MEANS AND VARIANCES OF TREATMENTS FROM A RANDOMIZED COMPLETE BLOCK DESIGN WITH MISSING PLOTS

BU-99-M W. T. Federer and O. Nissen December, 1958

ABSTRACT

The estimated treatment effects and variances have been developed for a number of cases in which plots are missing in the randomized complete block design. In particular, treatment effect and variance formulas are given for (i) one missing plot, (ii) two missing plots for one treatment, (iii) k missing plots for one treatment, (iv) two missing plots for two treatments in one block, (v) k missing plots for each of k treatments in one block, (vi) two treatments each with one missing plot in different replicates, and (vii) k treatments each with a missing plot in different replicates. Numerical examples are presented for cases (i) and (ii).

MEANS AND VARIANCES OF TREATMENTS FROM A
RANDOMIZED COMPLETE BLOCK DESIGN WITH MISSING PLOTS

BU-99-M

W. T. Federer and O. Nissen

December, 1958

Although the presence of a formula for estimating the yield of a missing observation is of frequent occurrence in statistical literature, the procedure for obtaining the estimated treatment effect or mean is rather infrequent. Therefore, in order to emphasize the estimation of treatment effects rather than the computational procedure for the analysis of variance and presumably the resulting tests of significance, the algebraic derivation of treatment effects is presented below for some situations where observations are missing.

One Missing Value

For the first situation consider a randomized complete block design with one missing observation, say for treatment 1 in replicate 1. The normal equations for this design are given on page 128 of Federer's "Experimental Design." These equations reduce to the following when the yield X_{11} is deleted from the array of yields:

$$(n-1)(\hat{\mu}+t_1)-r_1 = X_{1.}$$

$$n(\hat{\mu}+t_i) = X_{i.} \quad \text{for } i=2,3,\cdots,v$$

$$(v-1)(\hat{\mu}+r_1)-t_1 = X_{.1}$$

$$v(\hat{\mu}+r_j) = X_{.j} \quad \text{for } j=2,3,\cdots,n$$

$$(vn-1)\hat{\mu}-t_1-r_1 = X_{..}$$

From the above equation, the following solutions are obtained:

$$\hat{\mu}+t_i=\bar{x}_{i.} \quad \text{for } i=2,3,\cdots,v ;$$

$$\hat{\mu}+r_j=\bar{x}_{.j} \quad \text{for } j=2,3,\cdots,n \quad .$$

The solutions for $\hat{\mu}$, t_1, and r_1 are obtained from the following three equations:

$$(n-1)(\hat{\mu}+t_1)-r_1 = X_{1.}$$

$$(v-1)(\hat{\mu}+r_1)-t_1 = X_{.1}$$

$$(nv-1)\hat{\mu}-t_1-r_1 = X_{..}$$

Solving these equations we find that

$$\hat{\mu} = \frac{X_{..}(nv-v-n)+vX_{1.}+nX_{.1}}{nv(n-1)(v-1)} \quad ,$$

$$t_1 = \frac{1}{v(n-1)}\left\{ vX_{1.}+X_{.1}-X_{..} \right\} \quad ,$$

and

$$r_1 = \frac{1}{n(v-1)}\left\{ X_{1.}+nX_{.1}-X_{..} \right\} \quad .$$

Therefore, the mean for treatment one is

$$\hat{\mu}+t_1 = \frac{1}{v(n-1)}\left\{ vX_{1.}+X_{.1}-X_{..} \right\} + \frac{X_{..}(nv-v-n)+vX_{1.}+nX_{.1}}{nv(n-1)(v-1)}$$

$$= \frac{1}{n}\left\{ X_{1.}+\frac{vX_{1.}+nX_{.1}-X_{..}}{(n-1)(v-1)} \right\}$$

$$= \frac{1}{n}\left\{ X_{1.}+\hat{X}_{11} \right\} \quad ,$$

where $\hat{X}_{11}$ is the formula for a missing plot for treatment one in replicate one. This means that the estimated treatment mean for treatment one is obtained by summing the observations for treatment one and then adding the computed value for the missing plot to this total.

Also, it is stated that the treatment and block sums of squares are over-estimated if the missing plot yield is inserted in the yields and the analysis completed as for equal numbers. At least three methods for correcting these sums of squares are available (i.e., the method of fitting constants, covariance, and using a missing plot analysis together with an analysis where the squares are weighted by the number of observations). Since we now have all the t_i and can compute the $Q_{i.}=X_{i.}-$ means of blocks in which i^{th} treatment occurs, it is possible to obtain the adjusted sum of squares directly as $\sum_{i=1}^{v} t_i Q_i =$ treatment (eliminating blocks) effects. The analysis of variance table then becomes:

Source of variation	df	Sum of squares
Blocks (ignoring treatment)	n-1	$\Sigma \frac{X^2_{.j}}{v_j} - X^2_{..}(\Sigma v_j = vn-1)$
Treatments (eliminating blocks)	v-1	$\Sigma t_i Q_{i.}$
Residual	(n-1)(v-1)-1	by subtraction
Correction for mean	1	$X^2_{..}/(nv-1)$
Total (uncorrected)	nv-1	$\sum_{i=1}^{v}\sum_{j=1}^{r_i} X^2_{ij}$

The Residual sum of squares obtained by subtraction above is identical to that obtained from the missing plot analysis or a covariance analysis.

The variance of a difference between two treatment means for treatments 2 to v is $2\sigma^2_\epsilon/n$. The variance of a difference between treatment 1 and any other treatment, say 2, is obtained as:

$$E[t_1-t_2-E(t_1-t_2)]^2$$
$$=E\left\{\frac{nv-n+1}{n(v-1)(n-1)} \sum_{j=2}^{n}\epsilon_{ij} + \frac{1}{(n-1)(v-1)} \sum_{i=2}^{v}\epsilon_{i1} - \frac{1}{n(n-1)(v-1)}\left(\sum_{i=1}^{v}\sum_{j=1}^{n}\epsilon_{ij}-\epsilon_{11}\right) - \frac{1}{n}\sum_{j=1}^{n}\epsilon_{2j}\right\}^2$$
$$=\sigma^2_\epsilon\left\{\frac{2}{n} + \frac{v}{n(n-1)(v-1)}\right\}.$$

as given by Yates [1933].

We may obtain this variance in another manner. The solutions for the t_1 in matrix form are:

$$
\begin{pmatrix} t_1 \\ t_2 \\ t_3 \\ \vdots \\ t_v \end{pmatrix}
=
\begin{pmatrix}
\frac{v-1}{v(n-1)} & \frac{-1}{v(n-1)} & \frac{-1}{v(n-1)} & \cdots & \frac{-1}{v(n-1)} \\
\frac{-1}{v(n-1)} & \left(\frac{1}{n} - \frac{nv-n-v}{nv(n-1)(v-1)}\right) & \frac{-(nv-n-v)}{nv(n-1)(v-1)} & \cdots & \frac{-(nv-n-v)}{nv(n-1)(v-1)} \\
\frac{-1}{v(n-1)} & \frac{-(nv-n-v)}{nv(n-1)(v-1)} & \left(\frac{1}{n} - \frac{nv-n-v}{nv(n-1)(v-1)}\right) & \cdots & \frac{-(nv-n-v)}{nv(n-1)(v-1)} \\
 & & & & \\
\frac{-1}{v(n-1)} & \frac{-(nv-n-v)}{nv(n-1)(v-1)} & \frac{-(nv-n-v)}{nv(n-1)(v-1)} & \cdots & \left(\frac{1}{n} - \frac{nv-n-v}{nv(n-1)(v-1)}\right)
\end{pmatrix}
\cdot
\begin{pmatrix}
X_{1\cdot} - \sum_{j=2}^{n} \bar{x}_{\cdot j} = Q_{1\cdot} \\
X_{2\cdot} - \sum_{j=1}^{n} \bar{x}_{\cdot j} = Q_{2\cdot} \\
X_{3\cdot} - \sum_{j=1}^{n} \bar{x}_{\cdot j} = Q_{3\cdot} \\
\vdots \\
X_{v\cdot} - \sum_{j=1}^{n} \bar{x}_{\cdot j} = Q_{v\cdot}
\end{pmatrix}
$$

The variance of the difference t_1-t_2 is

$$V(t_1-t_2) = \sigma_\epsilon^2 \left\{\frac{v-1}{v(n-1)} + \frac{1}{n} - \frac{(nv-n-v)}{nv(n-1)(v-1)} - 2\left(\frac{-1}{v(n-1)}\right)\right\}$$

$$= \sigma_\epsilon^2 \left\{\frac{2}{n} + \frac{v}{n(n-1)(v-1)}\right\}$$

Likewise,

$$V(t_2-t_3) = \sigma_\epsilon^2 \left\{2\left(\frac{1}{n} - \frac{(nv-n-v)}{nv(n-1)(v-1)}\right) - 2\left(\frac{-nv-n-v}{nv(n-1)(v-1)}\right)\right\}$$

$$= 2\sigma_\epsilon^2/n \quad .$$

We may contrast the above matrix with the matrix obtained when all yields are present, i.e.,

$$\begin{pmatrix} t_1 \\ t_2 \\ t_3 \\ t_4 \\ \vdots \\ t_v \end{pmatrix} = \frac{1}{nv} \begin{pmatrix} v-1 & -1 & -1 & -1 & \cdots & -1 \\ -1 & v-1 & -1 & -1 & \cdots & -1 \\ -1 & -1 & v-1 & -1 & \cdots & -1 \\ -1 & -1 & -1 & v-1 & \cdots & -1 \\ \vdots & & & & & \\ -1 & -1 & -1 & -1 & \cdots & v-1 \end{pmatrix} \cdot \begin{pmatrix} X_{1.}-n\bar{x} \\ X_{2.}-n\bar{x} \\ X_{3.}-n\bar{x} \\ X_{4.}-n\bar{x} \\ \vdots \\ X_{v.}-n\bar{x} \end{pmatrix}$$

The variance of a difference between any two t_i is:

$$V(t_i-t_{i'},\ i\neq i') = \sigma_\epsilon^2\left\{\frac{v-1}{nv} + \frac{v-1}{nv} - 2\left(\frac{-1}{nv}\right)\right\} = 2\sigma_\epsilon^2/n.$$

The variances and covariances of effects are given on page 129 (loc. cited).

<u>Numerical example</u> -- Consider the following example taken from Federer's "Experimental Design," p. 510:

Treatment	Replicate I	II	III	Total
A	6	5	4	15
B	15	10	8	33
C	15	15	$\hat{X}_{33}$	$30+\hat{X}_{33}$
	36	30	$12+\hat{X}_{33}$	$78+\hat{X}_{33}$

From the ordinary missing plot formula, we find

$$\hat{X}_{33} = \frac{3(30)+3(12)-78}{(3-1)(3-1)} = 12.$$

Substituting this value in the table of yields and completing the analysis of variance by the method of missing plots results in the following:

Source of variation	df	Sum of squares
Replicate	2	24
Treatment	2	126
Error or residual	3	10
Total (corrected for mean)	7	160

Now, from the above formulae we find,

$$\hat{\mu} = \frac{78(9-3-3)+3(30)+3(12)}{3(3)(3-1)(3-1)} = 10 \quad ;$$

$$t_3 = \frac{3(30)+12-78}{3(3-1)} = 4 \quad ;$$

$$r_3 = \frac{30+3(12)-78}{3(3-1)} = -2 \quad ;$$

$$\hat{\mu}+t_3 = 10 + 4$$
$$= \frac{1}{3}(30+12) = 14.$$

The treatment (eliminating block) sum of squares is computed as:

$$\sum_{i=1}^{3} t_i Q_i = \frac{1}{3}(15-30)(15-(12+10-6))$$

$$+ \frac{1}{3}(33-3(10))(33-(12+10+6))+4(30-12-10)$$

$$= 65+5+32=102,$$

which agrees with the adjusted sum of squares for treatments in Tables XVI-18 and XVI-21 in "Experimental Design."

The variance of a difference for various pairs of means is:

Actual variance	Estimated variance
$V(\bar{x}_A-\bar{x}_B)=\frac{2}{3}\sigma_\epsilon^2$	$s^2_{\bar{x}_A-\bar{x}_B} = \frac{2}{3}(\frac{10}{3}) = \frac{20}{9}$
$V(\bar{x}_A-\bar{x}_C)=\sigma_\epsilon^2(\frac{2}{3} + \frac{3}{3(4)})$	$s^2_{\bar{x}_A-\bar{x}_C} = \frac{10}{3}(\frac{2}{3} + \frac{3}{12}) = \frac{110}{36} = \frac{55}{18}$
$V(\bar{x}_B-\bar{x}_C)=\sigma_\epsilon^2(\frac{2}{3} + \frac{3}{12})$	$s^2_{\bar{x}_B-\bar{x}_C} = \frac{10}{3}(\frac{2}{3} + \frac{3}{12}) = \frac{55}{18}$.

Two Missing Values for One Treatment

As a second illustration, consider the case where treatment one is missing in replicates one and two. The following normal equations for the effects (page 128, loc. cit.), after applying the equations $\Sigma r_i=\Sigma t_i=0$, are obtained:

$$(nv-2)\hat{\mu}-2t_1-r_1-r_2 = X_{..}$$

$$(n-2)(\hat{\mu}+t_1)-r_1-r_2 = X_{1.}$$

$$n(\hat{\mu}+t_i) = X_{i.} \quad i=2,3,\cdots, v$$

$$(v-1)(\hat{\mu}+r_1)-t_1 = X_{.1}$$

$$(v-1)(\hat{\mu}+r_1)-t_1 = X_{.2}$$

$$v(\hat{\mu}+r_j) = X_{.j} \quad i=3,4,\cdots,n$$

Solving for r_1 and r_2 in terms of t_1, $\hat{\mu}$, and the observations and then substituting the solutions for r_1 and r_2 in the equation for t_1 results in the following:

$$t_1\left(\frac{nv-n-2v}{v-1}\right) = X_{1.}+\bar{x}_{.1}+\bar{x}_{.2}-n\hat{\mu}$$

$$nt_i \qquad = X_{i.}-n\hat{\mu}\ , \quad i=2,3,\cdots,v \quad .$$

Now, divide through by the coefficient of t_i and the sum. Therefore, $\sum_{i=1}^{v} t_i=0$

$$= \frac{(v-1)(X_{1.}+\bar{x}_{.1}+\bar{x}_{.2}-n\hat{\mu})}{nv-n-2v} + \frac{1}{n}\sum_{i=2}^{v} X_{i.} - (v-1)\hat{\mu} \ .$$

Solving,

$$\hat{\mu} = \frac{X_{..}(vn-n-2v)+2vX_{1.}+n(X_{.1}+X_{.2})}{nv(v-1)(n-2)}$$

and we now have the solution for the individual t_i. Also, the equation in t_1 now may be written as:

$$t_1 = \frac{vX_{1.}+X_{.1}+X_{.2}-X_{..}}{v(n-2)}$$

and

$$\hat{\mu}+t_1 = \frac{X_{..}(nv-n-2v)+2vX_{1.}+n(X_{.1}+X_{.2})}{nv(v-1)(n-2)}$$

$$+ \frac{vX_{1.}+X_{.1}+X_{.2}-X_{..}}{v(n-2)} = \frac{1}{n}\left\{X_{1.}+\hat{X}_{11}+\hat{X}_{12}\right\} \quad ,$$

where $\hat{X}_{11}$ and $\hat{X}_{12}$ are obtained from the formulae on page 134 (loc. cit.).

The solutions for the t_i in matrix form are:

$$\begin{bmatrix} t_1 \\ t_2 \\ t_3 \\ \vdots \\ t_v \end{bmatrix} = \begin{bmatrix} \frac{v-1}{v(n-2)} & -\frac{1}{v(n-2)} & -\frac{1}{v(n-2)} & \cdots & -\frac{1}{v(n-2)} \\ -\frac{1}{v(n-2)} & \frac{1}{n}-\frac{nv-n-2v}{nv(n-2)(v-1)} & -\frac{(nv-n-2v)}{nv(n-2)(v-1)} & \cdots & -\frac{(nv-n-2v)}{nv(n-2)(v-1)} \\ -\frac{1}{v(n-2)} & -\frac{(nv-n-2v)}{nv(n-2)(v-1)} & \frac{1}{n}-\frac{nv-n-2v}{nv(n-2)(v-1)} & \cdots & -\frac{(nv-n-2v)}{nv(n-2)(v-1)} \\ & & & \vdots & \\ -\frac{1}{v(n-2)} & -\frac{(nv-n-2v)}{nv(n-2)(v-1)} & -\frac{(nv-n-2v)}{nv(n-2)(v-1)} & \cdots & \frac{1}{n}-\frac{nv-n-2v}{nv(n-2)(v-1)} \end{bmatrix} \cdot \begin{bmatrix} Q_{1\cdot} \\ Q_{2\cdot} \\ Q_{3\cdot} \\ \vdots \\ Q_{v\cdot} \end{bmatrix}$$

where $Q_{1\cdot} = X_{1\cdot} - \sum_{j=3}^{n} \bar{x}_{\cdot j}$ and $Q_{i\cdot} = X_{i\cdot} - \sum_{j=1}^{n} \bar{x}_{\cdot j}$ for $i=2,3,\cdots,v$.

The treatment (eliminating block effects) sum of squares is $\sum_{i=1}^{v} t_i Q_{i.}$; the various variances are:

$$V(t_1-t_i,\ i=2,3,\cdots,v) = \sigma_\epsilon^2\left\{ \frac{v-1}{v(n-2)} + \frac{1}{n} - \frac{nv-n-2v}{nv(n-2)(v-1)} \right.$$

$$\left. -2\left(-\frac{1}{v(n-2)}\right)\right\} = \sigma_\epsilon^2\left\{ \frac{2}{n} + \frac{2v}{n(n-2)(v-1)} \right\} \ ;$$

$$V(t_2-t_3) = \sigma_\epsilon^2\left\{ \frac{2}{n} - \frac{2(nv-n-2v)}{nv(n-2)(v-1)} \right.$$

$$\left. -2\left(-\frac{(nv-n-2v)}{nv(n-2)(v-1)}\right)\right\} = 2\ \sigma_\epsilon^2/n \ .$$

<u>Numerical</u> <u>example</u> -- Consider the previous example with treatment C missing in replicate I. The following yields remain:

Treatment	Replicate I	II	III	Total
A	6	5	4	$X_{1.}=15$
B	15	10	8	$X_{2.}=33$
C	$\hat{X}_{31}$	15	$\hat{X}_{33}$	$X_{3.}+\hat{X}_{31}+\hat{X}_{33}=15+\hat{X}_{31}+\hat{X}_{33}$
	$21+\hat{X}_{31}$	30	$12+\hat{X}_{33}$	63

From missing plot formulae given on page 134 of Federer [1955] we find

$$\hat{X}_{31} = \frac{vX_{3.}+(n-1)X_{.1}+X_{.3}-X_{..}}{(n-2)(v-1)}$$

$$= \frac{3(15)+(3-1)(21)+12-63}{(3-2)(3-1)=2} = 18$$

and

$$\hat{X}_{33} = \frac{vX_{3.}+X_{.1}+(n-1)X_{.3}-X_{..}}{(n-2)(v-1)}$$

$$= \frac{3(15)+21+(3-1)(12)-63}{(3-2)(3-1)=2} = 13.5$$

Substituting the missing plot values in the table of yields and completing the analysis of variance results in:

Source of variation	df	Sum of squares
Blocks	2	31.50
Treatments	2	166.50
Error	4-2=2	7.00
Correction for mean	1	992.25
Total uncorrected	7	1197.25

The "anova" using the "method of fitting constants" is:

Source of variation	df	Sum of squares
Blocks (ign. treat.)	2	25.50 = $\Sigma X^2_{.j}/n_{.j} - X^2_{..}/n_{..}$
Treatments (elim. blocks)	2	91.50 = $\Sigma t_i Q_{i.}$
Error	2	7.00 by subtraction
Correction for mean	1	567
Total uncorrected	7	691

where

$$\hat{\mu} = \frac{0+6(15)+3(21+12)}{3(3)(3-1)(3-2)=18} = 10.5$$

$$t_1 = \frac{15-(10.5)(3)}{3} = -5.5$$

$$t_2 = 11-10.5=0.5$$

$$t_3 = \frac{3(15)+21+12-63}{3(3-2)} = 5.0$$

$$Q_{1.} = 15-10.5-10-6=-11.5$$

$$Q_{2.} = 33-10.5-10-6=6.5$$

$$Q_{3.} = 15-10 \qquad =5.0$$

$$\Sigma t_i Q_{i.} = (-5.5)(-11.5)+(.5)(6.5)+(5)(5.0)=91.50 \quad .$$

Variances

$$s^2_{A-B} = \frac{7}{2}\left\{\frac{1}{3}+\frac{1}{3}\right\} = \frac{7}{3}$$

$$s^2_{A-C} = \frac{7}{2}\left\{\frac{2}{3}+1\right\} = \frac{35}{6}$$

$$s^2_{B-C} = \frac{7}{2}\left\{\frac{2}{3}+1\right\} = \frac{35}{6}$$

k Missing Values for One of the Treatments

Allan and Wishart [1930] and Yates [1933] were the first to present discussions for one and several missing values, respectively; Yates [1933] presents an iterative procedure for two or more missing values. Federer [1951] gave missing plot formulae for k different varieties missing in k different replicates and for k varieties missing in one replicate (or alternatively, k replicate yields missing for one variety). Some time later Thompson [1956] presented the same formulae with a numerical application.

For the case of $k(k<n)$ values missing for one of the treatments, say one, solutions for the t_i are given by:

$$\begin{pmatrix} t_1 \\ t_2 \\ t_3 \\ \vdots \\ t_v \end{pmatrix} = \begin{pmatrix} \frac{v-1}{v(n-k)} & -\frac{1}{v(n-k)} & -\frac{1}{v(n-k)} & \cdots & -\frac{1}{v(n-k)} \\ -\frac{1}{v(n-k)} & \frac{1}{n}-\frac{nv-n-kv}{nv(n-k)(v-1)} & -\frac{nv-n-kv}{nv(n-k)(v-1} & \cdots & -\frac{nv-n-kv}{nv(n-k)(v-1)} \\ -\frac{1}{v(n-k)} & -\frac{nv-n-kv}{nv(n-k)(v-1)} & \frac{1}{n}-\frac{nv-n-kv}{nv(n-k)(v-1)} & \cdots & -\frac{nv-n-kv}{nv(n-k)(v-1)} \\ \vdots & & & \vdots & \\ -\frac{1}{v(n-k)} & -\frac{nv-n-kv}{nv(n-k)(v-1)} & -\frac{nv-n-kv}{nv(n-k)(v-1)} & \cdots & \frac{1}{n}-\frac{nv-n-kv}{nv(n-k)(v-1)} \end{pmatrix} \cdot \begin{pmatrix} Q_{1\cdot} \\ Q_{2\cdot} \\ Q_{3\cdot} \\ \vdots \\ Q_{v\cdot} \end{pmatrix}$$

where $Q_{1\cdot}=X_{1\cdot}-\sum_{j=k+1}^{n}\bar{x}_{\cdot j}$, $Q_{i\cdot}=X_{i\cdot}-\sum_{j=1}^{n}\bar{x}_{\cdot j}$ for $i=2,3,\cdots,v$. The variances become:

$$V(t_1-t_i,\ i=2,\cdots,v)=\sigma_\epsilon^2\left\{\frac{2}{n}+\frac{kv}{n(n-k)(v-1)}\right\}$$

and

$$V(t_i-t_{i'},\ \text{for } i\neq i' \text{ both}>1)=2\sigma_\epsilon^2/n \ .$$

Two Missing Values for Two Treatments in One Block

The normal equations for this situation are, say for treatments 1 and 2 missing in block 1:

$$(nv-2)\hat{\mu}-t_1-t_2-2r_1 = X_{..}$$

$$(n-1)(\hat{\mu}+t_1)-r_1 = X_{1.}$$

$$(n-1)(\hat{\mu}+t_2)-r_1 = X_{2.}$$

$$n(\hat{\mu}+t_i) = X_{i.} \quad \text{for } i=3,\cdots,v$$

$$(v-2)(\hat{\mu}+r_1)-t_1-t_2 = X_{.1}$$

$$v(\hat{\mu}+r_j) = X_{.j} \quad \text{for } j=2,\cdots,r$$

From the above

$$\hat{\mu}+r_1=(X_{.1}+t_1+t_2)/(v-2)=\bar{x}_{.1}+\frac{t_1+t_2}{v-2}$$

Therefore, substituting the value for r_1 in the t_1 and t_2 equations we obtain:

$$(nv-v-2n+1)t_1-t_2=(v-2)X_{1.}+X_{.1}-n(v-2)\hat{\mu}$$

$$-t_1+(nv-v-2n+1)t_2=(v-2)X_{2.}+X_{.1}-n(v-2)\hat{\mu}$$

Solving

$$t_1 = \frac{nv-v-2n+1}{(nv-v-2n+1)^2-1}\left\{X_{1.}(v-2)+\frac{X_{2.}(v-2)}{nv-v-2n+1}\right.$$

$$\left.+\frac{(n-1)(v-2)}{(nv-v-2n+1)}(X_{.1}-n(v-2)\hat{\mu})\right\}$$

$$t_2 = \frac{nv-v-2n+1}{(nv-v-2n+1)^2-1}\left\{X_{2.}(v-2)+\frac{X_{1.}(v-2)}{nv-v-2n+1}\right.$$

$$\left.+\frac{(n-1)(v-2)}{(nv-v-2n+1)}(X_{.1}-n(v-2)\hat{\mu})\right\}$$

Also,

$$t_i=\bar{x}_{i.}-\hat{\mu} \quad \text{for } i=3,\cdots,v$$

Now $\Sigma t_i=0$

$$=\frac{X_{..}}{n}+\frac{v(X_{1.}+X_{2.})}{n(nv-v-2n)}+\frac{2X_{.1}}{nv-v-2n}-\frac{\hat{\mu}v(v-2)(n-1)}{nv-v-2n}$$

Solving,

$$\hat{\mu}=\frac{X_{..}(nv-v-2n)+v(X_{1.}+X_{2.})+2nX_{.1}}{nv(v-2)(n-1)=nv(A+1)}$$

Therefore, for $A=nv-v-2n+1$ and $A+1=(n-1)(v-2)$

$$t_1 = \left\{vX_{1.}+X_{.1}-X_{..}\right\}/v(n-1)$$

$$t_2 = \left\{vX_{2.}+X_{.1}-X_{..}\right\}/v(n-1)$$

$$\hat{\mu}+t_1 = X_{1.}\left\{\frac{v-2}{A+1}+\frac{v}{nv(A+1)}\right\}+X_{2.}\frac{v}{nv(A+1)}$$

$$+X_{.1}\left\{\frac{v-2}{v(A+1)}+\frac{2n}{nv(A+1)}\right\}+X_{..}\left(\frac{-v+2}{v(A+1)}+\frac{A-1}{nv(A+1)}\right)$$

$$=\frac{1}{n}\left\{X_{1.}+\frac{(v-1)X_{1.}+X_{2.}+nX_{.1}-X_{..}}{A+1}\right\}$$

$$=\frac{1}{n}\left\{X_{1.}+\hat{X}_{11}\right\} \quad .$$

Likewise,

$$\hat{\mu}+t_2=\frac{1}{n}\left\{X_{2.}+\hat{X}_{21}\right\}$$

where the formulae for the two missing plot values, $\hat{X}_{11}$ and $\hat{X}_{21}$, are those given by Federer [1955] in formulae (V-53) and (V-54).

The solutions for the t_i in matrix form are:

$$\begin{pmatrix} t_1 \\ t_2 \\ t_3 \\ \vdots \\ t_v \end{pmatrix} = \begin{pmatrix} \frac{v-1}{v(n-1)} & -\frac{1}{v(n-1)} & -\frac{1}{v(n-1)} & \cdots & -\frac{1}{v(n-1)} \\ -\frac{1}{v(n-1)} & \frac{v-1}{v(n-1)} & -\frac{1}{v(n-1)} & \cdots & -\frac{1}{v(n-1)} \\ -\frac{1}{v(n-1)} & -\frac{1}{v(n-1)} & \frac{1}{n}-\frac{vn-v-2n}{nv(n-1)(v-2)} & \cdots & -\frac{vn-v-2n}{nv(n-1)(v-2)} \\ \vdots & & & & \\ -\frac{1}{v(n-1)} & -\frac{1}{v(n-1)} & -\frac{vn-v-2n}{nv(n-1)(v-2)} & \cdots & \frac{1}{n}-\frac{vn-v-2n}{nv(n-1)(v-2)} \end{pmatrix} \cdot \begin{pmatrix} Q_{1.} \\ Q_{2.} \\ Q_{3.} \\ \vdots \\ Q_{v.} \end{pmatrix}$$

where $Q_{1.} = X_{1.} - \sum_{j=2}^{n} \bar{x}_{.j}$

$$Q_{2.} = X_{2.} - \sum_{j=2}^{n} \bar{x}_{.j}$$

$$Q_{i.} = X_{i.} - \sum_{j=1}^{n} \bar{x}_{.j} \quad \text{for } i=3,4,\cdots,v \quad .$$

The variance of a difference between the two means having a missing plot, say 1 and 2, is

$$V(\hat{\mu}+t_1-\hat{\mu}-t_2=t_1-t_2)=2\sigma_\epsilon^2/(n-1) \quad .$$

The variance of a difference between two treatments with no missing plots, say 3 and 4, is

$$V(t_3-t_4)=2\sigma_\epsilon^2/n \quad .$$

The variance of a difference between two treatment means, one with a missing plot and the other with no missing plots, say 1 and 3, is:

$$V(\hat{\mu}+t_1-\hat{\mu}-t_3=t_1-t_3)$$

$$= \sigma_\epsilon^2 \left\{ \frac{v-1}{v(n-1)} + \frac{1}{n} - \frac{nv-v-2n}{nv(n-1)(v-2)} - \frac{2(-1)}{v(n-1)} \right\}$$

$$= \sigma_\epsilon^2 \left\{ \frac{2}{n} + \frac{v-1}{n(n-1)(v-2)} \right\}$$

k Treatments Missing in One Replicate

Suppose that treatments 1,2,⋯,k (k<v) are missing in replicate one and that the remaining v-k treatments are present in all replicates. After substitution for the $\mu+r_j$, the solutions for the t_i are obtained by subtracting 1/(n-1)(v-k) times the sum of the first k normal equations from each of the last v-k equations, and then, dividing through by n to obtain

$$\begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_k \\ t_{k+1} \\ t_{k+2} \\ \vdots \\ t_v \end{pmatrix} = \begin{pmatrix} \frac{1}{n-1} & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & \frac{1}{n-1} & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & & & & & & & \\ 0 & 0 & \cdots & \frac{1}{n-1} & 0 & 0 & \cdots & 0 \\ -\frac{1}{a} & -\frac{1}{a} & \cdots & -\frac{1}{a} & \frac{1}{n} & 0 & \cdots & 0 \\ -\frac{1}{a} & -\frac{1}{a} & \cdots & -\frac{1}{a} & 0 & \frac{1}{n} & \cdots & 0 \\ \vdots & & & & & & & \\ -\frac{1}{a} & -\frac{1}{a} & \cdots & -\frac{1}{a} & 0 & 0 & \cdots & \frac{1}{n} \end{pmatrix} \cdot \begin{pmatrix} Q_{1\cdot} \\ Q_{2\cdot} \\ \vdots \\ Q_{k\cdot} \\ Q_{k+1} \\ Q_{k+2} \\ \vdots \\ Q_{v\cdot} \end{pmatrix}$$

where $a=n(n-1)(v-k)$, $Q_{i\cdot}=X_{i\cdot}-\sum_{j=2}^{n}\bar{x}_{\cdot j}$ for $i=1,2,\cdots,k$, and $Q_{i\cdot}=X_{i\cdot}-\sum_{j=1}^{n}\bar{x}_{\cdot j}$ for $i=k+1, k+2,\cdots,v$.

The above matrix may be rewritten as (remember that $\Sigma Q_{i\cdot}=0$):

$$\begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_k \\ t_{k+1} \\ t_{k+2} \\ \vdots \\ t_v \end{pmatrix} = \begin{pmatrix} \frac{v-1}{v(n-1)} & -\frac{1}{v(n-1)} & \cdots & -\frac{1}{v(n-1)} & -\frac{1}{v(n-1)} & -\frac{1}{v(n-1)} & \cdots & -\frac{1}{v(n-1)} \\ -\frac{1}{v(n-1)} & \frac{v-1}{v(n-1)} & \cdots & -\frac{1}{v(n-1)} & -\frac{1}{v(n-1)} & -\frac{1}{v(n-1)} & \cdots & -\frac{1}{v(n-1)} \\ \vdots & & & & & & & \\ -\frac{1}{v(n-1)} & -\frac{1}{v(n-1)} & \cdots & \frac{v-1}{v(n-1)} & -\frac{1}{v(n-1)} & -\frac{1}{v(n-1)} & \cdots & -\frac{1}{v(n-1)} \\ -\frac{1}{v(n-1)} & -\frac{1}{v(n-1)} & \cdots & -\frac{1}{v(n-1)} & \frac{1}{n}+b & b & \cdots & b \\ -\frac{1}{v(n-1)} & -\frac{1}{v(n-1)} & \cdots & -\frac{1}{v(n-1)} & b & \frac{1}{n}+b & \cdots & b \\ \vdots & & & & & & & \\ -\frac{1}{v(n-1)} & -\frac{1}{v(n-1)} & \cdots & -\frac{1}{v(n-1)} & b & b & \cdots & \frac{1}{n}+b \end{pmatrix} \cdot \begin{pmatrix} Q_{1\cdot} \\ Q_{2\cdot} \\ \vdots \\ Q_{k\cdot} \\ Q_{k+1} \\ Q_{k+2} \\ \vdots \\ Q_{v\cdot} \end{pmatrix}$$

where $b=(kn+v-nv)/nv(n-1)(v-k)$.

The various variances of differences between treatments are given below. The variance of a difference between two treatments each of which has a missing value in the same block, say t_1 and t_2, is:

$$V(t_1-t_2) = 2\sigma_\epsilon^2/(n-1) \quad .$$

The variance of a difference in means for two treatments which have no missing plots, say t_{k+1} and t_{k+2}, is:

$$V(t_{k+1}-t_{k+2}) = 2\sigma_\epsilon^2/n \quad .$$

The variance of a difference between two treatment means for a treatment with a missing plot and for one with no missing plots, say t_1 and t_{k+1}, is:

$$V(t_1-t_{k+1}) = \sigma_\epsilon^2\left\{\frac{2}{n} + \frac{v-k+1}{n(n-1)(v-k)}\right\} \quad .$$

Two Missing Values for Different Treatments in Different Blocks

Assume that treatment one is missing in replicate one and that treatment two is missing in replicate two. Then, the normal equations become (after using the restrictions that $\sum_{j=1}^{n} r_j = 0 = \sum_{i=1}^{v} t_i$):

$$(nv-2)\hat{\mu}-t_1-t_2-r_1-r_2 = X_{..}$$

$$(n-1)(\hat{\mu}+t_1)-r_1 = X_{1.}$$

$$(n-1)(\hat{\mu}+t_2)-r_2 = X_{2.}$$

$$n(\hat{\mu}+t_i) = X_{i.} \quad \text{for } i=3,4,\cdots,v$$

$$(v-1)(\hat{\mu}+r_1)-t_1 = X_{.1}$$

$$(v-1)(\hat{\mu}+r_2)-t_2 = X_{.2}$$

$$v(\hat{\mu}+r_j) = X_{.j} \quad \text{for } j=3,4,\cdots,n$$

After solving for r_1 and r_2 in terms of t_1, t_2 and the observations the following equations in the t_i result:

$$t_1 = \frac{v-1}{nv-n-v}\left\{X_{1.}+\frac{X_{.1}}{v-1}-n\hat{\mu}\right\}$$

$$t_2 = \frac{v-1}{nv-n-v}\left\{X_{2.}+\frac{X_{.2}}{v-1}-n\hat{\mu}\right\}$$

$$t_i = \frac{X_{i.}}{n}-\hat{\mu}, \quad \text{for } i=3,4,\cdots,v$$

Since the $\sum_{i=1}^{v} t_i = 0$, we sum the above equations and obtain

$$\hat{\mu} = \frac{v(X_{1.}+X_{2.})+n(X_{.1}+X_{.2})+X_{..}(nv-n-v)}{nv(nv-n-v+2)}$$

Also,

$$t_1 = \frac{v-1}{v(nv-v-n)(nv-n-v+2)}\Big\{v(n-1)(v-1)X_{1.}$$

$$+X_{.1}\Big(\frac{v(nv-n-v+2)-n(v-1)}{(v-1)}\Big)-vX_{2.}-nX_{.2}-X_{..}(nv-n-v)\Big\}$$

and

$$\hat{\mu}+t_1 = \frac{v-1}{nv-n-v}\Big\{X_{1.}+\frac{X_{.1}}{v-1}-n\hat{\mu}+\frac{nv-n-v}{v-1}\hat{\mu}\Big\}$$

$$= \frac{X_{1.}}{n}+\frac{vX_{1.}+nX_{.1}-nv\hat{\mu}}{n(nv-n-v)}$$

$$= \frac{1}{n}\Big\{X_{1.}+\hat{X}_{11}\Big\} ,$$

where $\hat{X}_{11}$ is obtained from formula (V-55) on page 134 (loc. cited).

Likewise,

$$t_2 = \frac{v-1}{v(nv-v-n)(nv-v-n+2)}\Big\{v(n-1)(v-1)X_{2.}$$

$$+X_{.2}\Big(\frac{v(nv-n-v+2)-n(v-1)}{v-1}\Big)-vX_{1.}-nX_{.1}$$

$$-X_{..}(nv-n-v)\Big\}$$

and

$$\hat{\mu}+t_2 = \frac{1}{n}\Big\{X_{2.}+\hat{X}_{22}\Big\} ,$$

where $\hat{X}_{22}$ is computed from formula (V-56) on page 134 of Federer [1955].

In matrix form, the solutions for the t_i are:

$$\begin{pmatrix} t_1 \\ t_2 \\ t_3 \\ t_4 \\ \vdots \\ t_v \end{pmatrix} = \begin{pmatrix} a & c & 0 & 0 & \cdots & 0 \\ c & a & 0 & 0 & \cdots & 0 \\ b & b & n^{-1} & 0 & \cdots & 0 \\ b & b & 0 & n^{-1} & \cdots & 0 \\ & & & & & \\ b & b & 0 & 0 & \cdots & n^{-1} \end{pmatrix} \cdot \begin{pmatrix} X_{1.} - \sum_{j=2}^{n} \bar{x}_{.j} = Q_{1.} \\ X_{2.} - \bar{x}_{.1} - \sum_{j=3}^{n} \bar{x}_{.j} = Q_{2.} \\ X_{3} - \sum_{j=1}^{n} \bar{x}_{.j} = Q_{3.} \\ X_{4.} - \sum_{j=1}^{n} \bar{x}_{.j} = Q_{4.} \\ \vdots \\ X_{v.} - \sum_{j=1}^{n} \bar{x}_{.j} = Q_{v.} \end{pmatrix}$$

where $a = (n-1)(v-1)^2/[(n-1)^2(v-1)^2-1]$,

$b = -1/n(nv-n-v+2)$, and

$c = -(v-1)/[(n-1)^2(v-1)^2-1]$.

The various variances are:

$V(t_1-t_2) = 2\sigma_\epsilon^2(a-c)$,

$V(t_1-t_3) = \sigma_\epsilon^2(a+n^{-1}-c)$, and

$V(t_3-t_4) = 2\sigma_\epsilon^2/n$.

<u>k Treatments With One Missing Plot in Each of k Replicates</u>

Suppose that treatments $1,2,\cdots,k$ ($k \leq$ the smaller of n and v) each have one yield missing in such a way that no two treatments have plots missing in the same replicate. The remaining v-k treatments have no missing yields in the n replicates. Then, after substituting for $\mu+r_j$ values in the normal equations for the treatments, subtract the sum of the first k equations divided by $(n-1)(v-1)+(k-1)$

from each of the last equations, and subtract the sum of the other k-1 equations divided by (n-1)(v-1)-(k-2) from each of the first k equations. After division by the coefficient for each of the t_i the following solutions are obtained:

$$\begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_k \\ t_{k+1} \\ t_{k+2} \\ \vdots \\ t_v \end{pmatrix} = \begin{pmatrix} A & C & \cdots & C & 0 & 0 & \cdots & 0 \\ C & A & \cdots & C & 0 & 0 & \cdots & 0 \\ \vdots & & & & & & & \\ C & C & \cdots & A & 0 & 0 & \cdots & 0 \\ B & B & \cdots & B & n^{-1} & 0 & \cdots & 0 \\ B & B & \cdots & B & 0 & n^{-1} & \cdots & 0 \\ \vdots & & & & & & & \\ B & B & \cdots & B & 0 & 0 & \cdots & n^{-1} \end{pmatrix} \cdot \begin{pmatrix} Q_{1\cdot} \\ Q_{2\cdot} \\ \vdots \\ Q_{k\cdot} \\ Q_{k+1\cdot} \\ Q_{k+2\cdot} \\ \vdots \\ Q_{v\cdot} \end{pmatrix}$$

where $$A = \frac{(v-1)[(n-1)(v-1)+k-2]}{(n-1)(v-1)[(n-1)(v-1)+k-2]-k+1}$$

$$B = -\frac{1}{n[(n-1)(v-1)+k-1]}$$

$$C = -\frac{(v-1)}{(n-1)(v-1)[(n-1)(v-1)+k-2]-k+1}$$

The variance of a difference between two treatments with missing plots, say t_1 and t_2, is:

$$V(t_1-t_2) = 2\sigma_\epsilon^2(A-C)$$

The variance of a difference between two treatments with no missing plots, say t_{k+1} and t_{k+2}, is:

$$V(t_{k+1}-t_{k+2}) = 2\sigma_\epsilon^2/n \quad .$$

The variance of a difference between a treatment with a missing plot and one with no missing plots, say t_1 and t_{k+1}, is:

$$V(t_1-t_{k+1}) = \sigma_\epsilon^2 \left\{ \frac{1}{n} + A - B \right\} \quad .$$

Literature Cited

Allan, F. E. and Wishart, J., 1930, A method of estimating yield of a missing plot in field experimental work. Jour.Agric. Sci. 20:399-406.

Federer, W. T., 1951, Evaluation of variance components from a group of experiments with multiple classifications. Iowa Agric. Exp. Sta. Res. Bul. 380:241-310.

Federer, W. T., 1955, Experimental design - theory and application. Macmillan, N. Y.

Thompson, H. R., 1956, Extensions to missing plot techniques. Biometrics 12:241-244.

Yates, F., 1933, The analysis of replicated expeiments when field results are incomplete. Emp. Jour. Exp. Agric. 1:129-142.