer:
Number of replication = 3, 3 for the least and most favourable cases respectively
```

Thus $n = 3$.

Example 5.

```
Three-way mixed classification (A>B) × C:
C fixed, A, B random, testing effects of factor C
Level of Factor C = 2
Alpha = 0.05, Beta = 0.2, Variance = 1, Precision = 1
Answer:
Levels of Factor A = 17, 17 for the least and most favorable cases respectively
```

Thus, $\alpha = 17$. The unique value follows from the fact that when a factor has only 2 levels the least favorable and the most favorable cases are identical.

REFERENCES

Bowman K. O. (1972). Tables of the sample size requirement. *Biometrika* **59**, 234.

Bowman K.O., Kastenbaum M. A. (1975). Sample size requirement: Single and double classification experiments. *Selected Tables in Mathematical Statistics, Vol III, American Mathematical Society*, Providence, Rhode Island.

Bratcher T.L., Moran M.A., Zimmer W.J. (1970). Tables of sample size in the analysis of variance. *Journal of Quality Technology* **2**, 156-164.

Das Gupta P. (1968). Tables of the non-centrality parameter of F-test as a function of power. *The Indian Journal of Statistics Series B* **30**, 73-82.

David H.A., Lachenbruch P.A., Brandis H.P. (1972). The power function of range and studentized range tests in normal samples. *Biometrika* **59**, 161-168.

Fox M. (1956). Charts of the power of the F-test. *Annals of Mathematical Statistics* **27**, 484-497.

Herrendörfer G., Rasch D., Schmidt K., Wang M. (1997). Determination of the size of an experiment for the F-test in the Analysis of Variance – mixed models. In: Wegman E.J. and Azen, P.A. (Eds.), Computing Science and Statistics, **29**, 2, 547-550. Interface Found. of North America.

Kastenbaum M.A., Hoel D.G., Bowman K.O. (1970a). Sample size requirements for one-way analysis of variance. *Biometrika* **57**, 421-430.

Kastenbaum M. A., Hoel D. G., Bowman K. O. (1970b). Sample size requirements for randomized block designs. *Biometrika* **57**, 573-577.

Lehmer E. (1944). Inverse tables of probabilities of errors of the second kind. *Annals of Mathematical Statistics* **15**, 388-398.

Lenth R.V. (1986). Computing non-central Beta probabilities. *Appl. Statistics* **36**, 241-243.

Mace A.E. (1964). *Sample-size determination.* Reinhold Publishing Corporation, New York.

Norton V. (1983). A simple algorithm for computing the non-central F-distribution. *Appl. Statistics* **32**, 84-85.

Pearson E.S., Hartley H.O. (1951). Charts of the power function for analysis of variance tests, derived from the non-central F-distribution. Biometrika, 38, 112-30.

Pearson E.S., Hartley H.O. (1972). *Biometrika Tables for Statisticians.* Vol. 2, Cambridge University Press, London.

Rasch D. (1995). *Mathematische Statistik.* J. Ambrosius Barth, Heidelberg - Leipzig.

Rasch D., Herrendörfer G., Bock J., Victor N., Guiard V. (1996a) *Verfahrensbibliothek Versuchsplanung und -auswertung.* Band I, Oldenbourg Verlag, Müchen-Wien.

Rasch D., Wang M., Herrendörfer G. (1997). Determination of the size of an experiment for the F-test in the analysis of variance (Model I). *Advances in Statistical Software 6.* The 9th Conference on the Scientific Use of Statistical Software, Heidelberg.

Rasch D. (1998). Determination of the size of an experiment. In Atkinson A.C., Pronzato L., Wynn H.P. (eds.), *MODA 5 - Advances in Model-Oriented Data Analysis and Experimental Design, Physika Verlag*, 205-212.

Tang P.C. (1938). The power function of the analysis of variance tests with table and illustrations of their use. *Stat. Res. Mem.* **2**, 126-149.

Tiku M.L. (1967). Tables of the power of F-test. *Journal of the American Statistical Association* **62**, 525-539.

Tiku M.L. (1972). More tables of the power of the F-test. *Journal of the American Statistical Association* **67**, 709-710.

Received 10 September 1998

Wyznaczanie wielkości doświadczenia dla analizy wariancji w modelach stałych i mieszanych

STRESZCZENIE

Przedstawiono algorytm i program komputerowy służący do wyznaczania minimalnej wielkości doświadczenia dla porównania poziomów czynnika stałego za pomocą analizy wariancji. Minimalna wielkość jest wyznaczana na podstawie dolnego ograniczenia różnicy pomiędzy największym i najmniejszym z porównywanych efektów, ryzyka związanego z testem F i założonej, wspólnej wariancji resztowej. Rozpatruje się przypadek najbardziej sprzyjający (mini-min) oraz najbardziej niesprzyjający (maxi-min). Klasyfikacja oraz wybrany model wpływają tylko na liczbę stopni swobody dla testu F oraz na związek pomiędzy rozmiarem eksperymentu oraz parametrem niecentralności.

SŁOWA KLUCZOWE: analiza wariancji, rozmiar minimalny, parametr niecentralności, liczba podklas.