Preliminary Version

MODEL CONSIDERATIONS AND VARIANCE COMPONENT ESTIMATION IN AUGMENTED COMPLETELY RANDOMIZED AND RANDOMIZED COMPLETE BLOCKS DESIGNS

BU-592-M* W. T. Federer and S. R. Searle August 1976

Abstract

Augmented experiment designs include a set of standard treatments and a set of new treatments, usually with unequal replication on the two sets. The standard treatments are designed in any standard design and the blocks, rows or columns, etc. are enlarged to include subsets of the new treatments. A statistical model considered herein is for the case where the standard treatments are considered to be fixed effects and the new treatments as random effects. These, together with standard assumptions about the linear model, are used to form the model for this situation. The expectations of the sums of squares for augmented completely randomized and randomized complete block designs are presented for several cases. A variance component for new treatments is then obtainable.

* In the Mimeo Series of the Biometrics Unit, Cornell University, Ithaca, New York, 14853.

Preliminary Version

MODEL CONSIDERATIONS AND VARIANCE COMPONENT ESTIMATION IN AUGMENTED COMPLETELY RANDOMIZED AND RANDOMIZED COMPLETE BLOCKS DESIGNS

BU-592-M* W. T. Federer and S. R. Searle August 1976

1. INTRODUCTION

Augmented designs were introduced by Federer [1956, 1960, 1963, 1972] in a series of papers. Additional papers have also been prepared on this class of designs (see Searle [1965], Steel [1958], Federer and Raghavarao [1975], and Federer, Nair, and Raghavarao [1975]). The purpose of this paper is to present models for augmented completely randomized and randomized complete block designs, to discuss variance component and random effects estimation, and to present variances for some of the estimated quantities.

2. MODEL CONSIDERATIONS FOR THE AUGMENTED COMPLETELY RANDOMIZED DESIGN

In an augmented experiment design, the treatment design will often consist of v_s check varieties, standards, or controls and v_n new varieties (or treatments). The standards are selected for specific comparisons and hence should be considered as fixed effects. The new varieties must usually be considered to be a random sample of lines from some specified population of lines. One might wish to estimate a component of variance for the new varieties as well as to estimate the effects for the individual lines.

---

* In the Mimeo Series of the Biometrics Unit, Cornell University, Ithaca, New York, 14853.

Suppose that one draws a simple random sample of experimental units (e.u.) from a single population of experimental units. A response equation and model for this situation would usually be formulated as follows for the j^{th} observation:

$$(2.1) \qquad Y_j = \mu + \epsilon_j \ , \quad E(\epsilon_j) = 0 \ , \quad E(\epsilon_{ij}) = 0 \ , \quad j \neq j' \ , \text{ and } \ E(\epsilon_j^2) = \sigma_\epsilon^2 \ .$$

If Y_j is not of the above form, perhaps some function of the observation is. Assuming (2.1) is appropriate, we now wish to apply the standards and the new varieties to the simple random sample of e.u.'s obtained. If the variety effect is additive, then some such model as the following might be appropriate:

$$(2.2) \qquad Y_{ij} = \mu_i + \epsilon_{ij}$$

where $E\mu_{i.} = \mu_{i.}$ for $i=1,2,\cdots,v_s$; $E(\mu_{i.} = \mu_n+\tau_i) = E(\mu_n) = \mu_n$, $i=v_s+1,\cdots,v_s+v_n$; τ_i are $IID(0,\sigma_\tau^2)$, $i=v+1,\cdots,v_s+v_n$; and ϵ_{ij} are $IID(0,\sigma_\epsilon^2)$.

Another model that could be appropriate for this situation is that the variety effects enter in a multiplicative manner as follows:

$$(2.3) \qquad Y_j\tau_i = Y^*_{ij} = (\mu+\epsilon_j)\tau_i = \mu^*_{i.} + \epsilon^*_{ij} \ ,$$

where $E(Y^*_{ij}) = E(\mu^*_{i.}) = \tau_i\mu$ for $i=1,2,\cdots,v_s$, ϵ_j are $IID(0,\sigma_\epsilon^2)$; ϵ^*_{ij} are $ID(0,(\tau^*_i)^2\sigma_\epsilon^2)$; and $E(\log \mu\tau_i) = E(\log \mu) + E(\log \tau_i)$ where the $\log \tau_i$ are $IID(0,\sigma^2_{\log \tau})$. A test for this model is described in Federer [1955], page 50, for a specified check variety. The extension for v_s checks is straightforward.

Let us consider the model in (2.2) and the expectation of the following sums of squares:

(i) sum of squares among standard or check varieties;

(ii) sum of squares for standards versus new varieties;

(iii) sum of squares among new varieties; and

(iv) sum of squares for remainder.

Furthermore, let $Y_{i.} = \Sigma_{j=1}^{r} Y_{ij}$ for $i=1,2,\cdots,v_s$ and let Y_{i1} be the yield of the i^{th} new variety, $i=v_s+1,\cdots,v_s+v_n$; let $Y_{..} = \Sigma_{i=1}^{v_s}\Sigma_{j=1}^{r} Y_{ij}$ and $Y_{.1} = \Sigma_{i=v_s+1}^{v_s+v_n} Y_{i1}$. The total number of observations will be $v_n + rv_s$. The expectations of the various sums of squares are given in Table 2.1.

If there are r_i experimental units assigned to the i^{th} variety, then the expectation of the various sums of squares are given in Table 2.2. Let $C_s = \Sigma_{i=1}^{v_s}\Sigma_{j=1}^{r_i} Y_{ij}$ be the total of the standard variety yields, let $C_n = \Sigma_{i=v_s+1}^{v_s+v_n}\Sigma_{j=1}^{r_i} Y_{ij}$ be the total of the new variety yields, let $\Sigma_{i=1}^{v_s} r_i = r_s$ be the total number of observations on the standard varieties, let $\Sigma_{i=v_s+1}^{v_s+v_n} r_i = r_n$ be the total number of observations on the new varieties, and let $r_s + r_n = r_.$ be the total number of observations. Using this symbolism the various sums of squares are as given in Table 2.2. Note, however, that the residual mean square among standards may have a different expected value, $\sigma^2_{\epsilon s}$, from that for the remainder mean square for the new varieties, $\sigma^2_{\epsilon n}$. If $\sigma^2_{\epsilon s} = \sigma^2_{\epsilon n}$, then a pooled remainder mean square with $(r_s-v_s) + (r_n-v_n)$ degrees of freedom should be used. If the residual variation for new varieties is suspected of being higher (lower) than that for check varieties, one could perform an F-test to test this hypothesis.

Table 2.1. Expectation of sums of squares in an analysis of variance for an augmented completely randomized design for model (2.2). (Equal replication.)

Source of variation	d.f.	Sum of squares	Expected value of sum of squares
Total	rv_s+v_n	$\sum_{i=1}^{v_s}\sum_{j=1}^{r} Y_{ij}^2 + \sum_{i=v_s+1}^{v_s+v_n} Y_{i1}^2$	—
Correction for mean	1	$(Y_{..}+Y_{.1})^2/(rv_s+v_n)$	—
Among standard varieties	v_s-1	$\sum_{i=1}^{v_s} \frac{Y_{i.}^2}{r} - \frac{Y_{..}^2}{rv_s}$	$(v_s-1)\sigma_\epsilon^2 + r\sum_{i=1}^{v_s}\mu_{i.}^2 - \frac{r}{v_s}\left(\sum_{i=1}^{v_s}\mu_{i.}\right)^2$
Standards vs. new varieties	1	$\frac{Y_{..}^2}{rv_s} + \frac{Y_{.1}^2}{v_n} - \frac{(Y_{..}+Y_{.1})^2}{rv_s+v_n}$	$\sigma_\epsilon^2 + \frac{rv_s}{rv_s+v_n}\sigma_\tau^2 + \frac{r}{v_s}\left(\sum_{i=1}^{v_s}\mu_{i.}\right)^2 + v_n\mu_n^2$ $- r\sum_{i=1}^{v_s}\mu_{i.} + v_n\mu_n)^2/(rv_s+v_n)$
Among new varieties	v_n-1	$\sum_{i=v_s+1}^{v_s+v_n} Y_{i1}^2 - \left(\sum_{i=v_s+1}^{v_s+v_n} Y_{i1}\right)^2/v_n$	$(v_n-1)(\sigma_\epsilon^2+\sigma_\tau^2)$
Remainder	$v_s(r-1)$	$\sum_{i=1}^{v_s}\sum_{j=1}^{r} (Y_{ij}-\bar{y}_{i.})^2$	$v_s(r-1)\sigma_\epsilon^2$

Table 2.2. Expectation of sums of squares in an analysis of variance for an augmented completely randomized design for model (2.2). (Unequal replication.)

Source of variation	d.f.	Sum of squares	Expected value of sum of squares
Total	r_n+r_s	$\sum_{i=1}^{v_s+v_n}\sum_{j=1}^{r_i} Y_{ij}^2$	–
Correction for mean	1	$(C_s+C_n)^2/r_.$	–
Among standard varieties	v_s-1	$\sum_{i=1}^{v_s}\frac{Y_{i.}^2}{r_i}-\frac{C_s^2}{r_s}$	$(v-1)\sigma^2_{\epsilon s}+\sum_{i=1}^{v_s} r_i\mu_{i.}^2-\left(\sum_{i=1}^{v_s} r_i\mu_{i.}\right)^2/r_s$
Standard vs. new varieties	1	$\frac{C_s^2}{r_s}+\frac{C_n^2}{r_n}-\frac{(C_s+C_n)^2}{r_.}$	$\sigma^2_{\epsilon s}(1-r_s/r_.)+\sigma^2_{\epsilon n}(1-r_n/r_.)+\sigma^2_\tau\sum_{i=v_s+1}^{v_s+v_n} r_i^2\left(\frac{1}{r_n}-\frac{1}{r_.}\right)+\mu_n^2(r_n-r_n^2/r_.)$ $+\left(\sum_{i=1}^{v_s} r_i\mu_{i.}\right)^2\left(\frac{1}{r_s}-\frac{1}{r_.}\right)-2r_n\mu_n\sum_{i=1}^{v_s} r_i\mu_{i.}/r_.$
Among new varieties	v_n-1	$\sum_{i=v_s+1}^{v_s+v_n}\frac{Y_{i.}^2}{}-\frac{C_n^2}{r_n}$	$(v_n-1)\sigma^2_{\epsilon n}+\sigma^2_\tau\left(r_n-\sum_{i=v_s+1}^{v_s+v_n} r_i^2/r_n\right)$
Remainder for standards varieties	r_s-v_s	$\sum_{i=1}^{v_s}\sum_{j=1}^{r_i}(Y_{ij}-\bar{y}_{i.})^2$	$(r_s-v_s)\sigma^2_{\epsilon s}$
Remainder for new varieties	r_n-v_n	$\sum_{i=v_s+1}^{v_s+v_n}\sum_{j=1}^{r_i}(Y_{ij}-\bar{y})^2$	$(r_n-v_n)\sigma^2_{\epsilon n}$

3. MODEL CONSIDERATIONS FOR THE AUGMENTED RANDOMIZED COMPLETE BLOCK DESIGN

Let the treatment design for the augmented randomized complete block be the same as described in the previous section. A cluster - simple random sample will be used to select the experimental units to which the standard and new varieties will be applied. Consider the model for the e.u.'s in this case to be:

(3.1) $$Y_{hj} = \mu_{.j} + \epsilon_{hj} , \quad j=1,2,\cdots,b = \text{number of blocks} ,$$

where a simple random sample of size k_j is obtained from the j^{th} subpopulation (cluster) selected for experimentation, $E(Y_{hj}|j) = \mu_{.j} = \mu + \beta_j$, β_j are IID$(0,\sigma^2_\beta)$, and the ϵ_{hj} are IID$(0,\sigma^2_\epsilon)$.

If the variety effects, τ_i, have only an additive effect when applied to the experimental units, then a model similar to (2.2) may be formulated as:

(3.2) $$Y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij} = \mu_{i.} + \beta_j + \epsilon_{ij} , \quad i=1,2,\cdots,v_s,v_s+1,\cdots,v_s+v_n ,$$

where the above conditions for (3.1) and (3.2) hold.

If the variety effects τ_i have some multiplicative effect, then one of the following two models, (3.3) or (3.4), might be appropriate:

(3.3) $$\begin{aligned} Y_{hj}\tau_i &= (\mu+\beta_j+\epsilon_{hj})\tau_i = Y^*_{ij} \\ &= \mu^*_{i.} + \tau_i\beta_j + \tau_i\epsilon_{hj} \\ &= \mu^*_{i.} + \beta^*_j + \epsilon^*_{ij} \end{aligned}$$

where the ϵ^*_{ij} have zero mean and variance $\tau^2_i\sigma^2_\epsilon$ and the β^*_j have zero mean and a function of σ^2_β and the τ^2_i as their variance.

(3.4) $$Y_{ij} = \mu + \beta_j + \tau_i + \beta_j\tau_i/\mu + \epsilon'_{ij}$$

$$= \mu_{i.}\mu_{.j}/\mu + \epsilon'_{ij} ,$$

where the ϵ'_{ij} are IID$(0,\sigma^2_\epsilon)$ and the remaining quantities are defined as in (3.2). The latter model is what might be called the Tukey-type-of-nonadditivity model.

In the following we shall confine ourselves to model (3.2) as the theory is not available for model (3.3), for example. The expectations of the various sums of squares under model (3.2) are given in Table 3.1 for the case of v_s standard varieties arranged in $b = r_s$ blocks of a randomized complete block design. The blocks are enlarged and augmented with some of the new varieties in such a manner as to retain equality of error variances within blocks. That is, the block size, k_j, could vary as long as heterogeneity of variances is not encountered. Using an indicator variable, $n_{ij} = 0,1$, to indicate the presence or absence of a variety in a given block j under the assumption that no variety occurs more than once in a block, we could rewrite the response equation of (3.2) as

$$n_{ij}Y_{ij} = n_{ij}(\mu+\tau_i+\beta_j+\epsilon_{ij}) ,$$

which is the usual response equation for a binary incomplete block design for which $n_{ij} = 0,1$. The standard form of the analysis of variance holds. Let $Q_{i.} = Y_{i.} - \sum_{j=1}^{b} n_{ij}\bar{y}_{.j}$ where $Y_{i.}$ = total yield of i^{th} variety and $\bar{y}_{.j}$ be the mean of all k_j responses in the j^{th} block; let $\bar{y}_{.js}$ be the block mean for the v_s standard variety yields in block j; let the total of all $bv_s + v_n$ observations be $Y_{..}$, of the bv_s standard variety yields only, be $Y_{..s}$, and of the v_n new variety yields only, be $Y_{..n}$; let the corresponding means be $\bar{y}_{..s}$ for standard yields and

Table 3.1. Expectation of sums of squares in an analysis of variance for an augmented randomized complete block design under model (3.2) and with equal replication on standard varieties ($b=r_s$) and on new varieties ($r_i=1$).

Source of variation	d.f.	Sum of squares	Expected value of sum of squares
Total	bv_s+v_n	$\sum_{i=1}^{v_s+v_n}\sum_{j=1}^{r_i} Y_{ij}^2$	—
Correction for mean	1	$Y_{..}^2/(bv_s+v_n)$	—
Among blocks for standard yields only	b-1	$v\sum_{j=1}^{b}(\bar{y}_{.js}-\bar{y}_{..s})^2$	$(b-1)(\sigma_\epsilon^2+v_s\sigma_\beta^2)$
Among standard varieties	v_s-1	$r\sum_{i=1}^{v_s}(\bar{y}_{i.}-\bar{y}_{..s})^2$	$(v_s-1)\sigma_\epsilon^2 + r\sum_{i=1}^{v_s}\mu_{i.}^2 - \frac{r}{v_s}\left(\sum_{i=1}^{v_s}\mu_{i.}\right)^2$
Standards vs. new varieties (eliminate blocks)	1	—	—
Among new varieties (eliminate blocks)	v_n-1	$\sum_{i=v_s+1}^{v_s+v_n}\hat{\tau}_i^2 - \left(\sum_{i=v_s+1}^{v_s+v_n}\hat{\tau}_i\right)^2/v_n$	$(v_n-1)(\sigma_\epsilon^2+\sigma_\tau^2)$
Standard variance X block = remainder	$(v_s-1)(b-1)$	$\sum_{i=1}^{v_s}\sum_{j=1}^{b}(Y_{ij}-\bar{y}_{i\cdot s}-\bar{y}_{.js}+\bar{y}_{..s})^2$	$\sigma_\epsilon^2(v_s-1)(b-1)$

$\bar{y}_{\cdot\cdot n}$ for new varieties. Furthermore, let the $\hat{\tau}_i = Y_{i\cdot n} - \bar{y}_{\cdot js}n_{ij}$ = new variety total minus the estimated block mean (from standard yields only) for the block containing the i^{th} new variety. Then, the sum of squares $\sum_{i=v_s+1}^{v_s+v_n} \hat{\tau}_i^2$ has the expection, given that $\mu = \sum_{i=1}^{v_s} \mu_{i\cdot}/v_s$, indicated in Table 3.1.

4. AN EXAMPLE OF AN AUGMENTED RANDOMIZED COMPLETE BLOCK DESIGN

A hypothetical numerical example of an augmented randomized complete block with $v_s = 4$, $b = 3$, and $v_n = 8$ is given in Federer [1956]. The analysis of variance table is reproduced in Table 4.1, putting it in the form of Table 3.1.

One could also utilize a sum of squares among new varieties within blocks and obtain the resulting expectation. The expected value of the sum of squares in the j^{th} block would be $(k_j-v_s-1)(\sigma_\epsilon^2+\sigma_\tau^2)$. The expected value of this sum of squares over all blocks would be $(\sigma_\epsilon^2+\sigma_\tau^2) \sum_{j=1}^{b} (k_j-v_s-1)$, with $\sum_{j=1}^{b} (k_j-v_s-1)$ degrees of freedom.

Table 4.1. Analysis of variance table for an augmented randomized complete block design with $b=3$, $v_s=4$, and $v_n=8$.

Source of variation	d.f.	Sum of squares	Mean square	Expected value of mean square
Total	20	133,652		
Correction for mean	1	132,845		
Among blocks for standard yields only	2	69.5000	34.75	$\sigma_\epsilon^2 + 4\sigma_\beta^2$
Among standard varieties	3	52.9167	17.64	$\sigma_\epsilon^2 + \sum_{i=1}^{4} \mu_{i.}^2 - \frac{1}{4}\left(\sum_{i=1}^{4} \mu_{i.}\right)^2$
Standards vs. new varieties (eliminating block effects)	1	-	-	-
Among new varieties (eliminating blocks)	7	289.8047	41.40	$\sigma_\epsilon^2 + \sigma_\tau^2$
Remainder	6	161.8333	26.97	σ_ϵ^2

REFERENCES

Federer, W. T. [1956]. Augmented (or hoonuiaku) designs. Hawaiian Planters' Record, 55:191-208.

Federer, W. T. [1960]. Augmented designs with two-, three-, and higher-way elimination of heterogeneity. BU-329-M in the Biometrics Unit Mimeo Series, Cornell University. (Also see Abstract in Biometrics 17:166.)

Federer, W. T. [1961]. Augmented designs with one-way elimination of heterogeneity. Biometrics 17:447-473.

Federer, W. T. [1963]. Procedures and designs useful for screening material in selection and allocation, with a bibliography. Biometrics 19:553-587.

Federer, W. T. [1967]. Experimental Design — Theory and Application (original edition in 1955 by The Macmillan Company, New York), Oxford and I.B.H. Publishing Company, Calcutta, Bombay, and New Delhi.

Federer, W. T. [1972]. Construction of classes of experimental designs using transversals in Latin squares and Hedayat's sum composition method. Chapter 7 in Statistical Papers in Honor of George W. Snedecor (edited by T. A. Bancroft), Iowa State University Press, Ames, Iowa.

Federer, W. T., R. C. Nair and D. Raghavarao. [1975]. Some augmented row-column designs. Biometrics 31:361-373.

Federer, W. T. and D. Raghavarao. [1975]. On augmented designs. Biometrics 31:29-35.

Searle, S. R. [1965]. Computing formulae for analyzing augmented random complete block designs. BU-207-M in the Biometrics Unit Mimeo Series, Cornell University.

Steel, R. G. D. [1958]. A class of augmented designs. Math. Res. Center Tech. Summary Report No. 56, University of Wisconsin.