SAMPLING STRUCTURES FOR SPLIT PLOT AND SPLIT BLOCK DESIGNS

BU-551-M* March 1975

by

Walter T. Federer

Abstract

The sampling structures, types of variation, and an additive and a multiplicative model are discussed for the split plot designs. Three types of heterogeneity for the whole plot experimental units and two types of heterogeneity of split plot or subplot experimental units are considered. The experiment designs appropriate for the above three types of heterogeneity among whole plots are the completely randomized, the randomized complete block, and a class of row-column designs of which the Latin square is a member. The randomized complete block and the Latin square designs were considered as appropriate experiment designs to control the heterogeneity among split plot experimental units. A plant experiment and an animal experiment are discussed, and it is shown how to use a split plot design to control the various sources of variation as described above. Other practical applications are discussed.

Likewise, the sampling structure and variation are considered for the split block, two-way whole plot, or strip-block design. Two situations are considered to control heterogeneity among the whole plots, one wherein the variation is such that both sets of whole plots require a randomized complete block design and one wherein the variation is such that one set of whole plot treatments requires a randomized complete block design and the other set of whole plot

---

* In the Mimeo Series of the Biometrics Unit, Cornell University.

treatments requires a Latin square design to control the heterogeneity. Practical applications of the split block design are discussed.

In the last section of the paper some aids are given to help in recognizing when the experiment was designed as a split plot, as a split block, or as some other experiment design. Four examples are presented to indicate the necessity of determining the nature of an appropriate error variance for a treatment mean or contrast. The appropriate error variance has to be determined, not defined as it is in statistical literature. Definitions of what constitutes an error variance (e.g., given that "σ^2 is the error variance, we now proceed . . ."<u>may</u> be alright in a classroom, but in applying statistical procedures to actual experimental data, it is essential to determine the sampling structure, the nature of variation and response prior to and after the application of treatments, and the appropriate error variance or mean square for treatments means and contrasts.

SAMPLING STRUCTURES FOR SPLIT PLOT AND SPLIT BLOCK DESIGNS

by

Walter T. Federer

BU-551-M*

March 1975

1. Introduction

Sampling structures for experiment designs controlling zero-way, one-way, and two-way elimination of heterogeneity have been discussed in previous papers (Federer [1975a, 1975b]). After some preliminary definitions, our purpose here is to discuss the nature of the variation and the sampling procedure when split plot and split block designs are considered to be appropriate for the experiment under consideration. The whole plot treatments of the split plot experiment are laid out in a completely randomized, in a randomized complete block, and in a Latin square experiment design with the split plot treatments being laid out in either a randomized complete block or in a Latin square experiment design. Combinations of additive and multiplicative treatment effects are considered. Practical examples of split plot designs are presented.

The split block or two-way whole plot designs are discussed for two situations only, i.e., both whole plot treatments in a randomized complete block design and one set of whole plots in a randomized complete block design and the second set of whole plots in a Latin square design. The sampling structures for which these are the appropriate experiment designs, are discussed. Examples where these designs are useful are presented.

It is sometimes difficult to recognize a split plot or a split block design. Some suggestions are made to aid in and some rules are given for recognizing the designs.

---

* In the Mimeo Series of the Biometrics Unit, Cornell University.

2. Some Preliminary Definitions Relative to a Split Plot Design

The treatment design usually associated with a split plot or split block experiment design is a two factor factorial with a levels of the first factor A, say $A_1, A_2, \cdots, A_a$, and with b levels of the second factor B, say $B_1, B_2, \cdots, B_b$. Occasionally, the levels of factor B are nested within each level of factor A, but the former treatment design is the usual one. Note also that the levels of a factor may be different amounts of a chemical, different brands of a product, different medical vaccines, different crop varieties, different eye shadow tints, and so forth. That is, there may or there may not be a natural ordering among the levels of a given factor. Also, the levels of a factor, say B, may themselves constitute a factorial treatment design. Suppose that the levels of factor A, the whole plot treatments, are applied to a different size experimental unit than are the levels of factor B, the split plot treatments. Then we need to define a whole plot experimental unit as the smallest amount of material to which one level of factor A, or one whole plot treatment, is applied. Then, a whole plot experimental unit is subdivided into b units and a level of factor B, the split plot treatment, is allotted to one of these b units, and this constitutes the split plot experimental unit. Hence, as opposed to the completely randomized, the randomized complete block, and the Latin square experiment designs, there are two different experimental units to be considered.

Since the whole plot experimental unit (wpeu) is b times larger than the split plot experimental unit (speu), the variance among wpeu's will generally be larger than the variance among speu's. Hence, the error variance for differences between whole plot treatment means will be larger than for differences between split plot treatment means. Thus, if less information is desired for factor A than for B, the levels of A may constitute the whole plot treatments. On the

other hand, it may be impossible to use levels of factor A in any other manner than as whole plot treatments, despite the relative amount of information desired. In summary then, if less information is desired and/or if it is impossible to have smaller experimental units, the levels of factor A become the whole plot treatments and the levels of factor B become the split plot treatments.

Depending upon the nature of variation for wpeu's, a completely randomized design, a randomized complete block design, a Latin square design, or some other design may be selected for the whole plot treatments. Likewise, an appropriate experiment design for controlling variation within wpeu's within <u>each</u> level of whole plot treatments, should be used for the split plot treatments.

A schemmatic layout for a split plot experiment design with the a = 3 whole plots in a randomized complete block design with r = 4 blocks and with the b = 4 split plot treatments allotted at random to the 4 speu's within each wpeu, is:

Block I			Block II			Block III			Block IV		
A_1	A_3	A_2	A_3	A_1	A_2	A_2	A_3	A_1	A_2	A_1	A_3
B_2	B_4	B_1	B_2	B_4	B_3	B_3	B_4	B_1	B_2	B_3	B_1
B_1	B_1	B_4	B_1	B_3	B_1	B_1	B_3	B_3	B_1	B_2	B_4
B_4	B_3	B_3	B_4	B_1	B_2	B_2	B_2	B_4	B_4	B_1	B_3
B_3	B_2	B_2	B_3	B_2	B_4	B_4	B_1	B_2	B_3	B_4	B_2

The whole plot for treatment A_i is 4 speu's. Each wpeu contains all B_j treatments.

3. The Sampling Structure of Split Plot Designs

We shall discuss four particular types of sampling structures for split plot designs, each of which requires a different experiment design for either the whole plot treatments or for the split plot treatments. For each situation we shall consider two types of treatment responses, additive and multiplicative, when a particular combination is applied to a split plot experimental unit (speu) or to a whole plot experimental unit (wpeu).

For the first situation, suppose that there are N subpopulations making up the total population. Then, suppose that the response for a randomly selected individual from a randomly selected subpopulation is expressed as:

$$Y_{hk} = \mu_h + \epsilon_{hk} = \mu + (\mu_h - \mu = \delta_h) + \epsilon_{hk} , \qquad (3.1)$$

where the h^{th} subpopulation and the k^{th} sampling unit within the h^{th} subpopulation are selected. Suppose further that the δ_h are independently and identically distributed with zero mean and variance σ_δ^2 (denote this by IID$(0,\sigma_\delta^2)$) and that the ϵ_{hk} are IID$(0,\sigma_\epsilon^2)$. Note that we are assuming that the variance is constant from subpopulation to subpopulation. Now suppose that we obtain a simple random sample of ra subpopulations, and a simple random sample of b sampling units from EACH of the ra subpopulations. Now we shall randomly allocate the level, say $j=1,2,\cdots,b$, of a treatment factor B to the sampling units in each subpopulation. Thus, the sampling situation is the same as for the randomized complete block design with ra complete blocks as described in BU-548-M. The yields or responses for the additive and multiplicative situations resulting from the application of the j^{th} level of treatment factor B are:

Additive case

$$Y_{hj} = \mu + \delta_h + \epsilon_{hj} + \beta_j \tag{3.2}$$

Multiplicative case

$$Y^*_{hj} = \beta^*_j(\mu+\delta_h+\epsilon_{hj}) = \mu^*_{\cdot j} + \beta^*_j\delta_h + \beta^*_j\epsilon_{hj} \tag{3.3}$$

The expected values in the two cases are assumed to be $E[Y_{hj}] = \mu + \beta_j = \mu_{\cdot j}$ and $E[Y^*_{ij}] = \mu^*_{\cdot j}$, that is, the fixed treatment effects case. $\mu_{\cdot j}$ and $\mu^*_{\cdot j}$ are defined in the same manner as in BU-548-M and BU-550-M.

The next step is to obtain a simple random sample of size r of the ra subpopulations samples of size b. Now the set of b experimental units from subpopulation h forms the wpeu for the first level of treatment factor A. Note that the sample size for a whole plot treatment, say A_i, is r, whereas the sample size for a split plot treatment, say B_j, is ra. Also, note that the speu is one sampling unit and that the wpeu is b speu's.

The experimental yields or responses of a speu after application of A_i to a wpeu and B_j to a speu may be expressed as follows for the additive and multiplicative effects situations. (Note that we shall use a multiplicative treatment effects model of the form $\alpha_i\beta_j = \alpha_i + \beta_j + \alpha\beta_{ij}$.):

A and B additive

$$Y_{hij} = \mu + \delta_{hi} + \alpha_i + \beta_j + \alpha\beta_{ij} + \epsilon_{hij} \tag{3.4}$$

A multiplicative, B additive

$$\begin{aligned} Y^*_{hij} &= \alpha^*_i(\mu+\delta_h+\epsilon_{hj}+\beta_j) \\ &= \mu^*_{\cdot i\cdot} + \alpha^*_i\delta_h + \epsilon^*_{hij} + \alpha^*_i + \beta_j + \alpha^*\beta_{ij} \end{aligned} \tag{3.5}$$

A additive, B multiplicative

$$Y'_{hij} = \mu^*_{\cdot j} + \beta^*_j \delta_{hi} + \epsilon^*_{hij} + \alpha_i + \alpha\beta^*_{ij} \tag{3.6}$$

A and B multiplicative

$$\begin{aligned} Y''_{hij} &= \alpha^*_i(\mu^*_{\cdot j} + \beta^*_j \delta_h + \beta^*_j \epsilon_{hj}) \\ &= \alpha^*_i \beta^*_j (\mu + \delta_h + \epsilon_{hk}) \\ &= (\alpha^*_i + \beta^*_j + \alpha^*\beta^*_{ij})(\mu + \delta_h + \epsilon_{hk}) \end{aligned} \tag{3.7}$$

With respect to levels of factor A in the above sampling situation, a completely randomized design (crd) resulted, with r wpeu's for each level. With respect to levels of factor B, a randomized complete block design (rcbd) or r blocks is used for each level of factor A. Thus, there are a such randomized complete block designs, one for each level of factor A. Note that ra randomizations were used for the levels of factor B, but that only r randomizations were used for levels of factor A. A key-out of the total degrees of freedom in the analysis of variance would be:

Source of variation	Degrees of freedom	Mean square
Total	rab	
Correction for mean	1	
Whole plot analysis		
Among wpeu's	(ra-1)	
Among levels of factor A=A	(a-1)	T_A ↖
Within levels of A	a(r-1)	E_A ↗
Split plot analysis		
Within wpeu's	ra(b-1)	
Among levels of factor B in A_1=B:A_1	(b-1)	T_{B1} ↖
B × blocks:A_1	(r-1)(b-1)	E_{B1} ↗
Among levels of B within A_2=B:A_2	(b-1)	T_{B2} ↖
B × blocks:A_2	(r-1)(b-1)	E_{B2} ↗
⋮		
Among levels of B within A_a=B:A_a	(b-1)	T_{Ba} ↖
B × blocks:A_a	(r-1)(b-1)	E_{Ba} ↗

The arrows indicate where F-tests may be performed given that model equation (3.4) and normality hold.

An alternative analysis of variance could be obtained as follows (This is the one universally presented in statistics texts):

Source of variation	Degrees of freedom	Mean square
Total	rab	
Correction for mean	1	
Whole plot analysis		
Among levels of factor A=A	(a-1)	T_A
Within levels of A=error (a)	a(r-1)	E_A
Split plot analysis		
Among levels of B over all levels of A=B	(b-1)	T_B
A x B	(a-1)(b-1)	$T_{A\times B}$
B x blocks within levels of A= error (b)	a(r-1)(b-1)	E_B

Note that the sums of squares $B:A_i$ summed over all i is equal to the B plus the A x B sum of squares in the above analysis. Likewise, the "error (b) sum of squares" is the sum of the B x blocks sums of squares over all levels of A.

The variance of a difference between two means for levels of A_i, that is $\bar{y}_{.i.} - \bar{y}_{.i'.}$ for $i \neq i'$, is $V(\bar{y}_{.i.}-\bar{y}_{.i'.}) = \frac{2}{rb}(\sigma_\epsilon^2+b\sigma_\delta^2)$ and is estimated as $\hat{V}(\bar{y}_{.i.}-\bar{y}_{.i'.}) = 2E_a/rb$. The variance of a difference between two B_j level means is $V(\bar{y}_{..j}-\bar{y}_{..j'}) = 2\sigma_\epsilon^2/ra$ and is estimated by $\hat{V}(\bar{y}_{..j}-\bar{y}_{..j'}) = 2E_b/ra$. The estimated variance of a difference between two B level means within the i^{th} A level is $\hat{V}(\bar{y}_{.ij}-\bar{y}_{.ij'}) = 2E_b/r$.

For the above situation, the whole plot treatments were in a completely randomized design and the split plot treatments were in a randomized complete block design within <u>each</u> whole plot treatment experimental units. For the second situation, suppose that the N subpopulations may be subdivided into groups of subpopulations which are more alike within groups than between groups. This means that the whole plot treatments should be in a randomized complete block

design since it will be necessary to control one source of variation among whole plots. Prior to the application of any treatment combination, the response equation may be represented as:

$$Y_{ghk} = \mu_{g\cdot} + \delta_{gh} + \epsilon_{ghk} = \mu + \rho_g + \delta_{gh} + \epsilon_{ghk} \qquad (3.8)$$

where the k^{th} sampling unit in the h^{th} subpopulation in the g^{th} group is obtained, the ρ_g are $IID(0,\sigma^2_\rho)$, the δ_{gh} are $IID(0,\sigma^2_\delta)$, and the ϵ_{ghk} are $IID(0,\sigma^2_\epsilon)$. Note this assumes a constant variance among sampling units regardless of subpopulation, and a constant variance among subpopulation means within groups.

If a level of B is randomly allocated to a sampling unit (the speu) from a randomly selected subpopulation, say gh, within the g^{th} group, if a simple random sample of r groups is obtained, and if a level of factor A is applied to the k speu's from subpopulation gh, the additive and multiplicative response equations for a yield from a speu may be expressed as:

A and B additive

$$Y_{gij} = \mu + \rho_g + \delta_{gi} + \alpha_i + \beta_j + \alpha\beta_{ij} + \epsilon_{gij} \qquad (3.9)$$

A multiplicative, B additive

$$\begin{aligned} Y^*_{gij} &= \alpha^*_i(\mu+\rho_g+\delta_{gh}+\beta_j+\epsilon_{ghj}) \\ &= \mu^*_i + \alpha^*_i\rho_g + \alpha^*_i\delta_{gi} + \beta_j + \alpha^*\beta_{ij} + \epsilon^*_{gij} \end{aligned} \qquad (3.10)$$

A additive, B multiplicative

$$Y'_{gij} = \alpha_i + \alpha\beta^*_{ij} + \beta^*_j(\mu+\rho_g+\delta_{gi}+\epsilon_{gik}) \qquad (3.11)$$

A and B multiplicative

$$Y''_{gij} = (\alpha^*_i+\alpha^*\beta^*_{ij}+\beta^*_j)(\mu+\rho_g+\delta_{gh}+\epsilon_{ghk})$$

$$= \alpha^*_i\beta^*_j(\mu+\rho_g+\delta_{gh}+\epsilon_{ghk}) \qquad (3.12)$$

$$= \alpha^*_i(\mu^*_j+\beta^*_j(\rho_g+\delta_{gh})+\epsilon''_{ghj})$$

Given that model equation (3.9) holds, the key-out of degrees of freedom in the analysis of variance is as follows:

Source of variation	Degrees of freedom	Mean square
Total	rab	
Correction for the mean	1	
Whole plot analysis		
Block = R	r-1	
Levels of A = A	a-1	T_A
R × A = error (a)	(r-1)(a-1)	E_A
Split plot analysis		
same as for the previous design		

For the third situation, suppose that a Latin square design is appropriate for whole plot treatments. Then, from BU-550-M, the yield equation, prior to application of treatments, may be expressed as:

$$Y_{fghk} = \mu + \rho_f + \gamma_g + \delta_{fgh} + \epsilon_{fghk} \qquad (3.13)$$

where the above is the response of a randomly selected sampling unit k from the h^{th} subpopulation of the f^{th} "row" grouping and g^{th} "column" grouping of groups of subpopulations and where the ρ_f are $IID(0,\sigma^2_\rho)$, the γ_g are $IID(0,\sigma^2_\gamma)$, the δ_{fgh} are $IID(0,\sigma^2_\delta)$, and the ϵ_{fghk} are $IID(0,\sigma^2_\epsilon)$. Here, again note the assumptions

of constant variance. If the treatment effects are additive, then the yield equation becomes:

$$Y_{fgij} = \mu + \rho_f + \gamma_g + \alpha_i + \delta_{fgi} + \beta_j + \alpha\beta_{ij} + \epsilon_{fgij} \qquad (3.14)$$

Given that the above model yield equation holds, the key-out of degrees of freedom in the analysis of variance is:

Source of variation	Degrees of freedom	Mean Square
Total	$rab=a^2b$	
Correction for mean	1	
Whole plot analysis		
"Rows"	a-1	
"Columns"	a-1	
Levels of A	a-1	T_A
Error (a)	(a-1)(a-2)	E_A
Split plot analysis		
same as for the previous design		

For the fourth situation involving heterogeneity among the wpeu's and the speu's prior to application of the treatments, consider the preceding situation, but, in addition, a third dimensional layering exists in the RC subpopulations described in section 3 of BU-550-M. In order to be realistic it appears best to consider this layering as a fixed effect rather than as a random effect. The experiment design for split plot treatments will be affected, but the experiment design for whole plot treatments will be unaffected by this additional stratification. The nature of the variation in response for a randomly selected sampling unit is considered to be:

$$Y_{fgh} = \mu + \rho_f + \gamma_g + \pi_\ell + \epsilon_{fghk\ell} \qquad (3.15)$$

where μ, ρ_f, γ_g and δ_{fgh} are defined in (3.13), π_ℓ represents the ℓ^{th} subsubpopulation layering in the fh^{th} subpopulation and $\epsilon_{fghk\ell}$ is a random deviation in the subsubpopulation. The $\sum_{\ell=1}^{b} \pi_\ell = 0$ and the $\epsilon_{fghk\ell}$ are $IID(0,\sigma_\epsilon^2)$. If the effects of factors A and B are both additive, the yield response becomes:

$$Y_{fgi\ell j} = \mu + \rho_f + \gamma_g + \alpha_i + \epsilon_{fgi} + \beta_j + \alpha\beta_{ij} + \pi_{i\ell} + \epsilon_{fgij\ell} \qquad (3.16)$$

Note that an i subscript has been added to π_ℓ in order to denote that the ordering of the π's is within each level of i. This need not be done, but the statistical analyses become considerably more complicated if it is not. Hence, in order to retain relatively simple statistical analyses we shall take account of orders within each level of factor A. The analyses of variance for the various levels of factor A given that r = b = a and that a Latin square design is used for each level of factor A, are:

Source of variation	Level of A: A_1 d.f.	A_1 s.s.	A_2 d.f.	A_2 s.s.	...	A_a d.f.	A_a s.s.	Total d.f.	Total s.s.
Total	b^2	T_1	b^2	T_2		b^2	T_a	ab^2	$\sum_{i=1}^{a} T_i$
Correction for mean	1	A_1	1	A_2		1	A_a	a	$\sum_{i=1}^{a} A_i$
Whole plots for A_i	b-1	W_1	b-1	W_2		b-1	W_a	a(b-1)	$\sum_{i=1}^{a} W_i$
Orders within A_i	b-1	O_1	b-1	O_2		b-1	O_a	a(b-1)	$\sum_{i=1}^{a} O_i$
Levels of B	b-1	B_1	b-1	B_2		b-1	B_a	a(b-1)	$\sum_{i=1}^{a} B_i$
Remainder	(b-1)(b-2)	R_1	(b-1)(b-2)	R_2		(b-1)(b-2)	R_a	a(b-1)(b-2)	$\sum_{i=1}^{a} R_i$

The above analyses of variance may be combined into a single analysis of variance table as follows:

Source of variation	d.f.	Sum of squares	Mean square
Total	$ab^2=a^3=b^3$	ΣT_i	
Correction for mean	1	M	
Whole plot analysis			
Whole plot treatments = A	a-1	ΣA_i - M	T_A
Within whole plots	a(b-1)	ΣW_i	
Rows	a-1	compute = R	
Columns	a-1	compute = C	
Remainder = error (a)	(a-1)(a-2)	ΣW_i - R - C	E_a
Split plot analysis			
Levels of B within A	a(b-1)=a(a-1)	ΣB_i	
B	(a-1)	compute = B	T_B
A x B	$(a-1)^2$	ΣB_i - B	T_{AxB}
Orders within A_i	a(b-1)=a(a-1)	ΣO_i	
Remainder within A_i = error (b)	a(a-1)(a-2)	ΣR_i	E_b

The only difference between this analysis of variance and the preceding one is that the former error (b) sum of squares has been subdivided into two parts in the present analysis of variance, that is "orders within A_i" and "Remainder within A_i", with the latter being considered to be the error (b) sum of squares for the present analysis.

If either or both factor A and B effects enter in a multiplicative manner for the last two situations described above, model yield equations similar to equations (3.10) to (3.12) may be formulated.

4. Examples of Split Plot Designs

One plant and one animal experiment designed as a split plot design for each of the four situations discussed in the previous section are discussed below:

Factor A in a completely randomized design and factor B in a randomized complete block within the i^{th} level of A.

Consider a situation wherein the levels of factor A represent a different nitrogeneous fertilizers and the levels of factor B represent b different varieties of wheat. The response of interest is yield of grain per acre in bushels. The nature of the fertilizer treatment application dictates rather large experimental units, whereas the experimental units for varieties can be made relatively small. Also, interest in this experiment is centered on varietal mean differences and on variety by fertilizer interactions, with less information being desired on fertilizer mean differences. Suppose that in the area of land available for experimentation, there are no visible gradients or logical groups of the whole plot experimental units that can be made. Hence, one fertilizer is randomly allotted to r of the ra wpeu's, a second fertilizer is randomly allotted to r of the remaining ra-r wpeu's, and so forth. Each wpeu contains b speu's. Randomly allot the b varieties to the b speu's within each wpeu, using a different randomization for each wpeu.

For the animal experiment, consider the situation wherein ra litters of swine are available for experimentation and where the litters are all from the same sire and same management regime. No knowledge is available for grouping the litters in any way, and the litters all have b animals per litter. Each litter is confined to one pen and given one of a diets, a level of factor A being a diet which contains a major nutritional element. The b levels of factor

B are minor nutritional elements with the speu being a single animal. Comparisons among minor elements are made within a litter and hence are more precise than comparisons among litters. The major element diet is randomly allocated to the litter and pen until r have been allocated for each diet; the minor element nutritional supplements are randomly allotted to an animal within each litter. The response of interest is weight gain over a 54-day period.

Factor A in a completely randomized design and factor B in a randomized complete block within the i^{th} level of factor A.

For the plant experiment in the preceeding example, suppose that the field in which the experiment is performed is terraced and that it is known that the area within a terrace is relatively uniform but that the differences between terraces is relatively large. A block of a wpeu's is set up on each terrace and the a nitrogeneous fertilizers are randomly allocated to the wpeu's in each block, there being r such randomizations. The design for varieties in the split plots is identical to that described above.

For the above animal experiment, consider the situation in which a set of a litters is obtained from each of r sires. The litters of each sire then forms a block for the a major element diets. These diets are randomly allocated to a litter from a given sire. The allocation of the minor element dietary supplements is identical to the first animal example above.

Factor A in a Latin square design and factor B in a randomized complete block design within the i^{th} level of factor A.

For the previous plant example, consider that in addition to a difference between terraces, there is a gradient along the terrace with the gradients being similar from terrace to terrace. The a fertilizer treatments are set up in a

Latin square design with the terraces being the "rows" and the position along the terraces being the "columns". A randomly selected Latin square arrangement is used. The design for the levels of factor B, the split plot treatments, is identical to the preceding two examples.

For the previous animal experiment, suppose that a sires are mated to a dams from a different breeds with a^2 litters being available. The different sires, then, would constitute the "columns" of the Latin square, and the different breeds of dams would constitute the "rows" of the Latin square. Again, a randomly selected Latin square arrangement is used to allocate the a major element diets to specified litters and pens. The experiment design for the b minor element supplements is identical to the preceding two situations.

Factor A in a Latin square design and factor B in a Latin square design within the i^{th} level of factor A.

Consider now, for the previous plant example, that in addition to variation along the terrace and between terraces, there is a gradient from the top of a terrace to the bottom of a terrace. Also, suppose that the variety plots within a fertilizer plot must be laid out parallel to a terrace rather than perpendicular to it. This means that the topside of a terrace yields differently from the bottomside. To take account of this differential gradient, the split plot treatments, levels of factor B or varieties, are laid out in a Latin square design within each level of factor A. (Note: It is essential to lay out the Latin square design within each level of factor A in order to preserve simplicity in the statistical analysis. Any other lay out considerably complicates the analysis.) The "rows" of this Latin square design are the wpeu's for the i^{th} level of factor A, and the "columns" are the orders from the top to the bottom of the terrace. A randomly selected Latin square arrangement is used for each of the a whole

plot treatments. The whole plot treatments, fertilizer treatments, are also laid out in a Latin square design as described in the preceding subsection.

For the animal experiment described in the preceding subsection, suppose that order of birth of animal in a litter affects the response in weight from a minor nutritional element. Then, for the i^{th} level of factor A, the "rows" of the Latin square design would be the litters and pens receiving the i^{th} level of factor A, and the "columns" would be order of birth of an animal within a litter. The levels of factor A would be designed in an a x a Latin square design as described in the previous subsection.

Four different experimental situations relative to variation in the population structure, have been discussed above. Many more such situations are possible. The above demonstrate the necessity of knowing the nature of variation in the response of interest in order to set up efficient experiments.

In another field of experimentation, consider the following situation. An experimenter obtains a 5' x 5' sample of 6 brands of rugs. At two other times she received two more such samples for the 6 brands of rugs. She believes that the rug samples represent a random sample of size three for each brand of rug. She plans to measure 1' x 5' rug samples at 5 different lengths of wear (1, 3, 5, 9, and 12 months) for amount of static electricity. The measurement is destructive. A long, heavily-trafficed hall is available for experimentation. She cuts each sample of rug into 5 1' x 5' samples, resulting in 90 such pieces. She places all 30 pieces from the first sample in the first 30' section, all 30 pieces from the second sample of 6 brands received in the middle 30' of the hall, and the third sample of 6 brands in the last 30' of the hall. The 30 pieces for each sample were randomly allotted to the 30 positions in the hall for each 30' segment. An appropriate analysis of variance for the above experiment would be:

Source of variation	d.f.	Mean square
Total	90	
Correction for mean	1	
Whole plot analysis		
Blocks = segments of hall	2	
Brands = A	5	T_A
Brands x blocks = error (a)	10	E_a
Split plot analysis		
Lengths of wear = B	4	T_B
A x B	20	T_{AxB}
B x blocks : A = error (b)	48	E_b

The question may arise as to why brands are considered to be whole plots. The answer is obtained from the sampling procedure. Remember that a single sample of rug 5' x 5' of each brand was obtained. Hence, as far as brands are concerned the wpeu is 5' x 5' and the length of wear treatment speu is 1' x 5'. The comparisons among lengths of wear are made within samples of 5' x 5' rugs.

The last example of a split plot design is taken from Warriner [1951] and involves a measure of readability of 8 high school physics texts. Each text is divided into 10 comparable parts. Then the following experiment using 30 members of a high school class was set up where each student rates all 8 texts for readability on the part allocated:

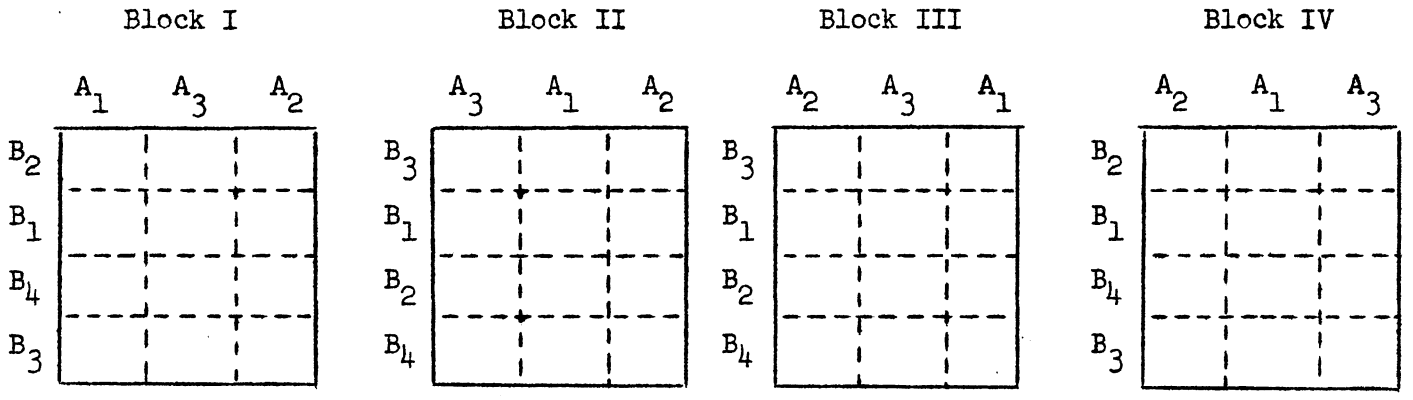
	Part of text					
	1	2	3	4	...	10
	Student	Student	Student	Student		Student
Text	1 2 3	4 5 6	7 8 9	10 11 12		28 29 30
1						
2						
3						
4						
5						
6						
7						
8						
Totals						

An analysis of variance for the above experiment is:

Source of variation	d.f.	Mean square
Total	240	
Correction for mean	1	
Whole plot analysis		
Among parts of text = P	9	T_P
Among students within parts	20	E_P
Split plot analysis		
Among texts = T	7	T_T
T x P	63	T_{TxP}
T x students within parts	140	E_T

5. Split Block Designs and Their Sampling Structure

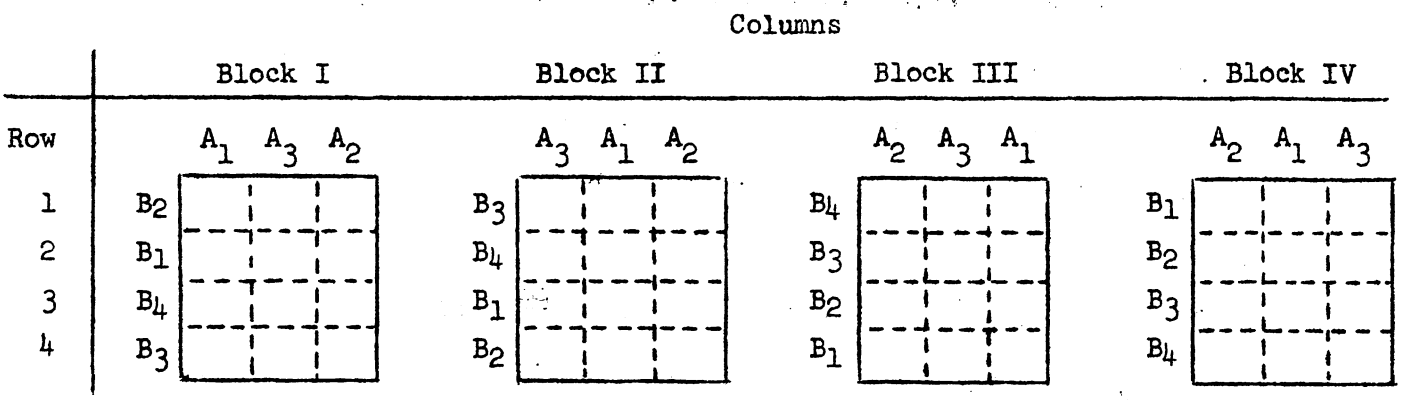
As stated previously, a split block design has two sets of whole plot treatments. To illustrate, consider a two factor factorial with a levels of factor A and b levels of factor B in all possible combinations. The whole plot experimental units for levels of factor A go across all levels of factor B and vice versa. An example for a = 3, b = 4, and r = 4 blocks is used to illustrate this:



There are three types of experimental units for this design. The experimental unit which receives a particular A_i and a particular B_j is denoted as the split block experimental unit (sbeu), whereas the smallest unit receiving one level of factor A is b = 4 sbeu's and is denoted as the whole plot experimental unit for factor A (wpeua). The smallest unit of experimental material receiving one level of factor B is a = 3 sbeu's and is denoted as the whole plot experimental unit for factor B (wpeub).

Let us consider two special cases of variation prior to application of levels of factors A and B. For the first case, a randomized complete block design is indicated for both levels of factors A and B, and for the second case, a randomized complete block design is indicated for factor A and a Latin square design of order r = b is indicated for levels of factor B. The above lay out is an example of the

first situation, and the one below for a = 3, r = b = 4 is an example for case 2:



Note that every level of B occurs once in each row and once in each block (column) of a 4 x 4 Latin square.

In certain types of experiments it is quite difficult, if not impossible, to utilize a split plot or randomized complete block design due to the nature of applying the treatments. For example, field application of fumigants, pesticides, fertilizers, cultivation methods, agronomic procedures, etc., are such that large plots are required. Note that when application is by airplane, the split block design may be the only realistic one to use. Suppose then that the levels of factor A are a pesticide sprays and the levels of factor B are cultivation methods with regular farm spraying and cultivating equipment. Further, suppose the plants are spaced x inches between rows and x inches between plants in a row. Hence, there are rows in two directions. Relatively large blocks, with one in each of r fields, are set up. The pesticide spray plots are laid out so that the length of the long narrow plot runs east and west. Then, the b cultivation methods are laid accross these plots so that the lengths of these plots run north and south. At measurement time the yield of an individual sbeu containing the combination A_iB_j is obtained.

A split block design has also been denoted as a two-way whole plot design and as a strip block design in statistical literature. This is the situation with several designs, that is they appear in statistical literature under a variety of aliases. Since this is true, the reader should always ascertain the exact design used, not the one purported to be used.

Both factors A and B in randomized complete block designs.

Consider the following variation and sampling structures. Suppose that there are N subpopulations each consisting of the variation structure described for a Latin square design in BU-550-M. The response model equation for the k^{th} element from the g^{th} subsubpopulation from subpopulation f is:

$$Y_{fghk} = \mu_{f...} + \rho_{fg} + \gamma_{fh} + \epsilon_{fghk} = \mu + \pi_f + \rho_{fg} + \gamma_{fh} + \epsilon_{fghk} \qquad (5.1)$$

A random sample of a columns and b rows in subpopulation f is made; a randomly selected sampling unit (sbeu) is obtained from subsubpopulation gh of subpopulation f. A random sample of r subpopulations is made with the same sampling in each subpopulation as described above. Then, in the sample from subpopulation f, the levels of factor A are randomly allotted to the columns and the levels of factor B to the rows. If the effects of adding or applying levels of factors A and B are additive we obtain the following model yield equation:

$$Y_{fijk} = \mu + \pi_f + \alpha_i + \gamma_{fi} + \beta_j + \rho_{fj} + \alpha\beta_{ij} + \epsilon_{gijk} \qquad (5.2)$$

Here it will be assumed that the γ_{fi} are $IID(0,\sigma^2_\gamma)$, the ρ_{fj} are $IID(0,\sigma^2_\pi)$, and the ϵ_{fijk} are $IID(0,\sigma^2_\epsilon)$.

An analysis of variance key-out of degrees of freedom for the above split block design would be:

Source of variation	Degrees of freedom	Mean square
Total	rab	
Correction for mean	1	
Blocks = R	(r-1)	
Levels of A = A	(a-1)	T_A
R × A = error (a)	(r-1)(a-1)	E_A
Levels of B = B	(b-1)	T_B
R × B = error (b)	(r-1)(b-1)	E_B
A × B	(a-1)(b-1)	T_{AXB}
R × A × B = error (ab)	(r-1)(a-1)(b-1)	E_{AXB}

Thus, there are three error terms in the above analysis of variance. The estimated variance of a difference between two means for level of factor A is $V(\bar{y}_{.i.}-\bar{y}_{.i'.}) = 2E_A/rb$, $i \neq i'$, and between two means for levels of factor B is $V(\bar{y}_{..j} -\bar{y}_{..j'}) = 2E_B/ra$, $j = j'$.

The levels of factor A have been randomized r separate times, once in each of r blocks. The same is true of levels of factor B. If the a × b factorial resulting in ab combinations has been allocated to experimental units in a randomized complete block design with v = ab and r blocks, <u>each</u> combination A_iB_j would have been involved in r randomizations; <u>however</u>, any level of A_i would have been randomly allotted rb times and any level of B_j ra times. Also, the difference in sizes and nature of wpeua and wpeub should be noted in split block designs.

One could conceivably consider another structure for population variation and sampling. Suppose that N subpopulations with means $\mu_{g..}$ comprise the total population with mean μ. Further, suppose that each sampling unit in subpopulation

g is of size b sbeu's by a sbeu's equal to ab sbeu's. A randomly selected sampling unit is obtained from subpopulation g. The levels of A are randomly applied in strips vertically and the levels of B are randomly applied in strips horizontally. This would produce the above split block design. Also, note that the expected value of $E_{A\times B}$ is σ^2_ϵ, the expected value of E_B is $\sigma^2_\epsilon + a\sigma^2_\rho$, and the expected value of E_A is $\sigma^2_\epsilon + B\sigma^2_\gamma$ under the above additive model equation, given the condition of equality of variances in subpopulations and in subsubpopulations. Note also that the above expectations indicate heterogeneity between vertical and horizontal strips of the sbeu considered in this paragraph. Otherwise, σ^2_ρ and σ^2_γ would be zero.

Factor A in a randomized complete block design and Factor B in a Latin square design.

Suppose that the following variation and sampling structures exist in the population of sbeu's prior to application of the combinations A_iB_j in the experiment. Let there be N subpopulations as defined for model yield equation (5.1). In addition, suppose that in rows, there is a layering of r = b subsubsubpopulations and that when we randomly select a columns and b rows in subpopulation f, the r = b layered subsubsubpopulations are also selected. The model yield equation, then, would be of the form:

$$Y_{feghk} = \mu + \pi_f + \delta_e + \rho_{fg} + \gamma_{fh} + \epsilon_{feghk} \qquad (5.3)$$

Upon application of A_iB_j under the additive effects situation, the above becomes:

$$Y_{feijk} = \mu + \pi_f + \delta_e + \gamma_{fi} + \alpha_i + \beta_j + \rho_{fej} + \alpha\beta_{ij} + \epsilon_{fegijk} \qquad (5.4)$$

We shall make the same assumptions about effects as in (5.2), but δ_e is considered to be a fixed effect.

A key-out of the degrees of freedom in an analysis of variance table is:

Source of variation	Degrees of freedom	Mean square
Total	ab^2	
Correction for mean	1	
Blocks = R	(b-1)	
Levels of A = A	(a-1)	T_A
R × A	(b-1)(a-1)	E_A
Levels of B	(b-1)	T_B
Rows for factor B design	(b-1)	
Remainder	(b-1)(b-2)	E_B
A × B	(b-1)(a-1)	$T_{A\times B}$
A × B × R	$(a-1)(b-1)^2$	$E_{A\times B}$

The estimated variances of a difference between A_i means or between B_j means is the same form as for the previous design.

6. Understanding and Recognizing the Split Plot and Split Block Designs

As is evident from the carpet example in section four, and from many examples given by Federer [1975c], it is not always easy to recognize the type of split plot or split block experiment design to associate with a particular experimental layout. We shall paraphrase some of the rules and procedures given in section 9 of the above cited paper in an attempt to aid the reader in understanding the nature of confounding and variation associated with split plot, split block and other similar types of experiment designs. First, however, let us consider a number of experimental situations in an attempt to acquaint the reader with what constitutes an appropriate error variance for a treatment contrast.

Example 1. Measuring one plant 100 times versus measuring each plant once for 100 randomly selected plants from a population of plants.

Suppose that we measure one plant 100 times for a characteristic Y. Then, suppose that the arithmetic mean, the variance, and a $(1-\alpha)\%$ confidence limit is computed. The confidence limit relates to the population mean of the one plant selected. Now in the second situation, we have a population of plants with mean μ. We obtain a simple random sample of 100 plants and compute the arithmetic mean, the variance for the 100 measurements, and a confidence interval on μ, the population mean. Note that the confidence interval is on the population mean rather than on the mean of one plant as in the previous case. Some experimenters make the mistake of believing that obtaining measurements as in the first situation allows them to construct confidence intervals on μ. Note that the variance in the first case contains only measurement error, whereas the variance in the second case contains a component due to measurement error and a component due to plant-to-plant variation.

Example 2. A single block of a randomized complete block design with k observations on each experimental unit.

The statistical analyses used in practice for one replicate of a randomized complete block design with k observations per experimental unit probably represent the most common mistakes made in experimentation. To illustrate what is done in practice and what should be done, let us consider the following example. An experimenter wishes to compare the amount of wet and dry weight produce after a six-week period from tomato plants (T), pigweed plants (P), and tomato plants and pigweed plants grown together (T+P). He plants one greenhouse flat of 8 tomato plants, one greenhouse flat of 8 pigweed plants, and one greenhouse flat of 4 tomato and 4 pigweed plants in alternate positions. The 3 greenhouse flats with 8 plants each were placed in a growth chamber at a specified light intensity. An incorrect statistical analysis for measurements from experiments of this nature is to consider that the above is a completely randomized design and to run a one-way analysis of variance as follows:

Source of variation	Degrees of freedom	Mean square
Total	24	
Correction for mean	1	
Among treatments	2	T
Within treatments	20	"E"

An F-test of T/"E" is used, and it is stated that hypotheses about equality of population treatment means is involved. But, suppose that one flat was filled with sand and the other two with highly fertile loam soil. This illustrates that flat to flat variation affects an estimated treatment mean. In this case, the effect of the particular flat and the effect of the treatment are completely

confounded. Hence, the "among treatments" mean square should be "among treatments plus flats" mean square. The within flat variation does not contain a between flats component. Also, the variation among plants within a flat may contain a component due to competition among plants within a flat. A correct form of an analysis of variance for this experiment would be:

Source of variation	Degrees of freedom	Mean square
Total	24	
Correction for mean	1	
Among blocks = R	0	
Among treatments = T	2	T
R × T = error (T)	0	E
Among plants within flats	21	

In this form it is obvious that no estimate of the error variance exists for comparing treatment means. The experimenter should use more than one block.

Example 3. A single replicate of a split plot design with k observations on the speu.

This situation is another frequently occurring situation in practice. The experiment in example 2 was part of an actual experiment conducted last year. The entire experiment was set up as follows:

Treatment or plant types	Light intensity (a_i) + growth chamber		
	a_1 + chamber 1	a_2 + chamber 2	a_3 + chamber 3
t_1 = tomato = T	8 plants	8 plants	8 plants
t_2 = pigweed = P	8 plants	8 plants	8 plants
t_3 = T + P	8 plants	8 plants	8 plants

The incorrect statistical analysis used was:

Source of variation	Degrees of freedom	Mean square
Total	72	
Correction for mean	1	
Light intensities = I	2	T_I
Plant types = T	2	T_T
T x I	4	T_{IxT}
"Error" = plants within flats	63	"E"

where the arrows indicate the F tests made. A correct partition and F tests would be:

Source of variation	Degrees of freedom	Mean square
Total	72	
Correction for mean	1	
Blocks = R	0	
Light intensities = I	2	T_I
R x I = error (I)	0	E_I
Plant types = T	2	T_T
T x I	4	T_{TxI}
T x R : I	0	E_T
Plants within flats	63	

Hence, no error terms exist for comparing treatment mean squares.

Another highly probable mistake made in both this and the previous example is that in pooling the plants within flat variances, it has been assumed that they are estimates of the same parameter. This is highly unlikely for plants

as divergent as pigweed and tomato.

Example 4. Finding an "error variance" for one replicate of a split plot design.

The actual situation for the carpet example in section four was that only one sample of each brand of rug was obtained by the experimenter. There were a total of 30 observations made for the 6 brands and 5 lengths of wear. Obviously, one could construct an analysis of variance for this single replicate of a split plot design as described in example 3. The question arises as to whether or not it is possible to construct an error variance from some of the treatment contrasts. Let us suppose that we know that brands 1 and 5 are quite similar and that brands 2, 3, and 6 are quite similar. This means that the variation within these two groups might be considered as mostly rug sample to rug sample variation. If this were the situation, then an analysis of the following form would be appropriate:

Source of variation	Degrees of freedom	Mean square
Total	30	
Correction for mean	1	
Brands	5	
B_1 = Brand 1+5 vs 2+3+6	1	
B_2 = Brand 4 vs rest	1	
B_3 = Brand 1 vs 5	1	E_B
B_4 = Among brands 2, 3, and 6	2	
Split-plot analysis		
Length of wear = L	4	
L x B_1	4	
L x B_2	4	
L x B_3	4	E_L
L x B_4	8 (12)	

In another situation let us suppose that brands are different with respect to static electricity and we can do no grouping, as above. Instead, let us suppose that amount of static electricity is a quadratic function of length of wear. Then, any interaction of higher degree polynomial terms with brands could possibly be considered as an error term for linear and quadratic components effects and interactions of length of wear. An analysis of variance as follows could be appropriate:

Source of variation	Degrees of freedom		Mean square
Total	30		
Correction for mean	1		
Brands = B	5		
Error for brands	0		
Length of wear = L			
Linear effect of length		1	
Quadratic effect of length		1	
Remainder (lack of fit)		2	
L × B	20		
Linear × brands		5	
Quadratic × brands		5	
Remainder × brands		10	E_L

The above four examples were designed to lead the reader into thinking about what constitutes an appropriate error variance for a specified treatment contrast. From these and other statements in this paper one is led to the following considerations:

I. It is essential to determine precisely the experimental unit for <u>each</u> treatment type (factor) and to define the sampling structure of the experimental units for each of the factors singly and for the factors jointly.

II. It is essential to determine precisely what the experimental procedure actually is rather than what it is purported to be.

III. It is essential to count the number of randomizations for each factor singly, for pairs of factors, for triplets of factors, etc. If the number of randomizations differs, one can expect different sizes of experimental units and some types of confounding.

IV. It is essential to determine if levels of factors are crossed or nested within levels of other factors.

V. An error variance for a treatment contrast of levels of a factor is determined from the variance among experimental units treated alike (i.e., same treatment) where the experimental units are those for the factor involved. Note that there may be several error variances in a given experiment.

VI. It is essential to have a complete and meaningful key-out of the degrees of freedom in an analysis of variance table prior to <u>any</u> statistical computations.

VII. It is desirable to recheck all assumptions involved in a statistical analysis, such as independence of observations, additivity of effects, equality of components of a mean square, and normality of residuals, to ascertain that they are satisfied for the material used in the experiment under consideration.

VIII. After computing means and effects, it is highly desirable to compute and to study residuals prior to the computation of sums of squares.

IX. It is very useful to prepare graphical displays of the data, the means and the residuals.

X. When fully satisfied with the preceding nine considerations, compute the sums of squares and mean squares and complete the statistical analyses.

Three procedures (algorithms I, II, and III) are discussed by Federer [1975c]. These have been helpful procedures to follow when analyzing complex data. One of the most complex designs analyzed by the author involved 203 lines in analysis of variance excluding the total sum of squares and the correction for the mean (see Federer and Farden [1955]). With the partitioning of the degrees of freedom for levels of factors into single degree of freedom contrasts several hundred lines in the analysis of variance would result. Without such considerations as the ten above, it is doubtful if one could have obtained a key-out of degrees of freedom.

7. Literature Cited

Federer, W. T. [1975a]. Design and blocking considerations in empirical investigations. BU-548-M in the Mimeo Series of the Biometrics Unit, Cornell University, February.

Federer, W. T. [1975b]. Blocking and design for one-way and two-way elimination or control of heterogeneity in the experimental material. BU-550-M in the Mimeo Series of the Biometrics Unit, Cornell University, March.

Federer, W. T. [1975c]. The misunderstood split plot. Applied Statistics (edited by R. P. Gupta), North Holland Publishing Company, Amsterdam, Oxford and American Elsevier Publishing Company, Inc., New York, pages 9 to 39.

Federer, W. T. and Farden, C. A. [1955]. Analysis of variance set-up for Joint Project 69. BU-67-M in the Mimeo Series of the Biometrics Unit, Cornell University.

Warriner, D. A. [1951]. An investigation of the effect of certain psychosemantic factors on the level of reading comprehension difficulty in high school chemistry and physics. Ph.D. Thesis, Cornell University.