UNIVERSITY OF NORTH CAROLINA
Department of Statistics
Chapel Hill, N. C.

DESIGN AND ANALYSIS OF EXPERIMENTS

Notes on lectures held during the fall semester (1963) at the Statistics Department, University of North Carolina, Chapel Hill, N.C.

by

Arthur Linder

Visiting Professor from the University of Geneva (Switzerland)

August 1964

This research was supported by the Mathematics Division of the Air Force Office of Scientific Research Grant No. AF-AFOSR-84-63.

Institute of Statistics
Mimeo Series No. 398

TABLE OF CONTENTS

1. INTRODUCTORY REMARKS.

1.1 Scope of statistical methods.

In order to see the subject of this course in proper perspective, let us examine the essential steps in any statistical investigation.

(a) The first step in a statistical investigation consists in the collection of data. This can be done in two ways, namely (1) by sample survey or (2) by conducting experiments. A complete census would be considered as the extreme limit of a sample survey.

Examples of (1) are statistics of prices, of salaries. Examples of (2) are field experiments, biological assays, studies of the yield of a chemical process as a function of temperature and duration of the process.

(b) The second step in a statistical analysis is the reduction and presentation of data. Averages and measures of dispersion or other suitable statistics are obtained which describe the essential features of the data. Representing data in the form of tables, diagrams and graphs helps to detect the essential features of statistical data and to communicate them to others.

(c) The third and last step in statistical analysis is a critical analysis. This may involve the estimation of parameters, setting of fiducial or confidence limits, testing of hypothesis, deciding on one of several courses of action. In this third step use is made of the calculus of probability. But statistical procedures cannot be fully understood or explained by probability theory alone; an essential ingredient of statistics is inductive reasoning which is part of logic or of philosophy of knowledge.

1.2 What is an experiment?

The characteristics of an experiment may be studied in considering a particular example. The following medical inquiry was conducted by a surgeon in Switzerland during the period 1929 to 1945. The object was to compare various treatments after operation in their possible effects in reducing the

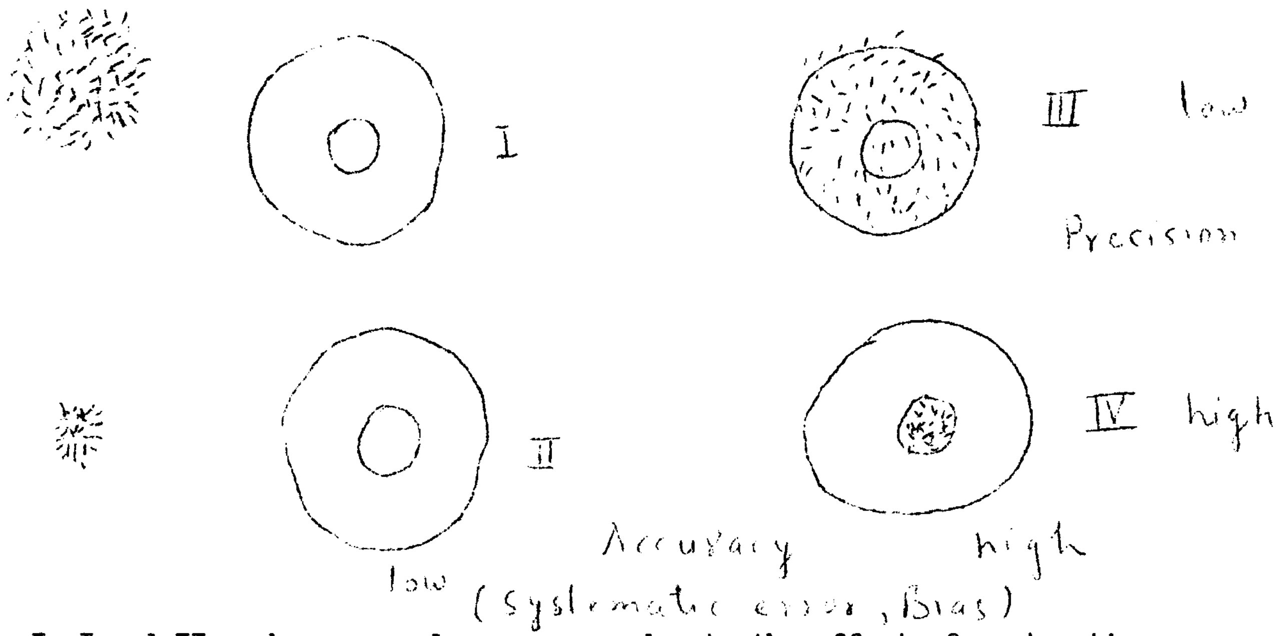
the death rates due to embolism. The results obtained are summarised in the following table.

Years	Treatment after operation	Operations performed	Postoperative deaths from embolism number	o/oo
1929-33	No specific treatment	4311	14	3.25
1934-38	Rest in bed, exercises, sympatol	4669	15	3.21
1941-45	Get up early	5041	1	0.19

The figures indicate a sharp decrease in the number of deaths from embolism coinciding with the introduction of the policy of getting up soon or immediately after operation. It may be concluded that this changed policy was the <u>cause</u> of the reduction of the mortality due to embolism. This interpretation of the results may be correct, but it must be remembered that many other causes may be responsible for the drop in the number of deaths. It may well be that the age distribution of the patients has changed, or the distribution according to social status, to severity of operation. Furthermore, there may have been changes in the kind of food given to the patients which could affect the death rate. For an experiment to be valid, these so-called extraneous causes should not affect the results. While the data summarised above could be considered as a basis for inferring the usefulness of getting up early after operations in preventing deaths from embolism, a statistician would generally warn to adopt such evidence without further study. He will remember the old advice that correlation and causation should not be confounded. In order to establish the relationship between cause and effect, it will be necessary to eliminate as far as possible the influence of any extraneous factor.

1.3 Aims of an experiment.

The object of an experiment is to give an answer to specific questions as accurately and as precisely as possible with a limited amount of time, money and resources in general. To illustrate what we mean by an accurate and precise experiment let us consider the distribution of the points of impact when shooting at a given target. The following figure shows four possible outcomes of shootings.



In I and II we have very low accuracy due to the effect of systematic errors; these errors act the same way in every shoot and it should be possible to detect the source of such systematic errors and to remove it. In I and III the precision is low due to substantial accidental or random errors. In IV both types of errors - systematic and random - are very small; this experiment could be considered as accurate and precise.

1.4 <u>Principles of experimental design</u>.

Consider the problem of conducting an experiment on a group of animals for the comparison of the effects of four different vitamins A, B, C and D.

First of all we should decide how to measure the effect of these vitamins, i.e. whether we should measure weights, or lengths, etc., whether it is sufficient to use only one character or more characters should be included. These points should be settled before the experiment is started.

In what follows we shall confine ourselves to the measurement of one character only. Supposing we have decided to rely on the weight of the animal, we have furthermore to see whether we will use the final weight, the weight increase per day or the weight increase per unit quantity of food.

Secondly it may be useful and often necessary to compare the animals receiving vitamins to a group receiving only the usual food without any vitamin added. This control group is of importance in clinical experiments. It helps to ascertain that an effect is more important than mere psychological effects.

Let us now consider the simplest experiment in which there are only two treatments O and A and these are applied to animals (e.g. mice or rats) which are, say, 8 weeks old. The effect of treatments may be measured by the increase in weight during an experimental period of 4 weeks.

Let us imagine that we use only <u>one</u> animal in each treatment.

Treatment	A	O
Animal	1	2

The results of this experiment can be plotted on a straight line:

A

O

Increase in weight

Looking at the results in the diagram we could conclude that treatment A is superior to treatment O. This is a rather over-hasty conclusion open to doubts. In fact, with only these results, we do not know whether a difference as great as that recorded between O and A could not arise if we would apply the same treatment (either O or A) to the animals 1 and 2 . We have to know something about the variability between animals treated alike in order to assess the effects of different treatments. And we should know it under the same conditions as those which prevail when we compare A to O. The inevitable conclusion is, that we should have several animals in each treatment A and O, which will give us the possibility of determining the experimental error.

<u>Principle I. Replication. In order to evaluate the experimental error, we have to apply each treatment to several experimental units.</u>

The main reason why it is necessary to replicate, lies in the fact that we cannot get a valid measure of the experimental error without replication. Replication will of course also give higher precision for comparing different treatments, but this is a subsidiary effect of replication,not the essential reason.

Let us then assume that an experiment is carried out, using e.g. 10 animals under each treatment.

Treatment	Animals									
A	1	2	3	4	5	6	7	8	9	10
O	11	12	13	14	15	16	17	18	19	20

If we plot the results, and if we get something like the following graph,

A

———————————————— Increase in weight

O

could we then be sure that A has caused a higher increase than O ? By no means. It may well be that the difference in weight increase between the animals in group A and those in group O is not due to the action of A, but to the effect of some other cause. It may be, to name only one possible source of bias, that in assigning the animals to the treatments A and O, the animals under A had a higher initial weight. This alone might explain the higher increase in weight of the animals in group A.

Any careful experimenter would avoid systematic errors of the kind just mentioned, but it is often difficult, especially in medical experiments, to avoid such biases. In some cases the animals (or the experimental units in general) may differ from each other in some characters which are hidden, so that it is impossible to take them into account, as we could for the initial weight.

The problem before us, is then to subdivide a number of experimental units into two or more treatment groups in such a way as to avoid any possible bias through some unknown characters of these units. The only solution to this problem is to assign the experimental units strictly at random to the different groups. The random allocation of experimental units to treatments will not avoid any bias caused by differences between experimental units, but it will be distributed at random to the different treatments. It will, in fact become part of the random errors. We may state principle II as follows:

Principle II (Randomization). The experimental units should be assigned to treatments strictly at random, so as to avoid bias due to differences between experimental units.

If we wish to assign units strictly at random, we have to use such mechanisms as have been used in games of chance, that is, shuffling of cards, throwing a die, etc. Such mechanisms may be avoided through the use of tables of random numbers.

Units should not be assigned in a haphazard way. If, for instance, two tranquilizers had to be compared on rats, and if 20 rats should be split up in two groups, one could proceed as follows. If the 20 rats were all in one cage, he could pick up one after another until he had 10, which he would assign to the first tranquilizer, while those remaining in the box would be assigned to the second tranquilizer. There is no doubt that by doing so the first tranquilizer would be applied to lazy, phlegmatic rats, while the second would be given to quicker rats of a more nervous disposition. The comparison of the two tranquilizers would most certainly be heavily biased through this faulty assignment of rats to the treatments.

The effect of randomization can be studied in different ways. A very simple procedure is to consider the effect of random assignment of 20 animals (experimental units in general) to two treatments. Assuming that 10 animals are such that they have a bias in weight increase of -1, and the other 10 a biase of +1. The first 10 animals may form a group I, the 10 others group II. The probability p(x) of assigning at random x animals from group I to treatment O (and 10 - x to treatment A) and consequently 10 - x animals of group II to treatment O (and x to treatment A) is given by

$$p(x) = \binom{10}{x}\binom{10}{10-x}\Big/\binom{20}{10}$$

and the probabilities for x = 1, 2, ..., 10, as well as the average bias are given in the following table.

Number of animals from group I under treatment O	Probability of obtaining x animals from group I under treatment O by random allocation	Bias of difference between averages for O and A
x	p(x)	(20 - 4x)/10
0	1/184,756	+ 2.0
1	100/k84,756	+ 1.6
2	2,025/k84,756	+ 1.2
3	14,400/184,756	+ 0.8
4	44,100/184,756	+ 0.4
5	63,504/184,756	+ 0.0
6	44,100/184,756	- 0.4
7	14,400/184,756	- 0.8
8	2,025/184,756	- 1.2
9	100/184,786	- 1.6
10	1/184,756	- 2.0

As will be seen from these figures, those cases where the bias is small or non-existant will occur most often. Cases with extreme bias will occur with s small probability only. Randomization may be considered as a safeguard against bias, or to put it in another way, to get experimental results which are as accurate as possible.

We mentioned earlier that one objective of design was to obtain an answer to the experimental question which is as precise as possible. To make the experiment precise we have to reduce the magnitude of random errors. This can be done in many different ways.

In our example the random errors arise because of small differences between animals. The animals may differ in age, initial weight, sex, and other constitutional characters. There are also differences in the experimental conditions under which the animals are kept: they may get different amounts of food, the quality of the food may differ. There will furthermore be differences in measuring the weight of the animals.

The experimenter can reduce the errors flowing from all these sources if he carries out the experiment with the utmost care. In some cases experimenters go very far in the way of reducing experimental errors. They may for instance through a selection program obtain a high degree of inbreeding in the animals used for experimentation. This homogenizing of experimental conditions has however an economic limit. Very soon the cost involved in the standardization of experimental conditions becomes prohibitive.

Even if there were not an economic limit for this tendency in obtaining for all experimental units similar conditions, this standardization should not be carried too far, because the results obtained will be only valid for the very particular experimental conditions chosen. The inductive basis of the experiment will be very narrow, we will not be able to make any inferences for experimental conditions not exactly alike to those of our experiment.

How could we achieve high precision in an experiment and at the same time maintain a sufficiently broad basis for inductive inference? To attain this objective R. A. Fisher (1935) introduced the device known as "local control" or "blocking".

For our example we would apply this device in the following way. We would choose, say, 10 groups of 2 animals so that the two animals in each group would come from the same litter, and be of the same sex. The two litter-mates would form a "block",

Litter (block)	Treatment 0	A
I	1	2
II	3	4
III	5	6
IV	7	8
V	9	10
VI	11	12
VII	13	14
VIII	15	16
IX	17	18
X	19	20

and the treatment should be assigned to them at random. The principle of "local control" may be stated like this:

Principle III (Local Control). Experimental units as similar as possible are grouped together to form blocks. Comparisons within blocks will therefore be of high precision while the differences which exist between blocks ensure that the inductive basis is not unduly narrow.

We are now in a position to give a design for the experiment originally considered, namely one with treatments A, B, C, D and a control group O. We would take five litter-mates of the same sex - which for mice or rats is not too difficult to achieve - and assuming that we had animals from six litters this would give the following plan

Litters (Blocks)	Treatments				
	O	A	B	C	D
I	1	2	3	4	5
II	6	7	8	9	10
III	11	12	13	14	15
IV	16	17	18	19	20
V	21	22	23	24	25
VI	26	27	28	29	30

The comparisons between the treatments would be made within litters, thus being of relatively high precision. Random assignment of treatments to animals within litters (blocks) would avoid systematic errors.

In the plan above we have used implicitely a principle whichI think should be added to principles I, II and III as given by R. A. Fisher (1935), namely a principle of symmetry. If we adhere to a certain symmetry in the plan of the experiment, we simplify at the same time the analysis of the results. Also some questions cannot be answered if the plan lacks symmetry; the investigation of interactions is a case in point.

Principle IV (Symmetry). In view of simplifying the analysis the plan of the experiment should be symmetrical. Symmetry also makes the experiment more comprehensive in most cases.

An experiment carried out by Baxter and Lundstedt (1940) - see also Youden (1950) - may serve to illustrate the drawbacks in a design lacking symmetry. The purpose of the experiment was to determine the atomic weight of iodine. Two iodine samples and five silver preparations were used in the experiment which gave the following results:

Iodine sample	Silver preparation	Ratio of reaction weights
I	A	1.176 422
	A	1.176 425
	B	1.176 441
	B	1.176 441
	C	1.176 429
	C	1.176 420
	C	1.176 437
	D	1.176 449
	D	1.176 450
	E	1.176 455

Iodine sample	Silver preparations	Ratio of reaction weights
II	A	1.176 399
	A	1.176 440
	A	1.176 418
	B	1.176 423
	B	1.176 413
	D	1.176 461

It may be that the ratio of reaction weights varies with the iodine sample and the silver preparations. This can better be seen when we transform the ratios of reaction weights according to the formula

$$z = 1{,}000{,}000 \left[(\text{ratio of reaction weights}) - 1.176\,400 \right]$$

and arrange the z-values in a two-way table.

Iodine sample	Silver preparation A	B	C	D	E	$T_{.k}$	$N_{.k}$
I	22	41	29	49	55		
	25	41	20	50		369	10
			37				
II	-1	23		61			
	40	13				154	6
	18						
$T_{j.}$	104	118	86	160	55	523	...
$N_{j.}$	5	4	3	3	1	...	16

Without going into any formal analysis, it seems that there would be differences between iodine samples as well as between silver preparations. However, it is not easy to make comparisons because the number of determinations is not the same for each silver preparation with each iodine sample. The silver preparations D and Eseem to give higher values. This will affect the averages of the iodine samples in such a way that I will be higher than II, and it will be very difficult to disentangle the effects of the iodine samples from those of the silver preparations.

If on the contrary, the experiment had been performed according to the following plan, it would be much easier to find out the effects of iodine samples and silver preparations.

Iodine sample	Silver preparation A	B	C	D	E
I	1	3	5	7	9
	2	4	6	8	10
II	11	13	15	17	19
	12	14	16	18	20

In this design the averages of all the silver preparations would be equally influenced by differences between iodine samples; and the averages for the two iodine samples would be equally influenced by differences between silver preparations. Furthermore, it would be easy to ascertain whether the iodine samples acted differently on different silver preparations, that is we could determine any possible interaction between iodine samples and silver preparations. The symmetry of design would make the experiment more comprehensive.

In summary, to get an experiment which is accurate, precise and comprehensive, principles I (replication), II (randomization), III (local control) and IV (symmetry) should be applied as far as possible.

The following chapters will show how these principles are used in different experimental situations. Throughout illustrations from actual research work will be given to emphasize the problems of design and to show how these experiments ought to be analysed.

2. SOME SIMPLE EXPERIMENTAL DESIGNS.

In this chapter we shall describe the basic designs which are obtained when we use the principles developed in chapter 1. As already mentioned, we get a valid design by applying the principles of replication and randomization; the design obtained is called a completely randomized design. If we use also the principles of local control and of symmetry, we obtain designs known under the names of randomized blocks designs and latin squares. These designs are described in the present chapter and their analysis will be given.

2.1 Completely randomized design.

These are the simplest designs where use is made of symmetry but where local control is ignored. If the experimental material is quite homogeneous, these designs work very well. But it is usually worthwhile to try to increase precision through local control.

Example. An experiment was conducted to compare the effect of 7 food-mixtures A, B, C, ..., G. Out of 42 rats six were allocated at random to each treatment (food-mixture). The weights of the rats after an experimental period of 56 days are given below, in grams. See Linder (1953), p. 18.

	Food-mixture (treatment)							
	A	B	C	D	E	F	G	
	119	123	130	144	159	134	156	
	90	121	163	172	172	146	183	
	102	159	159	165	210	161	146	
	85	138	140	143	171	149	169	
	113	173	121	179	232	124	147	
	136	138	142	146	190	137	...	
Total	645	857	855	949	1134	856	801	6097
Mean	107.5	142.8	142.5	158.2	189.0	142.7	160.2	148.7

One animal (in group G) died during the experimental period.

The analysis of this completely randomized experiment rests on several assumptions. One specific treatment is assumed to act alike for all animals assigned to it. The experimental errors occuring in the course of the experiment are assumed to act independently of the treatment effects. Those effects which depend on differences between the individual animals are also randomly distributed and independent of treatment effects because the animals were assigned at random to the seven treatments. The random assignment is either done by drawing cards, throwing a die or using a table of random numbers.

The assumptions mentioned may be formally set out in the relation

$$y_{ji} = \alpha + \tau_j + \delta_{ji}$$

where y_{ji} is the weight of the i-th animal under the j-th treatment, α the general average, τ_j the effect of treatment j and δ_{ji} a random effect for the i-th animal in the j-th treatment-group. For the validity of tests of certain hypothesis, and for the construction of fiducial or confidence limits, we will have to suppose further that the δ_{ji} are normally distributed with a common variance σ^2 and a mean equal to zero. The model given implies that treatment effects and random effects are additive.

The statistical analysis is simple and straightforward. First we have to get estimates of the treatment effects and of the standard error σ of individual values. To perform the analysis in general terms we use the following notations:

t = number of treatments,

N_j = number of units under treatment j,

N = total number of units,

T_j = sum of values y_{ji} for units under treatment j,

T = total sum of values for all units,

$\bar{y}_i$ = mean of y_{ji} under treatment j,

$\bar{y}$ = general mean of all y_{ji} .

We have

$$N = \underset{j}{S} N_j \quad ; \qquad T_j = \underset{i}{S} y_{ji} \quad ; \qquad T = \underset{j}{S} T_j = \underset{j}{S}\underset{i}{S} y_{ji}$$

$$\bar{y}_j = T_j/N_j \quad ; \qquad \bar{y} = T/N \quad .$$

We use the principle of least squares to estimate α and the τ_j , minimizing

$$U = \underset{j}{S}\underset{i}{S} (y_{ji} - \alpha - \tau_j)^2 \quad .$$

For $\partial U/\partial\alpha = 0$ we have

$$N\,a + N_1 t_1 + N_2 t_2 + \ldots + N_t t_t = T$$

where a is the estimate of α and t_j the estimate of τ_j $(j = 1, 2, \ldots, t)$.

For $\partial U/\partial\tau_j$ we get

$$N_j a + N_j t_j = T_j \quad , \qquad (j = 1, 2, \ldots, t).$$

If we assume that

$$N_1 t_1 + N_2 t_2 + \ldots + N_t t_t = 0 \quad ,$$

we obtain the simple estimates

$$a = \bar{y} \quad ; \qquad t_j = \bar{y}_j - \bar{y} \qquad (j = 1, 2, \ldots, t)$$

To estimate σ^2, we have to form

$$\underset{j}{S}\,\underset{i}{S}\, d_{ji}^2 = \underset{j}{S}\,\underset{i}{S}\,(y_{ji} - a - t_j)^2$$

$$= \underset{j}{S}\,\underset{i}{S}\,(y_{ji} - \bar{y}_i)^2$$

which is the sum of squares within treatments. It can be easily calculated by means of an analysis of variance, whereby the total sum of squares is divided in two parts: the sum of squares (SS) between treatments and the SS within treatments.

By definition we have

$$\text{Total SS} = \underset{j}{S}\,\underset{i}{S}\,(y_{ji} - \bar{y})^2$$

$$\text{SS between treatments} = \underset{j}{S}\, N_j(\bar{y}_j - \bar{y})^2$$

$$\text{SS within treatments} = \underset{j}{S}\,\underset{i}{S}\,(y_{ji} - \bar{y}_j)^2$$

and the relation holds

$$\underset{j}{S}\,\underset{i}{S}\,(y_{ji} - \bar{y})^2 = \underset{j}{S}\, N_j(\bar{y}_j - \bar{y})^2 + \underset{j}{S}\,\underset{i}{S}\,(y_{ji} - \bar{y}_j)^2 \quad ,$$

with a corresponding relation for degrees of freedom (DF):

$$N - 1 = (t - 1) + (N - t) \quad .$$

The analysis of variance may be written as follows:

Variance	Degrees of freedom (DF)	Sum of squares (SS)	Mean Square (MS)
Between treatments	$t - 1$	$\underset{j}{S}\, N_j(\bar{y}_j - \bar{y})^2$	D
Error	$N - t$	$\underset{j}{S}\,\underset{i}{S}\,(y_{ji} - \bar{y}_j)^2$	s^2
Total	$N - 1$	$\underset{j}{S}\,\underset{i}{S}(y_{ji} - \bar{y})^2$	...

with

$$D = \underset{j}{S} N_j(\bar{y}_j - \bar{y})^2/(t-1)$$

$$s^2 = \underset{j}{S}\,\underset{i}{S}(y_{ji} - \bar{y}_j)^2/(N-t) \quad .$$

The variance s^2 within the tredtments is an estimate of the error variance σ^2 . The standard error of the estimate $a + t_j$ is $s/\sqrt{N_j}$.

To test the hypothesis

$$\tau_1 = \tau_2 = \dots = \tau_t = 0$$

that is the hypothesis of no difference between the t treatments, we compare the calculated

$$F = D/s^2$$

with the tabulated F_p, for the given critical probability P on $n_1 = t - 1$ and $n_2 = N - t$ degrees of freedom. If $F > F_p$ we reject the hypothesis.

To test the hypothesis

$$\tau_1 = \tau_2 \quad ,$$

we calculate

$$t = \frac{(\bar{y}_1 - \bar{y}_2)}{s}\sqrt{\frac{N_1N_2}{N_1+N_2}}$$

and compare it with the tablulated value t_p, corresponding to a critical probability P and $n = N - t$ degrees of freedom. If $t > t_p$ we reject the hypothesis.

To summarize, the essential steps in the analysis of a completely randomized experiment are:

(1) Set out the results of the experiment in a table, giving also the totals T_j, T and the means $\bar{y}_j$.

(2) To fill in the analysis of variance table we calculate succèssively

$$\text{Total SS} = \underset{j}{S}\,\underset{i}{S}\, y_{ji}^2 - T^2/N = (N\,\underset{j}{S}\,\underset{i}{S}\, y_{ji}^2 - T^2)/N$$

$$\text{SS between treatments} = \underset{j}{S}(T_j^2/N_j) - T^2/N$$

$$\text{SS error} = \text{SS within treatments} = \text{Total SS} - \text{SS between treatments} \quad .$$

The analysis of variance table is then completed by calculating D and s^2.

Tests of hypothesis can be performed.

For the example of p. 14 we have the following analysis of variance:

Variance	DF	SS	MS	F
Between treatments	6	21,783.7	$D = 3,630.6$	9.97
Error	34	12,376.8	$s^2 = 364.0$	...
Total	40	34,160.5	...	...

With $P = 0.01$ and $n_1 = 6$, $n_2 = 34$ the F-value is between 3.29 and 3.47 so we conclude that there are differences in the effect of the treatments.

In testing differences between two treatments care must be taken not to select for instance the highest and lowest mean. The problem of multiple comparisons has to be borne in mind, though the methods developed to deal with it are still of a preliminary nature. They will not be discussed here.

2.2 Randomized block designs.

In the randomized block designs all the four principles developed in 1.4 are used. Specifically, in applying the principle of local control we form groups of experimental units which are (or at least seem to be) more homogeneous and apply the treatments at random to the units within each group, using a new randomization process within each group. The principle of symmetry calls for all groups to contain as many units as there are treatments.

Example. A variety trial comparing 5 varieties of wheat was carried out in 4 blocks of 5 plots. The five varieties Huron (H), Pilot (P), Regent (R), New-batch (N) and Coronation (C) were assigned at random to each plot within each block. See Linder (1953, p. 34).

Plan and yield of grain in 10 kg. per ha.

Block I

R	N	P	C	H
446	409	440	421	464

Block II

N	H	P	C	R
376	441	393	402	334

Block III

H	C	R	N	P
407	410	321	309	320

Block IV

C	N	H	P	R
327	296	376	351	343

The analysis of experiments arranged in randomized blocks is based on the following dditive model:

$$y_{jk} = \alpha + \tau_j + \beta_k + \delta_{jk}$$

where

y_{jk} = value for treatment j in block k,

α = general mean,

τ_j = effect of treatment j,

β_k = effect of block k,

δ_{jk} = random error for treatment j in block k.

It will be assumed that the δ_{jk} have a common variance σ^2 and an expected mean equal to zero. The following general symbols will be used:

t = number of treatments

$b = r$ = number of blocks or replications,

$N = rt$ = total number of experimental units,

$T_{j.}$, $\bar{y}_{j.}$ total and mean for block k,

T, $\bar{y}$ general total and mean.

The following relations may be noted:

$$\underset{k}{S}\, y_{jk} = T_{j.} \quad , \quad T_{j.}/r = \bar{y}_{j.}$$

$$\underset{j}{S}\, y_{jk} = T_{.k} \quad , \quad T_{.k}/t = \bar{t}_{.k}$$

$$\underset{j}{S}\,\underset{k}{S}\, y_{jk} = T \quad , \quad T/N = \bar{y}$$

To estimate the parameters α, τ_j and β_k in the model we apply the principle of least squares and demand

$$U = \underset{j}{S}\,\underset{k}{S}(y_{jk} - \alpha - \tau_j - \beta_k)^2$$

to be a minimum over the variation of α, τ_j and β_k . The estimates a, t_j and b_k are given by

$$a = \bar{y}$$

$$t_j = \bar{y}_{j.} - \bar{y}$$

$$b_k = \bar{y}_{.k} - \bar{y}$$

respectively. For d_{jk} we get

$$\begin{aligned} d_{jk} &= y_{jk} - a - t_j - b_k \\ &= (y_{jk} - \bar{y}_{j.}) - (\bar{y}_{.k} - \bar{y}) \\ &= (y_{jk} - \bar{y}_{.k}) - (\bar{y}_{j.} - \bar{y}) \end{aligned}$$

and for the interpretation of these quantities we refer to the following scheme:

Blocks	Treatments 1 2 ... j ... t	Mean
1 2 ... k .... r	y_{jk}	$\bar{y}_{.k}$
Mean	$\bar{y}_{j.}$	$\bar{y}$

All d_{jk} will be zero when the differences between treatment effects are the same in all blocks, or alternatively if the differences between block effects are the same for all treatments.

The total SS can be split up in three parts: SS between treatments, SS between blocks and SS for error. The following table shows the analysis of variance and the definitions for the different SS.

Variance	DF	SS	MS
Treatments	$t - 1$	$r \underset{j}{S}(\bar{y}_{j.} - \bar{y})^2$	D_T
Blocks	$r - 1$	$t \underset{k}{S}(\bar{y}_{.k} - \bar{y})^2$	D_B
Error	$(t - 1)(r - 1)$	$\underset{j}{S}\,\underset{k}{S}(y_{jk} - \bar{y}_{j.} - \bar{y}_{.k} + \bar{y})^2$	s^2
Total	$tr - 1$	$\underset{j}{S}\,\underset{k}{S}(y_{jk} - \bar{y})^2$	...

The standard error of the mean $a + t_j = \bar{y}_j$ for the j-th treatment is $s/\sqrt{r}$, and it will be seen that it is the same for any treatment.

To test the hypothesis that

$$\tau_1 = \tau_2 = \dots = \tau_t = 0$$

we calculate

$$F = D_T/s^2$$

and compare this to F_p with $n_1 = t - 1$, $n_2 = (t - 1)(r - 1)$. If $F > F_p$ we reject the hypothesis that all treatments have the same effect.

To test the hypothesis

$$\tau_1 = \tau_2$$

we calculate

$$t = \frac{\bar{y}_{1.} - \bar{y}_{2.}}{s} \sqrt{\frac{r}{2}}$$

and compare this to t_p with $n = (t - 1)(r - 1)$ degrees of freedom. We reject the hypothesis if $t > t_p$.

The steps in the analysis can be summarized, and illustrated by our example, as follows:

(1) Prepare a table of the results, showing the totals for treatments and blocks and, if necessary, the means for treatments.

Blocks	Yield in 10 kg					
	H	P	R	N	C	Total
I	464	440	446	409	421	2180
II	441	393	334	376	402	1946
III	407	320	321	309	410	1767
IV	376	351	343	296	327	1693
Total	1688	1504	1444	1390	1560	7586
Mean	422.0	376.0	361.0	347.5	390.0	379.3

(2) For the analysis of variance we calculate

$$A = \text{total SS} = (N \, S_j S_k \, y_{jk}^2 - T^2)/N = 49{,}872.2 \quad ,$$

$$B = \text{treatment SS} = (t \, S_j \, T_{j.}^2 - T^2)/N = 13{,}179.2 \quad ,$$

$$C = \text{block SS} = (r \, S_k \, T_{.k}^2 - T^2)/N = 28{,}201.0 \quad ,$$

$$\text{Error SS} = A - B - C = 8{,}492.0 \quad ,$$

thus obtaining the analysis of variance table:

Variance	DF	SS	MS	P
Treatments	4	13,179.2	$D_T = 3,294.8$	4.656
Blocks	3	28,201.0	$D_B = 9,400.3$	...
Error	12	8,492.0	$s^2 = 707.7$	...
Total	19	49,872.2	...	...

(3) To test the hypothesis

$$\tau_1 = \tau_2 = \ldots = \tau_t = 0$$

we calculate

$$F = D_T/s^2 = 4.656$$

and compare this with $F_{0.05} = 3.26$ with $n_1 = 4$, $n_2 = 12$. This indicates that the five varieties have given yields differing significantly.

To test the hypothesis

$$\tau_1 = \tau_2$$

that is whether there is a significant difference between the yields for the varieties H and P, we have to calculate

$$t = \frac{\bar{y}_{1.} - \bar{y}_{2.}}{s} \sqrt{\frac{r}{2}} = 2.446 \quad .$$

The tabulated value of t at P = 0.05 with n = 12 is found to be $t_{0.05} = 2.179$, so that we may consider the difference in yield for H and P as significant.

(4) In many cases it will be useful to calculate the residual d_{jk}; this should be done whenever a computer is used. Note that the sums of the residuals

$$d_{jk} = y_{jk} - \bar{y}_{j.} - \bar{y}_{.k} + \bar{y}$$

both with respect to j and k are all zero. If there is a distinct pattern in the residuals this indicates that the model is not fully appropriate. The d_{jk} should be distributed at random.

2.3 Latin square design.

In the randomized block designs we used similarity of experimental units to form the blocks. Often similarity is present in units contiguous in space or time. It has been found in field trials that the principle of local control can be applied in two directions, in other words, we form homogeneous blocks corresponding to two criteria of grouping.

If for instance, a field is situated on a slope, it may well be that the upper portion is dryer than the lower. There may also on the same field exist a trend in fertility in a direction across the field, due perhaps to a change in soil composition. If on such a field we had an area on which to accomodate 24 plots and we had to compare 6 varieties of wheat, we could form 4 blocks as explained in 2.2.

If, however, the area were sufficiently large to comprise 36 plots, we might prefer to use the following layout:

	1	2	3	4	5	6
1	A	B	C	D	E	F
2	F	A	B	C	D	E
3	E	F	A	B	C	D
4	D	E	F	A	B	C
5	C	D	E	F	A	B
6	B	C	D	E	F	A

This arrangement, in which each of the six letters occurs once in each row and once in each column, is called a latin square. An experiment carried out according to this plan would fulfill the principles of replication, local control and symmetry, but not the principle of randomization. The fertility increasing from left to right and from bottom to top there will be a bias between varieties A and B, B and C, etc. if we use the plan given above. We may call this a <u>diagonal latin square</u> because all the A occur on the main diagonal.

The bias occuring in the diagonal square was soon found out and it was proposed to use a latin square known as Knut-Vik-square, at which we arrive in shifting letters by two columns when we go from one row to the next. This procedure works only for odd numbers of treatments; with 5 treatments we would obtain the following arrangement.

Knut-Vik-square.

	1	2	3	4	5
1	A	B	C	D	E
2	D	E	A	B	C
3	B	C	D	E	A
4	E	A	B	C	D
5	C	D	E	A	B

In this design each variety is nicely spread over the whole field. It may be assumed that the bias, if any, is smaller than in the diagonal square. Knut-Vik-squares, however, suffer from another drawback, which will be considered later.

As already mentioned, the principle of randomization is absent from a diagonal square. A randomized latin square is readily obtained starting from a diagonal square. All we have to do is, first to make a random permutation of columns, then secondly a random permutation of rows, and finally assign the varieties (or treatments) at random to the letters. The diagonal square from which we start has to be selected from a number of different squares which it is impossible to transform into each other by the process of permutation outlined above. These fundamental squares are listed in Fisher and Yates (1938). The randomization procedure ensures the validity of tests of hypothesis and of fiducial or confidence limits.

For the statistical analysis of latin square designs we assume the following additive model:

$$y_{k\ell(j)} = \alpha + \tau_j + \beta_k + \gamma_\ell + \delta_{k\ell(j)}$$

where

$y_{k\ell(j)}$ = value on unit in row k and column ℓ with treatment j,

α = general mean,

τ_j = effect of treatment j,

β_k = effect of rwo k,

γ_ℓ = effect of column ℓ,

$\delta_{k\ell(j)}$ = random effect on unit in row k and column ℓ with treatment j.

It can be easily seen that the least-square estimates of α, τ_j, β_k and γ_ℓ are given by

$$a = \bar{y} \quad ,$$

$$t_j = \bar{y}_{(j)} - \bar{y} \quad ,$$

$$b_k = \bar{y}_{k.} - \bar{y} \quad ,$$

$$c_\ell = \bar{y}_{.\ell} - \bar{y} \quad ,$$

respectively, where

$\bar{y}_{(j)}$ = mean of values under treatment j,

$\bar{y}_{k.}$ = mean of values in row k,

$\bar{y}_{.\ell}$ = mean of values in column ℓ,

$\bar{y}$ = general mean.

The sums of squares to be considered are

$$A = \text{Total SS} = \underset{k}{S}\,\underset{\ell}{S}(y_{k\ell(j)} - \bar{y})^2 = \underset{k}{S}\,\underset{\ell}{S}\, y^2_{k\ell(j)} - T^2/N$$

$$B = \text{Treatment SS} = r\,\underset{j}{S}(\bar{y}_j - \bar{y})^2 = (\underset{j}{S}\, T_j^2)/r - T^2/N$$

$$C = \text{Rows SS} = r\,\underset{k}{S}(\bar{y}_{k.} - \bar{y})^2 = (\underset{k}{S}\, T^2_{k.})/r - T^2/N$$

$$D = \text{Columns SS} = r\,\underset{\ell}{S}(\bar{y}_{.\ell} - \bar{y})^2 = \underset{\ell}{S}\, T^2_{.\ell})/r - T^2/N$$

$$E = \text{Error SS} = \underset{k}{S}\,\underset{\ell}{S}(y_{k\ell(j)} - \bar{y}_j - \bar{y}_{k.} - \bar{y}_{.\ell} + 2\,\bar{y})^2 = A - B - C - D$$

and the relation

Total SS = Treatment SS + Rows SS + Columns SS + Error SS.

holds. The number of treatments, replications, rows and columns is r. The analysis of variance is set up in the following table

Variance	D.F.	SS	MS
Treatments	r-1	$t\,\underset{j}{S}(\bar{y}_j-\bar{y})^2$	D_T
Rows	r-1	$t\,\underset{k}{S}(\bar{y}_{k.}-\bar{y})^2$	D_R
Columns	r-1	$t\,\underset{\ell}{S}(\bar{y}_{.\ell}-\bar{y})^2$	D_C
Errors	(r-1)(r-2)	$\underset{k}{S}\,\underset{\ell}{S}(y_{k\ell(j)} - \bar{y}_j-\bar{y}_{k.}-\bar{y}_{.\ell}+2\bar{y})^2$	s^2
Total	r^2-1	$\underset{k}{S}\,\underset{\ell}{S}(y_{k\ell(j)}-\bar{y})^2$	...

Here again the hypothesis of equal treatment effects $\tau_1 = \tau_2 = \dots = \tau_r = 0$ is tested using

$$F = D_T/s^2$$

and comparing this to F_P with $n_1 = r-1$, $n_2 = (r-1)(r-2)$. And to test we use

$$t = \frac{\bar{y}_1 - \bar{y}_2}{s} \sqrt{\frac{r}{2}}$$

with $n = (r-1)(r-2)$.

The symmetry in the latin square design ensures the orthogonality between treatments, rows and columns.

The breakdown of the total variability into four portions may be used to show the differences between random, diagonal and Knut-Vik latin squares. Let us consider a field trial on a given area of land, subdivided into r^2 plots. If we carry out an uniformity trial, then

$$\tau_1 = \tau_2 = \dots = \tau_r = 0$$

and the total sum of squares will be the same whether we use a random, a diagonal or a Knut-Vik design. In the diagonal design, as we have already found, the treatment sum of squares will be inflated, it will not only contain the inevitable random errors, but also systematic errors. On the other hand in the Knut-Vik square the treatment sum of squares will be smaller than in a random square. The sums of squares for rows and for columns will be the same for all three designs, as also the total sum of squares. It follows from this that the error sum of squares will be too small in the diagonal square, too big in the Knut-Vik square. Consequently, the F-ratio will be too high in the diagonal square and too low in the Knut-Vik square. In other words: using the diagonal latin square we will declare too often that there are treatment differences, while using the Knut-Vik square we will not declare often enough that there are treatment differences. Only the randomized latin square gives the correct proportion of significant results.

Example. An experiment was carried out by forest engineers in 1946 to compare the relative efficiency of three types of saws, A, B, C,. Of each type two individual saws were used to insure that differences between types of saws were not smaller than between two saws of the same type. The correspondance between letters and types of saws is as follows:

Type of saw:	1	2	3
Saw no. 1:	A	B	C
Saw no. 2:	D	E	F

The saws were used by six teams of two workmen, some teams more, some less experienced. The wood cutting was performed on three species of trees: spruce, pine and larch, once with and once without bark. The experiment was carried out in a latin square design, properly randomized, with the six types of wood as rows and the six teams of workmen as columns. Actually the experiment was repeated twelve times. The following table shows the results of one of these experiments (Zehudes, Weber and Linder, 1951).

Cutting times in 1/100 minutes

Wood	Teams						Total
	1	2	3	4	5	6	
1	C 63	E 109	D 98	B 75	A 46	F 49	440
2	B 68	F 62	E 78	A 60	D 40	C 42	350
3	E 127	A 135	B 125	F 73	C 61	D 74	595
4	F 88	D 102	A 125	C 86	E 61	B 60	522
5	D 74	B 100	C 83	E 64	F 43	A 56	420
6	A 130	C 120	F 120	D 113	B 60	E 98	641
Total	550	628	629	471	311	379	2968

The analysis proceeds in the following way.

(1) The totals for treatments are set out:

A	552	D	501	A+D	1053
B	488	E	537	B+E	1025
C	455	F	435	C+F	890

(2) For the analysis of variance we get

Variance	DF	SS	MS
Saws	5	1,716.2	343.2
Types of wood	5	10,286.6	2,057.3
Teams of workers	5	14,426.2	2,885.2
Error	20	2,101.9	105.1
Total	35	28,530.9	...

(3) To test the hypothesis $\tau_1 = \tau_2 = \dots = \tau_6 = 0$ of equal effects of all saws we calculate

$$F = 343.2/105.1 = 3.27$$

and compare this to $F_{0.05} = 2.71$ with $n_1 = 5$, $n_2 = 20$.

(4) In this particular experiment the treatment differences may be resolved into orthogonal comparisons in view of the way in which the experiment was designed. In fact it seems indicated to test the differences between the two saws of the same type. Furthermore, as the saws of type C were of a construction essentially different from types A and B, we may also compare the saws of type A to those of type B and finally type C to types A and B combined. This gives 5 comparisons between 6 values and the following table shows the coefficients for these comparisons.

Saw	Total	Coefficients for comparison A-D	B-E	C-F	A+D -(B+E)	A+B+D+E -2(C+F)
A	T_1	+1	0	0	+1	+1
B	T_2	0	+1	0	-1	+1
C	T_3	0	0	+1	0	-2
D	T_4	-1	0	0	+1	+1
E	T_5	0	-1	0	-1	+1
F	T_6	0	0	-1	0	-2
Sum	T	0	0	0	0	0

We note that the sums of the coefficients are zero for all comparisons. Also the sum of the products for each pair of comparisons are zero. This means that the five comparisons are orthogonal.

For the comparison A-D we get 552 - 501 = +51 and the sum of squares corresponding to this comparison is $(51)^2/(2)(6) = 216.7$. The factor 2 in the denominator is the sum of the squares of the coefficients for that comparison in the above table, the factor 6 is the number of replications. For the five orthogonal comparisons which we have chosen, we get

Comparison	V	Divisor	Sum of squares
A-D	51	(2)(6) = 12	216.7
B-E	-49	(2)(6) = 12	200.1
C-F	20	(2)(6) = 12	33.3
(A+D) - (B+E)	28	(4)(16) = 24	32.7
A+B+D+E - 2(C+F)	298	(12)(6) = 72	1,233.4
Total (Saws)	...	...	1,716.2

The total of the SS is the SS between saws. This must be so because the comparisons are orthogonal. Each comparison can be tested by calculating the F-ratio. Alternatively we may calculate

$$F_{0.05}s^2 = (4.351)(105.1) = 457.3$$

$$F_{0.01}s^2 = (8.096)(105.1) = 850.9$$

and compare the SS for each comparison with these values. The only comparison which turns out to be significant is that between the saws of type 3 and the two other types.

(5) The residuals

$$d_{k\ell(j)} = y_{k\ell(j)} - \bar{y}_j - \bar{y}_{k.} - \bar{y}_{.\ell} + 2\bar{y}$$

may be calculated and inspected to find out whether they are randomly distributed over rows, columns and treatments. This topic will be taken up in 2.6, where the assumption of additivity is studied in detail.

In the experiment on saws the experimenter deliberately introduced heterogeneity between columns (teams of workers) and rows (types of wood). This heterogeneity was afterwards eliminated from the error by means of the analysis of variance. It seems at first sight that there is no point in proceeding in this way. However, it may be pointed out that by doing so, the experimenter obtained a much broader basis for inferences.

The principle of local control has been used in three different ways in the preceding chapters, as summarized in the following table:

Design	Local control
Completely randomized	No local control
Randomized blocks	One-way local control
Latin square	Two-way local control.

Orthogonal comparisons of treatment effects are often helpful, but they should only be used if the comparisons are meaningful. The possibility of making orthogonal comparisons should be born in mind when the experiment is designed. Two examples may illustrate the procedure.

The first concerns a grazing experiment carried out on an alpine pasture in Switzerland (Linder, 1953, p. 49). In a randomized block layout with 6 treatments and 5 replications the possible detrimental effect on grazing of applying fertilizers was investigated. The treatments were chosen as follows:

(1) No fertilizer
(2) Ca
(3) PKCa
(4) $PKCaN_1$ (0.7 kg/a for N)
(5) $PKCaN_2$ (1.4 kg/a for N)
(6) $PKCaN_3$ (2.1 kg/a for N)

Two cows were allowed to graze during about 2 hours in a fenced area where the grass had been cut, except on the 30 experimental plots. The following five orthogonal comparisons seem to be meaningful in this experiment.

Treatment	Coefficients for comparisons					Total grazing time 1/10 minutes
	(1)	(2)	(3)	(4)	(5)	
0	-5	0	0	0	0	223
Ca	+1	-4	0	0	0	317
PKCa	+1	+1	-3	0	0	390
$PKCaN_1$	+1	+1	+1	-1	-1	459
$PKCaN_2$	+1	+1	+1	0	+2	545
$PKCaN_3$	+1	+1	+1	+1	-1	608
Total	0	0	0	0	0	2.542

Comparisons (4) and (5) give the linear and quadratic component of the effect of N. From the analysis of variance we note that the SS for treatments is 20,545.5 and the error mean square is $s^2 = 907.4$ with 20 degrees of freedom.

As a second example we consider an experiment carried out to investigate the effect of fluoride in water, milk and solid food on dental caries in rats. The experiment was completely randomized, with 8 rats per treatment. The effect of treatments was measured as the number of a specific type of lesions in the in the enamel of molars, called "scores". The experiment was expected to answer the following questions: (a) Does fluoride act differently when added to fluid or solid food? (b) I there an effect of fluoride at all? (c) Is there a difference between animals receiving milk or water?

The constitution of the 6 treatments, the results and the 5orthogonal comparisons chosen are given in the following table, where f_0, f_1 and for "no fluoride" and "fluoride" in solid food; F_0 F_1 for "no fluoride" and "fluoride" in liquie food; while m_0 indicates that the animal got water, m_1 that it received milk.

Treatmtne	Fluoride in solid food	Fluoride in liquid food	milk vs. water	"Scores"	Coefficients for comparison (1)	(2)	(3)	(4)	(5)
1	f_0	F_0	m_0	90	+2	0	+1	+2	0
2	f_0	F_0	m_1	86	+2	0	-1	-2	0
3	f_0	F_1	m_0	84	-1	+1	+1	-1	+1
4	f_0	F_1	m_1	83	-1	+1	-1	+1	-1
5	f_1	F_0	m_0	77	-1	-1	+1	-1	-1
6	f_1	F_0	m_1	71	-1	-1	-1	+1	+1
Total	...	...	...	491	0	0	0	0	0

Comparison (1) gives the answer to question (b), comparison (2) answers (a) and comparison (3) answers (c). Comparisons (4) and (5) are subsidiary, completing the set of 5 orthogonal comparisons. (4) is the interaction between (1) and (3), (5) the interaction between (2) and (3). From the analysis of variance for the whole experiment the SS for treatments is 28.854 and the error mean square $s^2 = 3.209$ with 42 degrees of freedom.

2.4 Missing values.

In a completely randomized experiment missing values cause no special difficulty for the analysis. Of course, if only a few observations are left in one treatment, the precision will be impaired, but this does not mean that there is any difficulty in analyzing the experiment.

If, however, values are missing in a randomized block experiment, the remaining values will usually lack the symmetry which makes the analysis so simple. In this case we may proceed in several different ways.

If the missing values all lie in the same block, as it may happen in any field trial, the values from the whole block may be dropped, and the experiment analysed as explained in 2.2.

Similarly, if the missing values occur all in the same treatment, the experiment may be analysed without taking any value from this treatment. Such a case may occur in a variety trial where the seed of one specific variety has suffered damage and does not germinate. Or in an experiment with insecticides, one of the insecticides tested may have no effect at all.

If only one value is missing the simplest procedure is to estimate the missing value and then perform the analysis according to standard methods. F. Yates, following a suggestion of R. A. Fisher, has worked out the following procedure. The missing value is replaced by y. In the analysis of variance we then get for the residual sum of squares $Ay^2 - 2By + C$. Minimizing this sum of squares we get $y = B/A$. This value for y is inserted and the experiment analysed as if the data were complete. The only point in which the analysis differs from the standard method is that the error sum of squares has one degree of freedom less, as also the total sum of squares, corresponding to the estimated missing value.

Suppose the missing value occurs in block k and treatment j. Using the notation

$T^*_{1.}$ = Total for values under treatment 1 (incomplete);

$T^*_{.j}$ = Total for values in block 1 (incomplete);

T^* = General total (incomplete).

y^*_{kj} = Estimate of missing value.

Then the procedure outlined above gives

$$y^*_{kj} = \frac{t\, T^*_{k.} + r\, T^*_{.j} - T^*}{(t-1)(r-1)}$$

The same method can be adapted for the <u>latin square</u> design. If the missing value occurs in row k, column ℓ and belongs to treatment j, and with the notation

$T^*_{k.}$ = Total of values in row k (incomplete

$T^*_{.\ell}$ = Total of values in column ℓ (incomplete)

T^*_{j} = Total of values from treatment j (incomplete)

T^* = General total (incomplete)

$y^*_{k\ell(j)}$ = Estimate of missing value

we obtain by Yates' procedure

$$y^*_{k\ell(j)} = \frac{r(T^*_j + T^*_{k.} + T^*_{.\ell}) - 2T^*}{(r-1)(t-1)}$$

When more than one value is missing special methods were developed for special designs for which any standard textbook may be consulted. In some cases it may be necessary to apply a <u>general method</u> in order to analyze a well planned experiment in which some data are missing, or to analyze an experiment lacking symmetry.

The general method, which we will describe for a particular example consists in setting up a model. The parameters of this model are estimated by the method of least squares, or more generally by the method of maximum likelihood. The residual sum of squares S_1 with n_1 degrees of freedom is calculated. In order to test the hypothesis that the parameters of any subset of all parameters are equal, we determine new estimates under this hypothesis and again calculate the residual sum of squares S_2 with n_2 degrees of freedom. S_2 must be greater or equal to S_1, n_2 is greater than n_1. If the hypothesis is true, the expression

$$\frac{S_2-S_1}{n_2-n_1} : \frac{S_1}{n_1} = n_1(S_2-S_1)/(n_2-n_1)S_1$$

is distributed according to F with (n_2-n_1) and n_1 degrees of freedom.

Consider the example on the determination of the atomic weight of iodine, for which the data are given on p.12. The individual values depend on the silver preparations on one hand and on the iodine samples on the other hand. Consequently the i-th value of y_{jki} for silver preparation (j) and iodine sample (k) may be written as

$$y_{jki} = \alpha + \beta_{(j)} + \gamma_{(k)} + \delta_{(j)(k)i}$$

where α is a general mean, $\beta_{(j)}$ specifies the effect of the silver preparation (j) and γ (k) the effect of iodine sample (k). The random error associated to each observation is $\delta_{(j)(k)i}$, we will assume that it is independent of both silver preparation and iodine sample and is normally distributed. We are choosing $\beta_{(1)} = 0$ and $\gamma_{(\ell)} = 0$, so as to have no redundancy. Thus the individual values y, the totals T and the expected values $\alpha + \beta_{(j)} + \gamma_{(k)}$ are as follows:

Silver preparation

Iodine sample	A	B	C	D	E		
I	y_{11i}	y_{21i}	y_{31i}	y_{41i}	y_{51i}		Individual values
II	y12i	y_{22i}	y_{32i}	y_{42i}	y_{52i}		
I	T_{11}	T_{21}	T_{31}	T_{41}	T_{51}	$T_{.1}$	Totals
II	T_{12}	T_{22}	T_{32}	T_{42}	T_{52}	$T_{.2}$	
	$T_{1.}$	$T_{2.}$	$T_{3.}$	$T_{4.}$	$T_{5.}$	T	
I	$\alpha+\gamma$	$\alpha+\beta_1+\gamma$	$\alpha+\beta_2+\gamma$	$\alpha+\beta_3+\gamma$	$\alpha+\beta_4+\gamma$		Expected values
II	α	$\alpha+\beta_1$	$\alpha+\beta_2$	$\alpha+\beta_3$	$\alpha+\beta_4$		
I	N_{11}	N_{21}	N_{31}	N_{41}	N_{51}	$N_{.1}$	Number of
II	N_{12}	N_{22}	N_{32}	N_{42}	N_{52}	$N_{.2}$	observations
	$N_{1.}$	$N_{2.}$	$N_{3.}$	$N_{4.}$	$N_{5.}$	N	

We obtain estimates a, b_1, b_2, b_3, b_4 and c of the parameters α, β_1, β_2, β_3 β_4 and γ by minimizing

$$\underset{j}{S}\ \underset{k}{S}\ \underset{i}{S}(y_{jki} - \alpha - \beta_{(j)} - \gamma_{(k)})^2 \quad .$$

The derivatives of this sum of squares with respect to α, β_1, β_2, β_3, β_4 and γ give the following 6 linear equations, when put equal to zero.

$$\begin{aligned}
&T - Na - N_{2.}b_1 - N_{3.}b_2 - N_{4.}b_3 - N_{5.}b_4 - N_{.1}c = 0\\
&T_{2.} - N_{2.}b_1 - N_{2.}b_1 - N_{21}c = 0\\
&T_{3.} - N_{3.}a - N_{3.}b_2 - N_{31}c = 0\\
&T_{4.} - N_{4.}a - N_{4.}b_3 - N_{41}c = 0\\
&T_{5.} - N_{5.}a - N_{5.}b_4 - N_{51}c = 0\\
&T_{.1} - N_{.1}a - N_{21}b_1 - N_{31}b_2 - N_{41}b_3 - N_{51}b_4 - N_{.1}c = 0
\end{aligned}$$

Inserting the numerical values for the N's and T's this gives

$$
\begin{array}{llllllll}
16a & + 4b_1 & + 3b_2 & + 3b_3 & + b_4 & + 10c & = 523 \\
4a & + 4b_1 & & & & + 2c & = 118 \\
3a & & + eb_2 & & & + 3c & = 86 \\
3a & & & + 3b_3 & & + 2c & = 160 \\
a & & & & + b_4 & + c & = 55 \\
10a & + 2b_1 & + 3b_2 & + 2b_3 & + b_4 & + 10c & = 369
\end{array}
$$

These equations may be solved by standard methods. However, a method given by W. L.Stevens (1948) seems to be particularly appropriate for finding the solutions of such equations. The calculations proceed as follows:

(1) The frequencies are set out.

(2) From the original totals we compute the averages for iodine samples and obtain a first adjustment for iodine samples; $-36.900 + 25.667 = 11.233$

(3) Adjusted totals are obtained:

$369 + (-11.233)(10) = 256.670$ $\qquad$ $104 + (-11.233)(2) = 81.534$

$154 + (\ 0.000)(6) = 154.00$ $\qquad$ $118 + (-11.233)(2) = 95.534$

$86 + (-11.233)(3) = 52.301$

etc.

The averages from the adjusted totals are calculated for the silver preparations.

Adjustments of constants.

Iodine samples			Silver preparations		
(1) Frequencies					
I	10 = 2 + 2 + 3 + 2 + 1		A	5 = 2 + 3	
II	6 = 3 + 2 + 0 + 1 + 0		B	4 = 2 + 2	
			C	3 = 3 + 0	
			D	3 = 2 + 1	
			E	1 = 1 + 0	

Total	Mean	Adjustment	Total	Mean	Adjustment
(2) Original totals. First adjustment for iodine samples.					
369	36.900	- 11.233	104		
154	25.667	0.000	118		
523			86		
			160		
			55		
			523		
(3) Corrected totals. First adjustment for silver preparations.					
256.670			81.534	16.3068	0.0000
154.000			95.534	23.8835	- 7.5767
410.670			52.301	17.4337	- 1.1269
			137.534	45.8447	-29.5379
			43.767	43.7670	-27.4602
			410.670		
(4) Corrected totals. Second adjustment for iodine samples.					
151.599	15.160	+ 3.058	81.534		
109.308	18.218	0.000	65.226		
260.907			48.920		
			48.920		
			16.307		
			260.907		

Total	Mean	Adjustment	Total	Mean	Adjustment
(5) Corrected totals. Second adjustment for silver preparations.					
182.179			87.650	17.530	0.000
109.308			71.342	17.836	0.306
291.487			58.094	19.365	- 1.835
			55.036	18.345	- 0.815
			19.365	19.365	- 1.835
			291.487		
(6) Corrected totals. Th3rd adjustment for iodine samples.					
172.597	17.260	+ 0.720	87.650		
107.881	17.980	0.000	70.118		
@*).$&*			52.589		
			52.591		
			17.530		
			280.478		
(7) Corrected totals. Last adjustment for silver preparation					
179.797	17.980	0.000	89.090	17.818	0.000
107.887	17.980	0.000	71.558	17.889	- 0.071
287.678			54.749	18.250	- 0.432
			54.031	18.010	- 0.192
			18.250	18.250	- 0.432
			287.678		

(4) The adjustments obtained in (3) are used to get new adjusted totals.

$$256.670 + (2)(0.000) + (2)(-7.5767) + (3)(-1.1269) + (2)(-29.5379)$$
$$+ (1)(-27.4602) = 151.599, \text{ etc.}$$

After step (7) the adjustments get so small as to be negligible.

The estimates are obtained as the negative value of the sum of the adjustments in all steps. This is shown below:

Adjustment	Iodine sample		Silver preparation				
	I	II	A	B	C	D	E
First	−11.233	0.000	0.000	−7.577	−7.577	−29.538	−27.460
Second	+3.058	0.000	0.000	−0.306	−1.335	−0.815	−1.835
Third	+0.720	0.000	0.000	−0.071	−0.432	−0.192	−0.432
Total	7.455	0.000	0.000	−7.954	−3.394	−30.535	−29.727
Estimates	c=+7.455	...	...	b_1=+7.954	b_2=+3.394	b_3=+30.535	b_4=+29.727

The estimate a is obtained starting from the last general total 287.673 and using the adjustments obtained in (7).

Corrected total in (7)		287.673
Adjustments: Iodine samples	(10)(0.000)	0.000
	(6)(0.000)	0.000
Silver preparation	(5)(0.000)	0.000
	(4)(−0.071)	−0.284
	(3)(−0.432)	−1.296
	(3)(−0.192)	−0.576
	(1)(−0.432)	−0.432
Final corrected total		285.090

The estimate a is

$$\underline{a} = 285.090/16 = \underline{17.818}.$$

From the estimates a, b_j and c the estimated means for each class may be calculated. For instance the estimated mean for iodine sample I and silver preparation C is

$$17.818 + 7.455 + 3.394 = 28.667$$

and the observed mean is

$$(29 + 20 + 37)/3 = 28.667;$$

the following table shows the expected and observed means.

	Observed means		Expected means		Differences	
Iodine sample	I	II	I	II	I	II
Silver preparations						
A	25.3	17.8	23.5	19.0	+ 1.8	− 1.2
B	33.3	25.8	41.0	18.0	− 7.7	+ 7.8
C	28.7	21.2	28.7	...	0.0	...
D	55.8	48.3	49.5	61.0	+ 6.3	−12.7
E	55.0	47.5	55.0	...	0.0	...

As a test for the hypothesis $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \gamma = 0$ we can set up an analysis of variance in which the sum of squares for the constants is obtained as

$$SS(\text{constants}) = (104)(0.000) + (113)(+7954) + (86)(+3.394) + (160)(+30.535) + (55)(+29.727) + (369)(+7.455) + (154)(0.000) + (523)(+17.818) - 523^2/16 = 2\ 725.188$$

and where we show also thevariability between and within the classes formed by iodine samples and silver preparations.

Variance	D.F.	SS	MS	F	$F_{0.05}$
Constants	5	2 725.188	545.038	4.186	3.688
Residual	2	488.583	244.292	1.876	4.459
Between classes	7	3 213.771	459.110	3.526	3.499
Within classes	8	1 041.667	130.208	...	...
Total	15	4 255.438	...	...	...

It follows that as a whole the constants contradict the null hypothesis.

If we want to see whether there are differences between the iodine samples or between the silver preparations, or in both classifications, we have to adopt the procedure outlined in general above.

Consider first the case where we suppose that there are no differences between the two iodine samples. The model for this would be

$$y_{ji} = \alpha + \beta_{(j)} + \delta_{(j)i}$$

where j refers to the silver preparations. The sum of squares for the constants in this model is simply:

$$SS(\text{silver preparations}) = (104)^2/5 + (113)^2/4 + (86)^2/3 + (160)^2/3 + (55)^2/1 - (523)^2/16 = 2\ 572.304$$

with 4 degrees of freedom. If we compare this to the sum of squares for constants obtained previously, we get

	D.F.	S.S.
All constants	5	2 725.188
Silver preparations	4	2 572.304
Difference	1	152.884

The increase in the SS due to the adjustment of a constant for the difference between the iodine samples is 152.884 and this is almost equal to the mean square 130.208 within classes. From this we conclude that there is no significant effect due to the iodine samples.

In a similar way we can test the differences between silver preparations. to do this we set up the model

$$y_{(k)i} = \alpha + \gamma_{(k)} + \delta_{(k)i}$$

where k refers to the two iodine samples. The sum of squares for the constants in this model is

$$SS(\text{iodine samples}) = (369)^2/10 + (154)^2/6 - (523)^2/16 = 473.204 \quad .$$

From this we obtain

	D.F.	S.S.	M.S.
All constants	5	2 725,108	...
Iodine samples	1	473.204	...
Difference	4	2 251.904	562.996

The increase of the sum of squares due to the adjustment of constants for differences between silver preparations is quite substantial. If we compare the mean square 562.996 to the mean square 130.208 within classes, we get

$$F = 562.996/130.208 = 4.324$$

while the value for $n_1 = 4$, $n_2 = 8$ is. 3.838. Consequently we infer that there are significant differences between the silver preparations.

2.5 Information and design.

The concept of local control was introduced as a means of attaining higher precision in the comparisons. It would be useful to be able to measure the gain in precision obtained by grouping units into blocks. R. A. Fisher (1935) introduced the concept of information, which is useful in this connection.

We have seen that the standard error for a treatment mean is $s/\sqrt{r}$, the variance s^2/r. If we add more blocks to a randomized blocks experiment, the cost will be proportionate to the inverse of the variance, which could be called invariance. Fisher considers r/s^2 as the information with respect to a treatment mean.

The inverse $1/s^2$ of the error variance gives a measure of the precision attained in the experiment.

If we wish to evaluate the effect of forming blocks we have to compare the error of a randomized blocks experiment with the error of a completely randomized experiment making use of the same experimental units.

The respective analysis of variance assume the following form:

Completely randomized

Variance	D.F.	M.S.
Between treatments	$t-1$	...
Within treatments	$t(r-1)$	s^2_{CR}
Total	$tr-1$	...

Randomized blocks

Variance	D.F.	M.S.
Treatments	$t-1$	...
Blocks	$r-1$	...
Residual	$(t-1)(r-1)$	$s^2_{R.B.}$
Total	$tr-1$	...

The problem is, how to estimate $s^2_{C.R.}$ if we have performed a randomized blocks experiment and obtained the analysis of variance.

If we had the results of an uniformity trial, we would get for the treatment mean square also $s^2_{R.B.}$. In that case, the total mean square gives us the best estimate of $s^2_{C.R.}$ which is available. So we have

$$s^2_{C.R.} = \frac{[(t-1) + (t-1)(r-1)]s^2_{R.B.} + SS(\text{blocks})}{tr - 1}$$

$$s^2_{C.R.} = \frac{r(t-1)s^2_{R.B.} + SS(\text{blocks})}{tr - 1}$$

We may compare s^2_{CR} to s^2_{RB} or $1/s^2_{CR}$ to $1/s^2_{RB}$, this will give us the relative efficiency of the randomized blocks design compared to the completely randomized design.

$$1/s^2_{CR} : 1/s^2_{RB} = s^2_{RB}/s^2_{CR} \quad .$$

As the two designs have different degrees of freedom for the error variance, this should be taken into account. R. A. Fisher suggests. that two invariances or quantities of information should be compared according to the formula

$$E = \frac{(n_1+1)(n_2+3)}{(n_2+1)(n_1+3)} \cdot \frac{s_2^2}{s_1^2}$$

As an example consider the relative efficiency of randomized block designs in example of § 1.2.

In that example we had

$$t = 5, \quad r = 4, \quad s_{RB}^2 = 707.7, \quad SS(\text{blocks}) = 23,201.0 \quad .$$

Hence

$$s_{CR}^2 = \frac{(4)(4)(707.7) + 23,201.0}{(5)(4) - 1}$$

$$s_{CR}^2 = 2,080.2$$

and the relative efficiency E is

$$E = \frac{(13)(18)(2080.2)}{(16)(15)(707.7)} = 2.866 \quad .$$

The gain in precision through local control is considerable in our example: to obtain the same precision in a completely randomized experiment, we would have to use (20)(2.866) or approximately 57 plots instead of 20 in the randomized blocks experiment.

2.6 Test of non-additivity.

J. W. Tukey (1949, 1955) has shown how the assumption of additivity in the models used in 2.2 and 2.3 for randomized blocks and latin squares can be tested.

For the randomized block designs we consider the model

$$y_{jk} = \alpha + \tau_j + \beta_k + \delta_{jk} \quad ,$$

and we have seen that the expected value Y_{jk} is equal to

$$Y_{jk} = a + t_j + b_k = \bar{y}_{j.} + \bar{y}_{.k} - \bar{y}$$

and the residuals d_{jk}

$$d_{jk} = y_{jk} - Y_{jk} = y_{jk} - \bar{y}_{j.} - \bar{y}_{.k} + \bar{y} \quad .$$

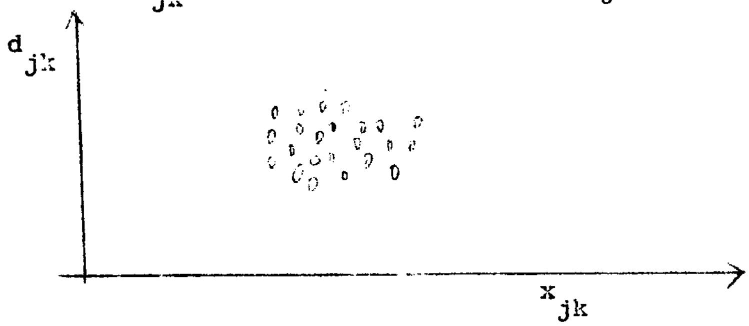
If the data are such that the additivity holds, then the d_{jk} should be independent of the Y_{jk}. Let x_{jk} be defined as

$$x_{jk} = Y_{jk}^2 - \underset{k}{S}\, Y_{jk}^2/r - \underset{j}{S}\, Y_{jk}^2/t + \sum_j \sum_k Y_{jk}^2/rt \quad .$$

Note that

$$\underset{j}{S}\, d_{jk} = \underset{k}{S}\, d_{jk} = 0 \quad , \quad \underset{j}{S}\, x_{jk} = \underset{k}{S}\, x_{jk} = 0 \quad .$$

We prepare a graph on which the d_{jk} are plotted against the x_{jk}, with r·t points.



Inspection of this graph may give a rough idea whether there is any correlation between x_{jk} and d_{jk}. Instead of plotting d_{jk} against x_{jk} we can also compute the linear regression of the d_{jk} against the x_{jk}. In particular, the sum of squares of the regression values is

$$\left[\underset{j}{S}\, \underset{k}{S} (x_{jk} d_{jk}) \right]^2 / \underset{j}{S}\, \underset{k}{S}\, x_{jk}^2 = S_R \text{ (say)}$$

with one degree of freedom. If we subdivide the residual or error sum of squares, say S_E into two parts, one being S_R, the other $S_E - S_R$, then Tukey has shown that

$$\frac{S_E - S_R}{(t-1)(r-1)-1} = s_o^2$$

Hence

$$F = (S_R)/s_o^2 \quad , \quad n_1 = 1 \quad , \quad n_2 = (t-1)(r-1) - 1$$

is distributed according to F, provided the d_{jk} are linearly independent of the x_{jk}.

The computations may be shortened considerably in the case of randomized blocks experiments or in a two-way table in general.

The analysis of variance may be written as

Variance	DF	SS	MS
Treatments	$t-1$	S_T	...
Blocks	$r-1$	S_B	...
Non-additivity	1	S_R	S_R
Remainder	$(t-1)(r-1)-1$	$S_E - S_R$	s_o^2
Total	$tr-1$	$S_T + S_B + S_E$	...

A similar analysis can be carried out for a latin square design, using

$$x_{k\ell(j)} = Y^2_{k\ell(j)} - \frac{\underset{(k,\ell)}{S} Y^2_{k\ell(j)}}{r} - \frac{S_\ell Y^2_{k\ell(j)}}{r} - \frac{\underset{k}{S} Y^2_{k\ell(j)}}{r} + 2\,\frac{\underset{k}{S}\,\underset{\ell}{S}\, Y^2_{k\ell(j)}}{r^2}$$

where $\underset{(k,\ell)}{S}$ indicates a sum over those k and ℓ to which treatments j corresponds.
Here again the values of

$$d_{k\ell(j)} = y_{k\ell(j)} - \bar{y}_j - \bar{y}_{k.} - \bar{y}_{.\ell} + 2\,\bar{y}$$

are plotted against $x_{k\ell(j)}$ and the regression sum of squares is calculated.

3. DESIGN WITH TWO SERIES OF COMPARISONS.

3.1 Comparisons of equal precision.

In chapter 2 we considered experiments designed to evaluate differences in one single series of comparisons, e.g. comparisons between vitamins, between varieties of wheat, between types of saws. Sometimes, however, the experimenter wishes to make two or even more series of comparisons in one single experiment.

An example may make this clear. In a field experiment the yield of three types of grasses was to be compared. These grasses were sown at the same time and together with wheat. It was then found advisable to compare the influence of wheat, of oats and of barley as cover cereals. So we have two series of comparisons: one between three inseed grasses and one between the three cover cereals and no cover cereal at all.

In view of principle IV, the analysis will be simpler, if the three types of grasses are sown under all cover cereals and without cover cereal. This gives 12 combinations.

Assuming that we have 48 plots at our disposal, the experiment could be arranged in 4 blocks with 12 plots in each block, the 12 combinations being arranged at random within blocks.

The number of grass-plants were counted in 9 squares of 1 square foot each. The layout and results (number of grass-plants per square foot) are as follows (Linder, 1953, p. 44)

Block I		Block II		Block III		Block IV	
WF	GM	OM	WR	WR	OR	GF	OR
51	50	45	46	76	54	49	77
OR	WM	HF	HM	HF	GF	HM	GM
81	46	34	30	53	55	53	42
WR	GR	OF	GF	HR	GR	HR	HF
80	80	46	51	70	79	82	50
HM	HF	OR	GR	OF	WF	WF	OM
59	47	55	45	47	50	55	47
OM	GF	WF	GM	HM	WM	GR	OF
67	56	61	31	29	32	75	52
OF	HR	HR	WM	GM	OM	WM	WR
73	66	68	28	35	43	50	70

Cereals: Wheat (W), Oats (H), Barley (G), None (O)

Grasses: R, F, M

Each of the 12 combinations occurs in every block. So we have a randomized block design with $t = 12$ treatments and $r = 4$ blocks. The analysis of variance is straightforward.

Variance	DF	SS	MS
Treatments	11	6 134.8	557.7
Blocks	3	2 305.7	768.6
Residual	33	2 425.6	$s^2 = 73.5$
Total	47	10 865.3	...

The error mean square $s^2 = 73.5$ is valid for all comparisons between the 12 treatment combinations. Specifically, we have

Comparison	standard error
Grasses within cereals	$s\sqrt{2/4}$
Means for grasses (over all cereals	$s\sqrt{2/16}$
Means for cereals (over all grasses)	$s\sqrt{2/12}$

The results of the experiment can be conveniently presented as in the following table:

Cereals	Sums of yields R	F	M	Totals
G	279	211	158	648
W	272	217	156	645
H	286	184	171	641
O	267	218	212	697
Total	1104	830	697	2631

According to this two-way table the 12 treatment combinations can be split up by analysis of variance, in which we use the plot as unit.

Variance	DF	SS	MS
Between cereals	3	172.2	57.7
Between grasses	2	5,383.6	2,691.8
Residual	6	577.7	96.3
Between treatments	11	6,146.6	...

The three sums of squares are computed as follows:

$$\text{SS (cereals)} = (648^2 + 645^2 + 641^2 + 697^2/12 - 2,631^2/48$$

$$\text{SS (grasses)} = (1104^2 + 830^2 + 697^2)/16 - 2,631^2/48$$

$$\text{SS (residual)} = 6146.6 - 173.2 - 5,383.6$$

The nature of the residual variability is the same as the error variability in a two-way table with treatments and blocks as entries.

Blocks	Treatments 1	2	...	j	...	t	Mean
1							
2							
...							
k				y_{jk}			$\bar{y}_{.k}$
...							
r							
Mean				$\bar{y}_{j.}$			$\bar{y}$

In this setup the residual sum of squares can be shown to be

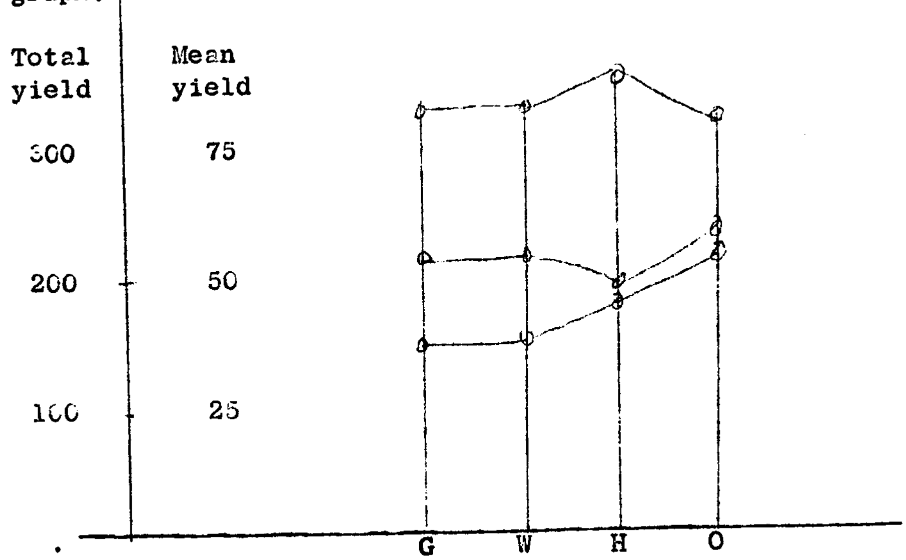
$$SS(\text{residual}) = \underset{j}{S}\,\underset{k}{S}(y_{jk} - \bar{y}_{j.} - \bar{y}_{.k} + \bar{y})^2$$

$$= \underset{j}{S}\,\underset{k}{S}\left[(y_{jk} - \bar{y}_{j.}) - (\bar{y}_{.k} - \bar{y})\right]^2$$

$$= \underset{j}{S}\,\underset{k}{S}\left[(y_{jk} - \bar{y}_{.k}) - (\bar{y}_{j.} - \bar{y})\right]^2$$

The terms $y_{jk} - \bar{y}_{j.} - \bar{y}_{.k} + \bar{y}$ vanish when the differences between blocks are the same for all treatments, or alternatively, when the differences between treatments are the same in each block.

In the above example the data from the table on p. 53 can be plotted in the following graph.



If the three curves were parallel, the residual SS would vanish; we would have <u>independence</u> between cereals and grasses. If we had pronounced non-parallelism, we would have <u>interaction</u> between cereals and grasses. Differences between grasses would vary from one cereal to another.

The residual SS in the preceding analysis of variance is called SS for interaction, and often denoted as interaction C·G (cereals·grasses).

The full analysis of variance for this experiment is

Variance	DF	SS	MS
Cereals	3	173.2	57.7
Grasses	2	5 383.6	2 691.8
Interaction C·G	6	577.7	96.3
Treatments	11	6 146.6	557.7
Blocks	3	2 305.7	768.6
Error	33	2 425.0	s^2 = 73.5
Total	47	10 865.3	...

From this analysis of variance we may draw the following conclusions:

(1) The interaction between grasses and cereals is not significant.

$$F = 96.3/73.5 = 1.310 \; (n_1 = 6, \; n_2 = 33 : \; F_{.05} = 2.4).$$

(2) The cereals have no effect on the yield of grass.

(3) The differences in yield between the grasses are highly significant.

The comparison between the control C and the cereals G, W, H is in the nature of the experiment. The corresponding SS is obtained as

$$[3(697) - 648 - 645 - 641]^2 / 12 \cdot 12 = 171.2 \quad .$$

This SS has one degree of freedom; the variability between G, W and H is 173.2 - 171.2 = 2.0, a very small amount indeed.

The test $F = 171.2/73.5 = 2.3$ ($n_1 = 1$, $n_2 = 33$: $F_{.05} \approx 4.1$) reveals that the difference between the controls and the cereals is not significant.

3.2 Comparisons with unequal precision (split plot design).

In the example of paragraph 3.1 we designed the experiment deliberately in such a way as to have the same precision for the two series of comparisons involved.

The experimenter might have asked us to design the experiment so as to have different precision for the two series of comparisons. There are two possibilities:

(a) The precision for the comparison between grasses should be higher than that between cereals;

(b) The precision for the comparison between cereals should be higher than that between grasses.

It is clearly a question of applying the principle III (Local Control) in an appropriate way. We know that in applying this principle we take into account the fact that some experimental units will give more similar results than others. In field trials, it is known that plots lying close to each other are generally less different than those far apart.

In the case of equal precision for both comparisons, we considered all 12 combinations of cereals with grasses as treatments and arranged them at random within blocks.

Case (a).

In this case we may proceed in the following way

Blocks	I	II	III	IV
W	F R M			
H	H F R			
G	M R F			
O	R F M			

Design (a)

We divide every block into four main plots on which the cereals are arranged at random. Each main plot is further subdivided into three subplots to which the grasses are allocated at random. The comparison of the grasses is done between subplots within plots, the comparison of the cereals between main plots within blocks. The comparison between grasses will therefore be of higher pre cision than that between cereals.

Case (b).

If we want to have higher precision in comparing cereals than grasses, we may divide the four blocks into three main plots to which the grasses are allocated at random. Each main plot is subdivided into four subplots which receive the cereals at random.

Design (b)

	Block I	II	III	IV
R	W H			
	G O			
M	O H			
	G W			
F	G W			
	O H			

It is to be expected, that in (a) comparisons between grasses are more precise than in the design of 3.1, while cereals are compared with lower precision. In (b) the reverse is to be expected.

Split plot designs may also be constructed on the basis of a latin square. The plots of the latin square will serve as main plots and will be subdivided into subplots.

The split plot design can also be used in industrial and other applications. In some such applications days may be considered as main plots, hours as subplots, weeks as blocks.

As an example of a split plot design we consider the following data, showing the influence of the date of cutting (or harvesting) and the influence of varieties of grasses on the yield. As the interest of the experimenter was primarily to investigate the effect of the date of harvesting, the three dates were compared between subplots within plots, while the 3 varieties or mixtures of grass were compared between main plots within blocks.

The experiment was actually carried out in 10 blocks, of which we have chosen only 2, to simplify the presentation. The original data were given in grams. Inspection of the data showed that in view of the high variability between subplots and plots, it was sufficient to consider the values rounded 100 g (Linder, 1953, p. 72).

Yields in 100 g per subplot of 3.5 m^2.

7	B	68	S	70	R	65	R	75	B	115	S	134	2
2	S	90	R	64	B	87	S	63	B	80	R	66	5
5	B	53	S	59	R	47	R	62	B	80	S	83	8
4	B	79	S	77	R	59	R	49	S	44	B	63	7
8	R	74	S	87	B	40	S	91	B	85	R	71	6
3	S	113	R	77	B	100	R	84	B	98	S	98	1
1	S	114	R	87	B	118	B	110	S	95	R	71	3
6	R	64	S	83	B	88	S	117	B	96	R	66	4
			Block I							Block II			

S = early date, B = intermediate date, R = late date.

1, 2, 3, ..., 8 = varieties or mixtures of grass.

In the analysis, we will have to consider two errors, one for the comparison of varieties between plots (within blocks) and one for comparing dates of harvest between subplots (within plots).

(a) Comparison of varieties.

Since each plot contains three subplots with each of three harvesting dates, we can analyse this as a randomized block design.

Variance	DF	SS	MS
Varieties	7	9 265	1 324
Blocks	1	143	143
Error (a)	7	2 895	$s_a^2 = 414$
Between plots	15	12 303	...

In order to compare with the subsequent analysis for dates of harvest, the SS and MS are computed on the basis of subplots.

$$\text{SS (plots)} = (319^2 + 280^2 + \ldots + 251^2 + 225^2)/3 - 3909^2/48$$

$$\text{SS (varieties)} = (599^2 + 535^2 + \ldots + 476^2)/6 - 3909^2/48$$

$$\text{SS (blocks)} = (1913^2 + 1996^2)/24 - 3909^2/48 = (1996 - 1913)^2/48$$

In order to test whether there are any differences between varieties we have to compute

$$F = 1324/414 = 3.20 \quad (n_1 = 7,\ n_2 = 7,\ F_{0.05} \quad 3.8)$$

The differences between grasses are not significant.

(b) Comparisons between dates of harvest (S, B, R).

The total SS between subplots has 47 DF. These are 15 DF between plots and 32 DF (16.2) between subplots within plots.

Of these 32 DF between subplots within plots, there are 2 between dates of harvest.

If we make up a table showing varieities and dates of harvest

Varieties	S	B	R	Total
1	212	216	171	599
2				
3				
4				
5				
6				
7				
8	170	170	136	476
Total	1413	1410	1081	3909

We may compute the interaction V·D between varieties and dates of harvest. This has 2·7 = 14 degrees of freedom, which belongs also to the 32 DF within plots.

There is also an interaction between blocks and dates with 1·2 = 2 DF.

Date	Block I	Block II	Total
S	693	725	1 413
B	683	727	1 410
R	537	544	1 081
Total	1 913	1 996	3 909

This SS (B·D) is considered as part of the error SS, because we will expect, in field experiments at least, that the treatments have the same effect in each block. Experimental fields should not vary markedly from one block to the other with respect to effects of treatment. They may of course vary, but not qualitatively.

The 32 - 14 - 2 - 2 = 14 remaining DF can be identified as interaction between varieties, dates and blocks (BDB). It can be considered either as interaction of VD with blocks, or interaction of VB with dates, or interaction of DB with varieties.

It is instructive to put these results in tabular form:

Variance	Degrees of freedom between Blocks	Plots within Blocks	Sub-plots within plots
Blocks (B)	1	...	...
Varieties (V)	...	7	...
Dates (D)	...	...	2
BV	...	7	...
BD	...	...	2
VD	...	...	14
BVD	...	...	14
Total	1	14	36

The error (b) for comparisons between dates or for judging VD is given by BVD and BD.

$$SS(D) = (1418^2 + 1410^2 + 1081^2)/16 - 3909^2/48 = \underline{4\ 622}$$

$$SS(V \cdot D) = (212^2 + 216^2 + \ldots + 136^2)/2 - 3909^2/48$$

$$= \underset{SS(V)}{9\ 275} - \underset{SS(D)}{4\ 622} = \underline{1\ 532}$$

$$SS(total) = 68^2 + 70^2 + \ldots + 66^2 - 3909^2/48 = \underline{19\ 594}$$

$$SS(subplots\ within\ plots) = 19\ 594 - 12\ 303 = \underline{7\ 291}$$

$$SS(error(b)) = 7\ 291 - 4\ 622 - 1\ 532 = \underline{1\ 136}$$

The complete analysis of variance is shown below:

Variance	DF	SS	MS
Varieties (V)	7	9 265	1 324
Blocks (B)	1	143	143
Error (a) [VB]	7	2 895	s_a^2 = 414
Between plots	15	12 303	...
Dates (D)	2	4 622	2 311
Interaction V·D	14	1 532	109
Error(b)[B·D + BVD]	16	1 136	s_b^2 = 71
Between subplots within plots	32	7 291	...
Total	47	19 594	...

It may be seen that the dates are compared with much higher precision than the varieties.

Interaction V·D: $F = 109/71 = 1.54$ ($n_1 = 14$, $n_2 = 16$: $F_{.05} = 2.4$)

The interaction VD is not significant.

Differences between dates:

$$F = 2\,311/71 = 32.55 \; (n_1 = 2,\; n_2 = 16 : F_{.001} = 10.97$$

Differences between the dates are highly significant. R gives significantly lower yield than S and B.

The SS for VB, DB and BVD could be computed directly and simply because there are only 2 blocks.

As an example let us consider B·D, for which we take the table of results:

Date	Block I	Block II	II-I	$(II-I)^2$
S	693	725	32	1 024
B	683	727	44	1 936
R	537	544	7	49
Total	1 963	1 996	83	3 009

The terms of $(II-I)^2$ are SS within dates, with 3 DF. The interaction BD has 2 DF.

$$3\ 009 : 16 = 188$$

This SS contains SS (blocks) = 143; consequently

$$SS(B\cdot) = 188 - 143 = \underline{45}$$

The SS (BVD) and SS (BV) can be obtained similarly.

Gain in precision.

We can obtain an estimate of the gain (or loss) in precision of our actual experiment with respect to an experiment in which both series of comparisons would have the same precision.

A design giving equal precision to both comparisons would give the following analysis of variance:

Variance		DF		MS
Blocks		1		
Varieties		7		
Dates	Treatments	2	23	
Interaction V·D		14		
Error (B·T)		$n_1 = 23$		s_1^2
Total		47		

If we consider our experiment as an uniformity trial, we would have

Variance	DF		MS	SS
Blocks	1			
Varieties	7	= 14		
Error (a)	7		$413.568 = s_2^2$	5 789, 952
Between plots	15			
Dates	2			
Interaction V·D	14	= 32		
Error (b)	16		$71.000 = s_3^2$	2 272,000
Subplots within plots	32			
Total	47	46		8 061,952

$s_1^2 = 8\ 061{,}952/46 = \underline{175.260}$.

The gain in precision for dates of harvesting can be evaluated as explained in 2.5, comparing s_3^2 to s_1^2; similarly the loss in precision for varieties or mixtures of grass is obtained as a comparison of s_2^2 to s_1^2.

The model underlying the analysis outlined in this paragraph, is

$$y_{ijk} = \alpha + \pi_i + \tau_j + \delta_{ij} + \rho_k + (\tau\rho)_{jk} + \epsilon_{ijk}$$

where

y_{ijk} = observed value for main plot treatment j and subplot treatment k in replication i

α = overall mean; π_i = effect due to i-th replicate;

τ_j = effect of j-th treatment for main plots;

δ_{ij} = random component of error, associated with j-th main plot treatment in i-th replicate (main plot error);

ρ_k = effect of k-th treatment for subplots;

$(\tau\beta)_{jk}$ = effect of interaction between j-th main plot treatment and k-th subplot treatment;

ϵ_{ijk} = random component of error, associated with main plot treatment j, subplot treatment k and replicate i (subplot error).

4. FACTORIAL DESIGNS.

4.1 Introduction. Designs with two levels.

In many cases we are interested in studying the effect of one or preferably more factors. We are investigating here the case when these factors are quantitative variables.

A typical example would be an experiment designed to study the effect on tensile strength of the content of carbon, silicon and manganese in cast steel. The content of these elements can be varied in a continuous way.

Forty years ago, when I was at college, we learnt that the right way to perform such an investigation would be to vary one factor at a time. We would first vary the carbon content, holding constant silicium and manganese. In a second step we would change the silicium content while holding carbon and manganese constant. And finally we would investigate the effect of manganese, while carbon and silicium would remain unchanged.

This procedure of studying only one factor at a time has one major drawback. If we study the effect of carbon for a low content of silicium, it may well be that we get something different from what we would observe with a high silicium-content. With this one-factor-at-a-time procedure we would never find out whether there is any interaction between two factors, or whether the factors act independently.

Even if we knew that the factors were independent, it would be more economical, as we will soon see, to adopt a different procedure than that professed by the physicists forty years ago, and I am afraid still used today.

An entirely different procedure was introduced by R. A. Fisher in his book "The Design of Experiments". We will describe it in terms of the investigation mentioned above.

In its simplest form a factorial design would investigate the effect of two or three factors, and we would choose two levels at which to study the effects of each factor.

If we use 0 and 1 as symbols for lower and upper levels of a factor, then the factorial design according to R. A. Fisher should comprise all the combinations of the factors at levels 0 and 1. In our example this would give the following eight treatment combinations.

Treatment Combination	Levels of factors C	S	M	Symbol
1	0	0	0	(1)
2	1	0	0	c
3	0	1	0	s
4	0	0	1	m
5	1	1	0	cs
6	1	[illegible]	1	cm
7	0	1	1	sm
8	1	1	1	csm

We will use the symbol both to mean one particular combination and the effect observed with that combination, i.e. the tensile strength obtained.

In order to discuss the advantages of such a design, we will study the results, not of an experiment, but of a collection of data obtained by astatistically trained engineer of SULZER BROTHERS (Winterthur, Switzerland), see

Linder, 1953, p. 86. Over the years they had collected chemical analyses and tensile strength for hundreds of cast steel samples. Prior to complete analysis by means of regression methods, the engineer picked out 32 samples according to a factorial design, in which the lower and upper levels of the three factors were defined as follows:

Factor		Level (Content in %) 0	1
Carbon	C	.14 - .15	.26 - .27
Silicon	S	.26 - .29	.56 - .58
Manganese	M	.45 - .49	.55 - .58

From each of the 8 combinations he selected 4 samples.

Treatment Combination	Tensile strength in kg/mm^2 Individual values 1	2	3	4	Total
(1)	44.4	44.4	45.1	43.8	177.7
c	50.5	51.5	49.2	51.9	203.1
s	48.1	45.4	45.7	48.8	188.0
cs	53.2	54.5	52.1	54.0	213.8
m	43.3	43.2	46.5	44.7	177.7
cm	52.1	49.2	51.9	51.3	204.5
sm	48.7	47.0	50.0	49.5	195.2
csm	53.8	58.8	56.2	56.6	225.4

(1) From these data, we may study the influence or the effect of carbon on tensile strength.

One has to be careful about conclusions because this was not a properly designed experiment. It may well be, that what we will consider as effect of carbon is in fact the effect of some element or substance contained in the cast steel, which for some reason is associated with carbon. The lack of randomization should be born in mind.

The influence of C can be evaluated in four ways

								Total
c	203.1	cs	213.8	cm	204.5	csm	225.4	846.8
(1)	177.7	s	188.0	m	177.7	sm	195.2	738.6
c-(1)	25.4	cs-s	25.8	cm-m	26.8	csm-sm	30.2	108.2

To take the differences is an arbitrary way (but the simplest) to measure the effect of factors.

108.2/16 = 6.7625 The average effect of change of carbon from from level 0 to level 1 is called the <u>main effect</u> of G.

(2) In the same way, we may study the effect of S.

								Total
s		cs		cm		csm		
(1)		c		m		cm		
s-(1)	10.3	cs-c	10.7	sm-m	17.3	csm-cm	20.9	59.4

The average effect of S, or main effect is

59.4/16 = 3.7125 .

We see that the same 32 values give us information of equal precision on the influence or effect of C and S.

(3) The same observations can be used to determine the effect of M.

								Total
m		cm		sm		csm		
(1)		c		s		cs		
m-(1)	.00	cm-c	1.4	sm-s	7.2	csm-cm	11.6	20.2

The average or main effect of M is

20.2/16 = 1.2625

(4) <u>Interactions</u>

If we look a little more closely at the components of the main effect C, we see that the values 25.4, 25.8, 26.8 and 30.2 are quite near each other. This is not so in the case of S and M.

In fact, if we consider the component values of S, we find that s-(1) = 10.3 and cs-c = 10.7 differ only to a small extent from each other, and that also sm-m = 17.5 and csm-cm = 20.9 are rather near each other. But there is a marked difference between the two groups, which we may evaluate as follows

				Increase
s-1	10.3	sm-m	17.5	
cs-c	10.7	csm-cm	20.9	
Total	21.0		38.4	17.4

17.4:8 = 2.175 is the interaction between S and M. It shows how much the effect of S is higher if M is on the higher level.

If we consider the effect of M, we see that this is higher when S is on the higher level.

				Increase
m-(s)	0.0	sm-s	7.2	
cm-c	1.4	csm-cs	11.6	
Total	1.4		18.8	17.4

The interaction is symmetrically defined. If we call it SM, it may also be written MS .

(5) In the same way as we computed the interaction SM, we may compute CS.

Effect of C at				
lower		higher		Increase
level of S				
c-(1)	25.4	cs-s	25.8	.4
cm-m	26.8	csm-sm	30.2	3.4
Total	52.2		56.0	3.8

The interaction CS is therefore 3.8:8 = .475.

(6) Similarly, the interaction between C and M may be obtained.

Effect of C at lower level of M		Effect of C at higher level of M		Increase
c-(1)	25.4	cm-m	26.8	1.4
cs-s	25.8	csm-sm	30.2	4.4
Total	51.2		57.0	5.8

Interaction CM: 5.8: 8 = .725 .

This could again be obtained in evaluating the effect of M at higher level of C against lower level of C.

(7) The interaction CS can be obtained at the low level of M and at the high level of M.

(a) Effect of C at lower level of M at

lower level of S		higher level of S		Increase	
c-(1)	25.4	cs-s	25.8	.4	CS for M low

Effect of C at higher level of M at

lower level of S		higher level of S			
cm-m	26.8	csm-sm	30.2	3.4	CS for M high
Increase				3.0	

This gives the interaction between (CS) and M which is called <u>second order interaction or interaction between three factors</u>. It is symmetrically defined and may be written

$$\text{CSM: } 3.0/4 = .750.$$

Remark concerning the definition of main effects and interactions.

Main effects and interactions are in fact linear combinations of the 8 treatments. They can be displayed in the following table:

		Main Effects			Interactions			
Treatment combinations	Total	C	S	M	CS	CM	SM	CSM
(1)	+	-	-	-	+	+	+	-
c	+	+	-	-	-	-	+	+
s	+	-	+	-	-	+	-	+
sc	+	+	+	-	+	-	-	-
m	+	-	-	+	+	-	-	+
cm	+	+	-	+	-	+	-	-
sm	+	-	+	+	-	-	+	+
csm	+	+	+	+	+	+	+	+
Sum	8	0	0	0	0	0	0	0
Sum of squares		8	8	8	8	8	8	8

Note that the coefficients in CS are the products of those in C and S, that the coefficients in CSM are the products of those in C and SM etc. These comparisons are mutually orthogonal. If we compute the SS for each comparison, the sum will be equal to SS (treatments) with 7 DF. The 7 DF between treatment combinations can be resolved into 7 comparisons, which are three main effects and 4 interactions.

Algebraically, the main effects and interactions may be written as follows:

$$C = (c-1)(s+1)(m+1), \qquad S = (c+1)(s-1)(m+1)$$
$$M = (c+1)(s+1)(m-1),$$
$$CS = (c-1)(s-1)(m+1), \qquad CM = (c-1)(s+1)(m-1)$$
$$SM = (c+1)(s-1)(m-1), \qquad CSM = (c-1)(s-1)(m-1)$$

In some textbooks the convention is introduced, that

$$C = \frac{1}{4}[(c-1)(s+1)(m+1)], \qquad CS = \frac{1}{4}[(c-1)(s-1)(m+1)]$$

$$CSM = \frac{1}{4}[(c-1)(s-1)(m-1)] \quad ,$$

where c, cm, csm etc. are defined as averages over replications. We are not going to use these coefficients.

To sum up: Factorial experiments permit the evaluation of main effects and interactions; all with the same precision because each evaluation is based on all 32 observations .

Calculation of main effects and interactions. Yates' method.

F. Yates (1937) has proposed a convenient way for calculating the main effects and interactions in factorial designs with factors at two levels. The totals for the treatment combinations are set out in a definite sequence. The first row contains the total for that combination where all the factors are at the lower level. The next row contains the combination c (any of the three factors), the third row s (one of the remaining factors), the fourth cs, the fifth m (the last factor), and then follow the preceding rows combined with m. The columns (1), (2) and (3) are formed as sums and differences of two items in the preceding columns. Inspection of the following table, in which column (3) contains the total and the main effects and interactions, will make the procedure clear.

Treatment combinations	Total	Sums and differences (1)	(2)	(3)	Effects
(1)	177.7	380.8	782.6	585.4	Total
c	203.1	401.8	802.8	108.2	C
s	188.0	382.8	51.2	59.4	S
cs	213.8	420.6	57.0	3.8	CS
m	177.7	25.4	21.0	20.2	M
cm	204.5	25.8	38.4	5.8	CM
sm	195.2	26.8	.4	17.4	SM
csm	225.4	30.2	3.4	3.0	CSM

Testing the significance of main effects and interactions.

In our example the error if formed by the differences between the four observations in the same treatment combination. The corresponding sum of squares and mean square are readily obtained by analysis of variance.

Variance	DF	SS	MS
Treatment combinations	7	500.109	....
Error	24	47.390	$s^2 = 1.975$
Total	31	547.499	....

The sums of squares for main effects and interactions are all obtained in the same way. The sum of squares SS(C) is

$$SS(C) = (108.2)^2/8.4 = 365.851.$$

This gives the following comprehensive analysis of variance:

Variance	DF	SS	MS
C	1	365.851**	
S	1	110.261**	
M	1	12.751*	
CS	1	.451	
CM	1	1.051	same
SM	1	9.461*	
CSM	1	.281	
Treatments	7	500.109	
Error	24	47.390	s^2=1.975
Total	31	547.499	

The MS for 1 DF are the same as the SS. Instead of computing all F ratio, we may compute $F_{0.05}\, s^2$ or $F_{0.01}\, s^2$ with n_1=1, n_2=24 ($F_{0.05}$ = 4.260, $F_{0.01}$ = 7.823) n

$$F_{0.05}s^2 = 8.412 \qquad F_{0.01}s^2 = 15.447 .$$

All main effects and the interaction SM are significant. In view of the significant interaction between S and M it would appear justified, to consider the effects of S and M separately for the two levels of the other factor. This can be done in tabular form

		Totals m: 0	Totals m: 1	Increase: Total	Increase: Mean
s	0	380.8	382.2	1.4	.175
	1	401.8	420.6	18.8	2.350
Increase, total		21.0	38.4	17.4	....
Increase, mean		2.625	4.800	....	2.175 ($= \frac{1}{8}$ SM)

These increases may be tested separately although they are not orthogonal to all other for significance. Each of them could be considered as a linear comparison into which 16 items enter, 8 with a factor of +1, 8 with a factor of -1. The corresponding SS is e.g.

$$(1.4)^2/16 = 1.96/16 = .1225$$

$$(18.8)^2/16 = 353.44/16 = 22.09$$

and this shows that the first is not significant while the second is so at the 1% level. Factorial designs can be arranged as randomized blocks or latin square designs. The factorial design concerns the treatments, not the local control.

<u>Example</u>. Factorial design arranged in randomized blocks. Fertilizer trial.

Block I	n	nk	np	(1)	npk	k	pk	p
	61	67	56	43	31	79	43	37
Block II	(1)	npk	pk	k	np	n	p	nk
	34	57	66	57	39	48	61	64
Block III	nk	(1)	n	npk	p	np	pk	k
	45	42	47	42	42	44	44	56
Block IV	(1)	k	np	npk	nk	p	pk	n
	23	54	60	53	46	52	60	78

<u>Example.</u> Fertilizer trial arranged in a latin square (Goulden, 1952, p. 205)

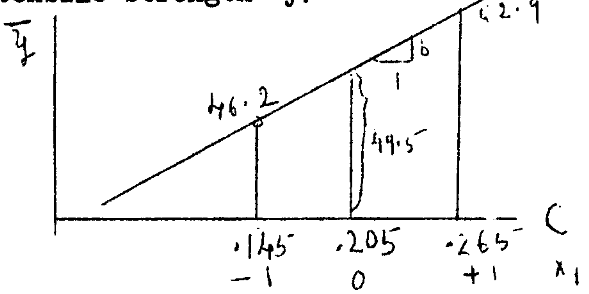
p	n	np	k	nk	(1)	npk	pk
188	122	183	158	114	115	194	189
n	nk	pk	npk	p	k	np	(1)
129	73	174	172	197	120	190	156
nk	np	n	p	(1)	npk	pk	k
107	175	104	180	98	166	175	143
pk	k	npk	(1)	n	np	p	nk
183	126	142	122	114	145	169	161
np	(1)	nk	n	pk	p	k	npk
179	128	133	113	165	156	109	167
k	pk	(1)	np	npk	n	nk	p
149	182	128	171	158	95	89	206
npk	p	k	pk	np	nk	(1)	n
190	189	112	171	179	86	102	145
(1)	npk	p	nk	k	pk	n	np
175	204	208	164	168	185	136	230

<u>Further remarks on the interpretation of factorial designs.</u>

In our example the increase in tensile strength is related to the changes in carbon contents, silicon and manganese content.

To study this relationship we would usually set up a regression equation. In our case tensile strength would be considered as the dependent variable y, the contents in the three elements as independent variates x_1, x_2, x_3.

Let us first consider the simple regression between carbon content C and tensile strength y.



$Y = \bar{y} + b(x_1 - \bar{x}_1)$

If we choose x_1 as indicated in the figure, we get $\bar{x}_1 = 0$, and $Y = \bar{y} + b\, x_1$

846.8/16 = 52.9; 738.6/16 = 46.2; 1585.4/32 = 49.5.

In order to compute b, we have to find

$$\underset{i}{S}\,(x_{1i} - \bar{x}_1)^2 = \underset{i}{S}\, x_{1i}^2 = 16(-1)^2 + 16(+1)^2 = 32$$

and

$$\underset{i}{S}\,(x_{1i} - \bar{x}_1)(y_i - \bar{y}) = \underset{i}{S}\, x_{1i}\, y_i = 846.8 - 738.6 = +108.2$$

We note that $\underset{i}{S}(x_{1i} - \bar{x}_1)(y_i - \bar{y}) = C$ (in our former notation). We further have

$$b = +108.2/32 = +3.4$$

and

$$\underline{Y = 49.5 + 3.4\ x_1}\ .$$

In the same manner we get

$$\underset{i}{S}(x_{2i} - \bar{x}_2)^2 = \underset{i}{S}(x_{3i} - \bar{x}_3)^2 = 32$$

$$\underset{i}{S}(x_{2i} - \bar{x}_2)(y_i - \bar{y}) = \underset{i}{S}\, x_{2i}\, y_i = +59.4\ (= S)$$

$$\underset{i}{S}(x_{3i} - \bar{x}_3)(y_i - \bar{y}) = \underset{i}{S}\, x_{3i}\, y_i = +20.2\ (= M)$$

and for

S: $b = +59.4/32 = +1.9,\qquad Y = 49.5 + 1.9\ x_2$

M: $b = +20.2/32 = +0.6,\qquad Y = 49.5 + 0.6\ x_3$

If we consider the three factors jointly, as we should, we could set up an equation

$$Y = \bar{y} + b_1(x_1 - \bar{x}_1) + b_2(x_2 - \bar{x}_2) + b_3(x_3 - \bar{x}_3)$$

$$Y = \bar{y} + b_1x_1 + b_2x_2 + b_3x_3 \quad ,$$

because we have chosen $\bar{x}_1 = \bar{x}_2 = \bar{x}_3 = 0$.

In this equation, b_1, b_2, b_3 would be the solution of

$$b_1 S_{11} + b_2 S_{12} + b_3 S_{13} = S_{14}$$
$$b_1 S_{12} + b_2 S_{22} + b_3 S_{23} = S_{24}$$
$$b_1 S_{13} + b_2 S_{23} + b_3 S_{33} = S_{34}$$

if we use the convention

$$S_{11} = \underset{i}{S} (x_{1i} - \bar{x}_1)^2 \text{ etc}$$
$$S_{12} = \underset{i}{S} (x_{1i} - \bar{x}_1)(x_{2i} - \bar{x}_2) \text{ etc}$$
$$S_{14} = \underset{i}{S} (x_{1i} - \bar{x}_1)(y_i - \bar{y}) \text{ etc.}$$

Because of the choice of the levels of x_1, x_2, x_3, we note that $S_{12} = S_{13} = S_{23} = 0$. So we get the equations we have already solved, and

$$Y = 49.5 + 3.4\, x_1 + 1.9\, x_2 + .6\, x_3$$

If we want to go back to

C = carbon content
S = silicon content
M = manganese content

we have to take into account that

$$x_1 = (C - .2050)/.0600$$
$$x_2 = (S - .4225)/.1475$$
$$x_3 = (M - .5175)/.0475$$

and this gives

$$Y = 49.5 + 3.4\,(C - .2050)/.0600$$
$$+ 1.9\,(S - .4225)/.1475$$
$$+ .6\,(M - .5175)/.0475$$

$$\underline{Y = 32.36 + 56.67\, C + 12.88\, S + 13.28\, M}$$

This would be the correct formula provided the factors acted independently of each other. We know that this is not so, because the interaction between S and M is significant. This can be taken into account by a change in the regression equation. In fact we may bring in a term $x_2 . x_3 = x_4$ so as to get

$$Y = \bar{y} + b_1(x_1 - \bar{x}_1) + b_2(x_2 - \bar{x}_2) + b_3(x_3 - \bar{x}_3) + b_4(x_4 - \bar{x}_4)$$

As we have already seen, the interaction comparison S-M is orthogonal to C, S and M; it follows that $x_4 = x_2 . x_3$ is orthogonal to x_1, x_2 and x_3, so that $S_{14} = S_{24} = S_{34} = 0$. Also $\bar{x}_4 = 0$, and $S_{44} = 32$. In order to get b_4, we have to determine

$$S_{4y} = \underset{i}{S}\,(x_{4i} - \bar{x}_4)(y_i - \bar{y}) = \underset{i}{S}\,x_{4i}y_i = +17.4 \;(= SM)$$

and from this

$$b_4 = +17.4/32 = 0.54375$$

so the regression equation is

$$\underline{Y \;=\; 49.5 + 3.4\,x_1 + 1.9\,x_2 + 0.6\,x_3 + 0.5\,x_2\,x_3}$$

If one would like to give this equation in the original units he would have to replace $0.5 \cdot x_2 x_3$ by

$$0.5 \cdot \frac{(S - 0.4225)}{0.1475} \cdot \frac{(M - 0.5175)}{0.0475}$$

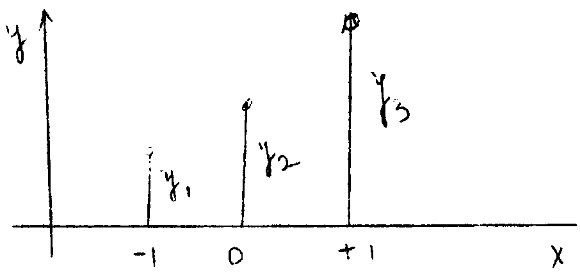
4.2 Designs with more than two levels.

Designs with 2 levels are useful when we explore whether some factors have any effect and whether they interact with each other. Often, however, we want to learn more about the relationship between factors and effects. One thing which may be of interest, is whether the response to any of the factors is approximately linear or not. In order to investigate this, we must have at least three levels.

With two factors each at three levels, we will have $3^2=9$ treatment combinations. If we consider a randomized blocks design it will be necessary to accomodate 9 plots within each block; in a latin square design it will be necessary to have 9 lines and 9 columns.

In order to investigate whether the regression of effect (response) against any factor is linear or not, we will have to compute a quadratic regression.

We consider the special case in which the x-values are arranged symmetrically so that we may take them as -1, 0 and +1. For simplicity, we further assume that we have at each level of x the same number of observations, and that the totals are y_1, y_2, y_3 respectively.



We can fit a quadratic regression by least squares. The regression equation may be written as

$$Y = a + b_1 x + b_2 x^2$$

and we have to find a_1 b_1 and b_2 which make

$$\underset{i}{S} (y_i - a - b_1 x - b_2 x^2)^2 = \text{minimum.}$$

This gives

$$a = \bar{y} = ab_2/3$$

$$b_1 = (y_3 - y_1)/2$$

$$b_2 = (y_1 + y_3 - 2y_2)/2 .$$

The regression coefficients are linear combinations of the three values y_i and they are orthogonal. The sum of squares (with 2 D.F.) between the three y-values can be split up in two portions with 1 DF each, according to two orthogonal comparisons $y_3 - y_1$ and $y_1 + y_3 - 2y_2$.

As an example we consider a fertilizer trial with potatoes arranged in randomized blocks (Rothamsted Experimental Station, Report 1932, p. 177). The two factors and the levels are

N		K	
0	n_o	0	k_o
0.4 cwt N per acre	n_1	1.0 cwt K_2O per acre	k_1
0.8 cwt N per acre	n_2	2.0 cwt K_2O per acre	k_2

Design and yields in lb. on plots of 1/60 acre

Block I (left), Block II (right)

n_1k_o 630	n_ok_o 620	n_2k_2 575	n_1k_1 601	n_2k_o 583	n_1k_o 650
n_1k_2 598	n_ok_2 554	n_2k_1 510	n_ok_o 558	n_2k_2 575	n_1k_2 601
n_ok_1 619	n_2k_o 579	n_1k_1 571	n_ok_2 597	n_2k_1 618	n_ok_1 602
n_1k_1 623	n_2k_o 613	n_1k_o 608	n_ok_o 599	n_ok_1 637	n_1k_2 649
n_1k_2 582	n_ok_2 619	n_ok_1 566	n_2k_o 618	n_1k_o 644	n_1k_1 642
n_2k_1 583	n_ok_o 555	n_2k_2 541	n_2k_1 525	n_ok_2 558	n_2k_2 631

Block III (left), Block IV (right)

The first step in the analysis consists in arranging the results in tabular form.

Treatment combination	Block I	II	III	IV	Total
$n_o k_o$	620	558	555	599	2332
$n_o k_1$	619	602	566	637	2424
$n_o k_2$	554	597	619	558	2328
$n_1 k_o$	630	650	608	644	2532
$n_1 k_1$	571	601	623	642	2437
$n_1 k_2$	598	601	582	649	2430
$n_2 k_o$	519	583	613	618	2393
$n_2 k_1$	510	618	583	525	2236
$n_2 k_2$	575	575	541	631	2322
	5256	5385	5290	5503	21434

From this we get an analysis of variance between treatments, blocks, and error.

Variance	DF	SS	MS
Treatments	8	15,011.1	1,876.4
Blocks	3	4,068.8	1,356.3
Error	24	23,984.7	s^2=999.4
Total	35	43,082.6	...

It is not necessary to test for treatment differences, because the experiment was designed in such a way that the effect of each factor and the interactions could and should be tested. The treatment differences as tested by

$$F = 1,876.4/999.4 = 1.9$$

are something like an average effect which is of no particular interest. In our case it turns out to be nonsignificant, because for $n_1=8$, $n_2=24$: $F_{0.05} = 2.355$.

There are 8 degrees of freedom between treatments. These can be resolved into 2 DF for factor N, 2 DF for factor K and 4 DF for interaction NK.

The 2 DF for N can be split up into a linear and a quadratic component; the same can be done for the 2 DF for K. We will call these N_ℓ , N_q, K_ℓ and K_q. The 4 DF for interaction NK, can be split up as

$$N_\ell \; K_\ell, \quad N_\ell K_q, \quad N_q \; K_\ell \quad \text{and} \quad N_q \; K_q$$

All these comparisons are mutually orthogonal, as will be seen from the table at the top of p. 82. They may be computed as follows.

	n_o	n_1	n_2	Total	N_ℓ	N_q
k_o	2332	2532	2393	7257	+ 61	+ 339
k_1	2424	2437	2236	7097	- 188	+ 214
k_2	2328	2430	2322	7080	- 6	+ 210
Total	7084	7399	6951	21434	- 133	+ 763
K_ℓ	- 4	-102	- 71	-177	- 67	- 129
K_q	+188	- 88	-243	-143	- 431	- 121

From the totals for n_o, n_1 and n_2 we get the linear component of N as $6951 - 7084 = -133$, the N_q as

$$2.7399 - 7084 - 6957 = +763 \; .$$

This shows that the whole effect of N is a decrease in yield, while there is a marked nonlinearity in that the n_1 value is much higher than the average of n_o and n_2. In the same way we get $K_\ell = -177$, $K_q = -143$. K also gives a decrease of yield when applied in high quantities. The interactions are obtained simply by computing N_ℓ and N_q for each level of K, and then computing from the values so obtained, which are shown in the two last columns, the component effects for K, namely K_ℓ and K_q. This gives

$$N_\ell \; K_\ell = -67 \; , \quad N_\ell \; K_q = -431 \; , \quad N_q \, K_\ell = -129, \quad N_q \, K_q = -121.$$

This can be checked by computing K_ℓ , K_q for each column and then computing N_ℓ , N_q for values obtained in each column.

What we have done so far can be summarized in the table which follows:

Treatment combinations	N_ℓ	N_q	K_ℓ	K_q	$N_\ell K_\ell$	$N_\ell K_q$	$N_q K_\ell$	$N_q K_q$
$n_o k_o$	-1	-1	-1	-1	+1	+1	+1	+1
$n_o k_1$	-1	-1	0	+2	0	-2	0	-2
$n_o k_2$	-1	-1	+1	-1	-1	+1	-1	+1
$n_1 k_o$	0	+2	-1	-1	0	0	-2	-2
$n_1 k_1$	0	+2	0	+2	0	0	0	+4
$n_1 k_2$	0	+2	+1	-1	0	0	+2	-2
$n_2 k_o$	+1	-1	-1	-1	-1	-1	+1	+1
$n_2 k_1$	+1	-1	0	+2	0	+2	0	-2
$n_2 k_2$	+1	-1	+1	-1	+1	-1	-1	+1
Total	0	0	0	0	0	0	0	0
Sum of squares	6	18	6	18	4	12	12	36

We are now in a position to calculate the SS for each of the comparisons; e.g.

$$SS(N_\ell) = (133)^2/4.6 = 737.042$$

$$SS(N_q K_q) = (121)^2/4.36 = 101.674$$

The complete analysis of variance is shown below.

Variance	DF	SS	MS
N_ℓ	1	737	
N_q	1	8086*	
K_ℓ	1	8305	same
K_q	1	284	
$N_\ell K_\ell$	1	281	
$N_\ell K_q$	1	3870	
$N_q K_\ell$	1	347	
$N_q K_q$	1	102	
Treatments	8	15011	...
Blocks	3	4087	...
Error	24	23985	$s^2=999.4$
Total	35	43083	...

$n_1=1$, $n_2 = 24$: $F_{0.05} = 4.260$; $F_{0.05}\ s^2 = 4,257$

The only effect which is significant, is N_q. It may be indicated in this particular case, to test separately the increase from n_o to n_1 and the decrease from n_1 to n_2; also the decrease from k_o to k_1.

A 2.4 factorial design.

Fertilizer trial with cabbage (Rothamsted Experimental Station, Report 1932, p. 167).

N: n_o 0
n_1 .2 cwt N per acre
n_2 .4 cwt N per acre
n_3 .8 cwt N per acre

D: d_o 0
d_1 15 tons per acre

Plan and yields (green weights in lb.)

n_2d_1 230	n_3d_o 258	n_2d_o 242	n_od_1 194	n_od_o 145	n_1d_1 268	n_3d_1 285	n_1d_o 185
n_od_1 167	n_1d_1 176	n_3d_1 314	n_2d_o 196	n_2d_1 230	n_1d_o 198	n_3d_o 231	n_od_o 136
n_3d_o 240	n_2d_o 189	n_2d_1 280	n_od_o 127	n_1d_1 240	n_od_1 212	n_1d_o 148	n_3d_1 298
n_1d_1 155	n_od_1 142	n_1d_o 194	n_2d_1 222	n_3d_o 255	n_3d_1 318	n_od_o 118	n_2d_o 208
n_2d_o 180	n_2d_1 212	n_od_1 204	n_3d_1 285	n_1d_o 171	n_od_o 152	n_1d_1 258	n_3d_o 251
n_od_o 124	n_1d_o 176	n_3d_o 247	n_1d_1 155	n_3d_1 283	n_2d_o 170	n_2d_1 254	n_od_1 208
n_1d_o 177	n_3d_1 306	n_od_o 146	n_3d_o 249	n_od_1 203	n_2d_1 260	n_2d_o 184	n_1d_1 206
n_3d_1 275	n_od_o 145	n_1d_1 289	n_1d_o 213	n_2d_o 248	n_3d_o 274	n_od_1 249	n_2d_1 258

The 8 treatment combinations are arranged in a latin square design.

The usual analysis of variance gives $s^2 = 430.7$ as error mean square. The treatment combinations can be presented as in the following table:

	n_o	n_1	n_2	n_3	Total
d_o	1093	1462	1617	2005	6177
d_1	1519	1747	1946	2364	7636
Total	2672	3209	3563	4369	13813
Difference	486	285	329	359	1459

The 3 DF between n_o, n_1, n_2, n_3 can be split up in 3 orthogonal comparisons; N_ℓ , N_q, N_c in the following way

	N_ℓ	N_q	N_c
n_o	-3	+1	-1
n_1	-1	-1	+3
n_2	+1	-1	-3
n_3	+3	+1	+1
Sum	0	0	0
Sum of squares	20	4	20

This is valid, if the x-values are equidistant. We are assuming that the increase from 0.4 to 0.8 is equal in effect to an increase from 0.0 to 0.2 and from 0.2 to 0.4.

It is easy to compute N_ℓ , N_q and N_c, as well as D, $N_\ell D$, $N_q D$ and $N_c D$, each with 1 degree of freedom.

Note: The levels of N are 0, 0.2, 0.4 and 0.8, thus clearly not equidistant. Strictly speaking the orthogonal comparisons are not applicable. If one uses them nevertheless, he will find that N_q and N_c are unimportant, while N is significant - as also D. The interactions are equally not significant. In view of this it would seem legitimate to adjust a straight line to n_o, n_1, n_2, n_3, taking into account the exact values 0, 0.2, 0.4 and 0.8 for x. Doing this, it turns out that an increase in N-fertilizer of 0.1 cwt per acre corresponds to an increase in weight of 12.5 lb on 1/60 acre.

As the number of factors and the number of levels for each factor increases, it becomes more and more difficult to use principle III (Local Control) efficiently. The number of treatment combinations would be

Factors	Levels	Combinations
3	2	$2^3 = 8$
4	2	$2^4 = 16$
5	2	$2^5 = 32$
6	2	$2^6 = 64$
2	3	$3^2 = 9$
3	3	$3^3 = 27$
3	4	$4^3 = 64$

How it is possible to use the principle of factorial design, that is to take all treatment combinations and still to have a manageable size of experiment?

Higher order interactions as error.

One possibility to keep the size of the experiment within a reasonable limit, is to have only one replication and to use higher order interactions as error. This will be a valid approach if the higher order interactions are in fact not significant. In some types of experiment it should be permitted to make this assumption.

As an illustration consider the following example, showing the effect of cyanide, temperature and oxygen content of water on the reaction time of a species of fish (Phoxinus laevis Ag.). The factors and their levels are: (Linder, 1953, p. 97)

Cyanide	(B)	.16 mg [CN]' per ℓ	b_o
		. 8 " " " "	b_1
		4 " " " "	b_2
		20 " " " "	b_3
		100 " " " "	b_4
Temperature (T)		5° C	t_o
		15° C	t_1
		25° C	t_2
Oxygen content (S) of water		1.5 mg O_2 per ℓ	s_o
		3.0 " " " "	s_1
		9.0 " " " "	s_2

Sums of the logarithms (multiplied by 10) of reaction times of 10 fishes (Phoxinus laevis Ag.)

		b_o	b_1	b_2	b_3	b_4	Total
t_o	s_o	201	150	131	130	97	709
	s_1	246	164	138	136	102	786
	s_2	271	170	149	127	99	816
		718	484	418	393	298	2311
t_1	s_o	124	104	86	89	60	463
	s_1	158	111	99	91	74	533
	s_2	207	117	81	87	72	564
		489	332	266	267	206	1560
t_2	s_o	79	63	50	51	32	275
	s_1	129	54	51	52	46	332
	s_2	142	93	62	51	52	400
		350	210	163	154	130	1007
Total	s_o	404	317	267	270	189	1447
	s_1	533	329	288	279	222	1651
	s_2	620	380	292	265	223	1780
		1557	1026	847	814	634	4878

It has repeatedly been found that the regression between the concentration of a poison and reaction time is non-linear and heteroscedastic, but that i it becomes linear and homoscedastic if we take the logarithms of both reaction times and concentration of the poison.

For each of the 45 treatment combinations 10 fishes were observed under exactly identical conditions. It would be tempting to consider the variability between these 10 fishes in each treatment combination as a measure of the experimental error. This would, however, not be correct because the variability between the treatment combination would be much higher, even under the assumption of no effect for all factors, than is the variability between fishes tested simultaneously. If the 10 fishes had been tested separately one at a time, perhaps at different times, then would the variability between them constitute a correct measure of the experimental error. We preferred to consider the total reaction time of 10 fishes as the experimental unit.

It may be mentioned that similar situations may be found when experiments are performed on insects such as bees, flies and the like.

The analysis of variance could be presented as follows:

Variance	DF
B	4
T	2
S	2
BT	8
BS	8
TS	4
BTS	16
Total	44

MS(BTS) = s^2 = error variance under the provision that BTS is not significant against a correctly determined error.

4.3 Confounding

Another way of keeping the size of blocks low while adhering to the principle of factorial design consists in <u>confounding.</u>

We may use the split-plot design so that unimportant comparisons, e.g. higher order interactions, will have lower precision, while main effects and interactions of low order are determined with higher precision. In a 2^3-experiment it may be less important to know the interaction between all three factors with the same precision as the main effects or the two-factor interactions. In a fertilizer trial with N, P, K each at two levels, we have the treatment combinations

$$(1),\ n,\ p,\ k,\quad np,\ nk,\ pk,\quad npk$$

For the interaction NPK, we have

$$NPK = n + p + k + npk - (1) - np - nk - pk$$

If we form two blocks with 4 plots in each replication we could accomodate

I n, p, k, npk in one block, and

II (1), np, nk, pk in the other block

The comparison NPK will then be less precise, but on the other hand all other comparisons will be more precise, because they will be comparisons within blocks. For instance

N	=	n	+ np	+ nk	+ npk	- (1)	- p	- k	- pk
I		+			+		-	-	
II			+	+		-			-

If all yields were higher in Block I for the same amount, this would not affect N, because there are 2 plus and 2 minus signs in this block. This is a consequence of the orthogonality of main effects and interactions. The same is true for P, K, NP, NK and PK.

Let us assume we had 3 replications; what would the analysis of variance become?

(a) Between blocks:

Variance	DF
NPK	1
Replicates (R)	2
Error (a)	2
Blocks	5

What is the exact meaning of the error term? It is the interaction between NPK and replicates, which we could write as (R)(NPK)

Error (a): (R)(NPK) .

We could evaluate this by determining NPK in each replicate and calculate the SS between these three values, this would give SS[(R)(NPK)].

(b) Within blocks:

Variance	DF
N,P,K	3
NP, NK, PK	3
Error (b)	12
Total	18

The 12 DF for Error (b) could be computed as

(R)(N), (R)(P), (R)(K), (R)(NP), (R)(NK), (R)(PK)

each with 2 degrees of freedom.

In this case NPK is completely confounded with differences between blocks. The precision for evaluating NPK will be low indeed, first because of the small degree of freedom for error (a) and second because it is evaluated from differences between blocks, which are usually large.

It is also possible to confound interactions only partially with differences between blocks. If we have 4 replicates, we may confound NP, NK, PK and NPK each in one replicate with the difference between the two blocks. This means that we will be able to ascertain all interactions in three out of four replicates with good precision.

Another possibility is to confound two interactions; this will result in a further reduction of block size.

We consider first an example of complete confounding.

Fertilizer trial with asparagus. (T. Wishart, 1940, p. 25).

Repl.	Block									Block
1	I	npk 40	k 82	p 66	n 47	pk 50	(1) 47	nk 33	np 23	II
2	III	nk 22	np 48	pk 58	(1) 59	k 53	p 72	n 41	npk 35	IV
3	V	(1) 67	np 13	nk 8	pk 10	npk 34	k 24	p 37	n 13	VI
4	VII	n 28	p 9	k 3	npk 23	np 11	(1) 25	nk 2	pk 55	VIII

In all four replications, the interaction NPK is indistinguishably confounded with differences between blocks. The analysis proceeds as usual as far as the computation of main effects and interactions is concerned. That is, we use Yates' method.

Treatment combinations	Blocks I	II	III	IV	V	VI	VII	VIII	Total	Sums and Differences (1)	(2)	(3)	
(1)	...	47	59	...	67	...	...	25	198	327	606	1138	Total
n	47	...	...	41	...	13	28	...	129	279	532	-296	N
p	66	...	...	72	...	37	9	...	184	227	-158	30	P
np	...	23	48	...	13	...	...	11	95	305	-138	36	NP
k	82	...	...	53	...	24	3	...	162	-69	-48	-74	K
nk	...	33	22	...	8	...	...	2	65	-89	78	20	NK
pk	...	50	58	...	10	...	...	55	173	-97	-20	126	PK
npk	40	...	...	35	...	34	23	...	132	-41	56	76	NPK
Total	235	153	187	201	98	108	63	93	1138				
Sum	388		388		206		156		1138				
Difference(NPK)	+82		+14		+10		-30		+76	(check)			

(a) Analysis between blocks.

$$SS(\text{blocks}) = [(235^2 + 153^2 + \dots + 93^2).8 - 1138^2]/32 = 6,502.4$$

$$SS(NPK) = 76^2/32 = 180.5$$

$$SS(\text{repl.}) = [(388^2 + 388^2 + 206^2 + 156^2)4 - 1138^2]/32 = 5,512.4$$

$$SS[R(NPK)] = SS[\text{error (a)}] = 6,502.4 - 180.5 - 5,512.4 = 809.5$$

(b) Analysis within blocks.

$$SS(\text{total}) = [(47^2 + 66^2 + \dots + 55^2 + 93^2/32 - 1138^2]/32 = 14,547.9$$

$$SS(\text{within blocks}) = 14,547.9 - 6,502.4 = 8,045.5$$

$$SS(N) = (296)^2/32 = 2,738.0$$

$$SS(P) = (30)^2/32 = 28.1$$

$$SS(K) = (-74)^2/32 = 171.1$$

$$SS(NP) = (36)^2/32 = 40.5$$

$$SS(NK) = (20)^2/32 = 12.5$$

$$SS(PK) = (126)^2/32 = 496.1$$

$$SS(\text{Treatments, without NPK}) = 3,486.3$$

$$SS[\text{Error (b)}] = 8,045.5 - 3,486.3 = 4,559.2$$

The analysis of variance is given in the following table

Variance	DF	SS	MS	F	$F_{0.05}$
NPK	1	180.5	180.5	.67	10.1
Replications	3	5512.4	1837.5	...	...
R(NPK), Error (a)	3	809.5	269.8	...	...
Blocks	7	6502.4	...	...	...
N	1	2738.0	2738.0	10.8	4.4
P	1	28.1	28.1	...	...
K	1	171.1	171.1	...	...
NP	1	40.5	40.5	...	...
NK	1	12.5	12.5	...	...
PK	1	496.1	496.1	2.0	4.4
Treatments (without NPK)	6	3486.3	...	...	...
Error (b)	18	4559.2	253.3	...	...
Plots within blocks	24	8045.5	...	...	...
Total	31	14547.9	...	...	...

As a check we may compute SS(treatments with NPK)

$$= [(198^2 + 129^2 + \dots + 132^2)\cdot 8 - 1138^2]/32 = 3{,}666.9$$

$$= 180.5 + 2{,}738.0 + 28.1 + \dots + 496.1 = 3{,}486.3 + 180.5 = 3{,}666.8$$

The only effect which is significant is that of N, which is negative, i.e. the higher level of N has given lower yields.

In this experiment the error mean square for the main effects and two-factor-interactions is 253.3; not much smaller than that for NPK. In other cases the difference may be very marked indeed. It is possible to evaluate the gain in precision for N, P, K, NP, NK, PK due to confounding NPK in the same way as shown for the split-plot design.

Partial confounding in a 2^3 design.

We discuss the design and analysis using a fertilizer trial with potatoes (Linder, 1953, p. 114).

Plan and yields

Replication 1 (Blocks I, II) — Replication 2 (Blocks III, IV)

Block I		Block II		Block III		Block IV	
(1)	pk	n	p	np	k	nk	p
196	346	275	348	275	297	98	218
np	nk	npk	k	npk	(1)	n	pk
181	258	383	334	372	262	90	245

Block V		Block VI		Block VII		Block VIII	
npk	p	k	pk	n	pk	p	np
288	218	243	322	222	339	85	44
nk	(1)	np	n	(1)	npk	k	nk
147	183	290	158	178	329	151	41

Replication 3 (Blocks V, VI) — Replication 4 (Blocks VII, VIII)

Inspection of the plan reveals that the interactions confounded are:

Replication:	1	2	3	4
Interaction confounded:	NPK	NP	NK	PK

Formally, we may compute each main factor and each interaction within each of the four replications.

Replication	N, P, K	NP	NK	PK	NPK
1	uncon-	u	u	u	c
2	founded	c	u	u	u
3	with block	u	c	u	u
4	differences	u	u	c	u

In the analysis of variance we have 7 degrees of freedom for the treatments. We have 7 d.f. for differences between blocks. These are formed of 3 d.f. between replications and 4 d.f. between blocks within replications, which actually are the confounded interactions. The remaining 17 degrees of freedom can be shown to consist of

9 d.f. as interactions (R)(N), (R)(P), (R)(K)

8 " " " (R)(NP), (R)(NK), (R)(PK), (R)(NPK)

The first step in the analysis consists as usual in the computation of treatment totals.

Treatment	1		2		3		4		
combination	I	II	III	IV	V	VI	VII	VIII	Total
(1)	196	...	262	...	183	...	178	...	819
n	...	275	...	90	...	158	222	...	745
p	...	348	...	218	218	...	...	85	869
np	181	...	275	...	...	290	...	44	790
k	...	334	297	...	...	243	...	151	1025
nk	258	...	...	98	147	...	...	41	544
pk	346	...	...	245	...	322	339	...	1252
npk	...	383	372	...	288	...	329	...	1372
Total	981	1340	1206	651	836	1013	1068	321	7416
	-	+	+	-	+	-	+	-	
Difference	+359		+555		-177		+747		
Confounded	NPK		NP		NK		PK		

Second step: Yates' method, taking into account confounding.

Treatment combination	Total	Sums and Differences			Effect		Sum of Squares
(1)	819	1564	3223	7416	...	Total	...
n	745	1659	4193	-514	-514	N	$(-514)^2/32 =$ 8256
p	869	1569	-153	+1150	+1150	P	$(1150)^2/32 =$ 41328
np	790	2624	-361	+ 596	+596-555	NP	$(+41)^2/24 =$ 70
k	1025	-74	+95	+ 970	+970	K	$(+970)^2/32 =$ 29403
nk	544	-79	+1055	- 208	-208+177	NK	$(-31)^2/24 =$ 40
pk	1252	-481	-5	+ 960	+960-747	PK	$(+213)^2/24 =$ 1890
npk	1372	+120	+601	+ 606	+606-359	NPK	$(+247)^2/24 =$ 2542

Corrections for interactions:

NP is estimated from replications 1, 3 and 4. Consequently, we have to subtract +555 from +596 to get the correct value +41. Similarly, for the other interactions. The SS is obtained as $(41)^2/24 = 70$, because NP is based on the results of 24 plots.

Analysis of variance.

Treatments account for 7 degrees of freedom. There are 7 d.f. between blocks, of which 3 are between replications and 4 within replications. These 4 are actually the confounded interactions NP, NK, PK and NPK. The total SS has 31 degrees of freedom of which 7 are between blocks and 24 within. Consequently, we have 17 d.f. for error. These can be interpreted as consisting of R(N), R(P), R(K) with 3 d.f. each; and R(NP), R(NK), (R)(PK) and (R)(NPK) with 2 d.f. each.

Variance	DF	SS	MS
N	1	8,256	8,256*
P	1	41,328	41,328**
K	1	29,403	29,403**
NP	1	70	70
NK	1	40	40
PK	1	1,890	1,890
NPK	1	2,542	2,542
Treatments	7	83,529	...
Blocks	7	182,574	26,082
Error	17	18,921	1,113
Total	31	285,024	...

$n_1 = 1,\ n_2 = 17 \qquad F_{0.05} = 4.45 \qquad F_{0.01} = 8.40$

The main effects of N, P and K are significant, the interactions are not significant.

Confounding in 3^3-designs.

In the 3^3-designs it is usual to resort to partial confounding instead of complete confounding. The reason for this is that the interactions can be split up on a formal way, and some of the component degrees of freedom can be confounded with differences between blocks.

Let us consider a 3^2-design. The q treatment combinations of the two factors A and B can be presented as in the following table.

	b_o	b_1	b_2	
a_o	1	4	7	(12)
a_1	2	5	8	(15)
a_2	3	6	9	(18)
	(6)	(15)	(24)	(45)

We have 8 degrees of freedom in all, 2 for A, 2 for B and 4 for AB.

The formal subdivision of A·B may be defined in the following way:

$$I_1 = 1+5+9 \quad (15) \qquad T_1 = 1+6+8 \quad (15)$$
$$I_2 = 2+6+7 \quad (15) \qquad T_2 = 2+4+9 \quad (15)$$
$$I_3 = 3+4+8 \quad (15) \qquad T_3 = 3+5+7 \quad (15)$$

The I component and the T component of the interaction are orthogonal to the main effects A and B.

Treatment combinations	A		B		AB			
	ℓ	g	ℓ	g	I_1/I_2	$I_3/(I_1+I_2)$	T_1/T_2	$T_3/(T_1+T_2)$
1	-1	-1	-1	-1	+1	-1	+1	-1
2	0	+2	-1	-1	-1	-1	-1	-1
3	+1	-1	-1	-1	0	+2	0	+2
4	-1	-1	0	+2	0	+2	-1	-1
5	0	+2	0	+2	+1	-1	0	+2
6	+1	-1	0	+2	-1	-1	+1	-1
7	-1	-1	+1	-1	-1	-1	0	+2
8	0	+2	+1	-1	0	+2	+1	-1
9	+1	-1	+1	-1	+1	-1	-1	-1
Total	0	0	0	0	0	0	0	0
SS	6	18	6	18	6	18	6	18

The orthogonality is a consequence of the existence of a 3x3 graeco-latin square.

$A\alpha$	$C\beta$	$B\gamma$
$B\beta$	$A\gamma$	$C\alpha$
$C\gamma$	$B\alpha$	$A\beta$

The latin and greek letters are orthogonal to each other as also to rows and columns.

If we consider 1, 2, ... 9 as values of a variable, we get the totals given in brackets, and the analysis of variance is:

Variance		DF	SS
A	lin.	1	6.
	quadr.	1	0.
	total	2	6.
B	lin.	1	54.
	quadr.	1	0.
	total	2	54.
AB	I_1/I_2	1	0.
	$I_3/(I_1+I_2)$	1	0
	T_1/T_2	1	0
	$T_3/(T_1+T_2)$	1	0
	Total	4	0
Total		8	60

There is complete additivity between A and B with these values.

If we changed the data we would get something quite different. As an example we might have

	b_0	b_1	b_2	
a_0	1	3	12	16
a_1	2	6	24	32
a_2	3	9	36	48
	6	18	72	96

$I_1=43$, $I_2=23$, $I_3=30$
$T_1=34$, $T_2=41$, $T_3=21$

Variance		DF	SS
A	linear	1	170.7
	quadratic	1	0.0
B	linear	1	726.0
	quadratic	1	98.0
AB	I_1/I_2	1	66.7
	$I_3/(I_1+I_2)$	1	2.0
	T_1/T_2	1	8.2
	$T_3/(T_1+T_2)$	1	60.5
Total		8	1132.0

In this case the breakdown of the AB interaction is purely formal.

In a 3^2-experiment we could confound components of I or of T. This would give experiments in which the interaction AB would be partially confounded, and the block size would be reduced from 9 to 3 units.

	I components confounded			T components confounded		
Blocks	I	II	III	I	II	III
	(00)	(10)	(20)	(00)	(10)	(20)
	(11)	(21)	(01)	(21)	(01)	(11)
	(22)	(02)	(12)	(12)	(22)	(02)
	I_1	I_2	I_3	T_1	T_2	T_3

We write (21) instead of a_2b_1. In the plan where I is confounded we could still evaluate the interaction AB, using the T-components, which are not confounded. In the other plan the I-components are not confounded.

In a 3^3 design we have 27 treatment combinations for the three levels of three factors A, B, C. They may be written as $a_0\ b_1\ c_2$ or shortly (012) etc. There are 8 degrees of freedom for ABC.

It is possible to subdivide the 27 treatment combinations into three groups of 9 so that the totals give components of ABC, with 2 degrees of freedom. There are four possibilities to do this, accounting for the 8 degrees of freedom for ABC.

We know that we can obtain the interaction ABC considering the interaction AB for each level of C. We can do the same for each component I and T. Performing then the operation I or T we get the possibilities I, I; I, T; T, I and T, T which will be denoted by W, X, Y and Z, as proposed by Yates (1937).

If we consider the I components for AB at three levels of C, we have

$$\begin{array}{c|ccc|ccc|ccc}
 & & c_0 & & & c_1 & & & c_2 & \\
 & b_0 & b_1 & b_2 & b_0 & b_1 & b_2 & b_0 & b_1 & b_2 \\
a_0 & 0 & . & . & . & . & 0 & . & 0 & . \\
a_1 & . & 0 & . & 0 & . & . & . & . & 0 \\
a_2 & . & . & 0 & . & 0 & . & 0 & . & . \\
 & & I_1 & & & I_2 & & & I_3 &
\end{array}$$

This we could write in a single square as

$$\begin{array}{cccc}
 & c_0 & c_1 & c_2 \\
I_1 & 0 & . & . \\
I_2 & . & 0 & . \\
I_3 & . & . & 0
\end{array} \qquad I_1$$

and if we book the component I_1 in this table, this would give W_1

$a_0\ b_0\ c_0$, $a_1\ b_1\ c_0$, $a_2\ b_2\ c_0$

$a_1\ b_0\ c_1$, $a_2\ b_1\ c_1$, $a_0\ b_2\ c_1$

$a_2\ b_0\ c_2$, $a_0\ b_1\ c_2$, $a_1\ b_2\ c_2$

which could also be written as

$$\begin{array}{cccc}
 & b_0 & b_1 & b_2 \\
a_0 & 0 & 2 & 1 \\
a_1 & 1 & 0 & 2 \\
a_2 & 2 & 1 & 0
\end{array} \qquad W_1$$

where 0, 1, 2 are the levels of C. This would be W_1.

If we apply to the table

$$\begin{array}{cccc}
 & c_0 & c_1 & c_2 \\
I_1 & . & 0 & . \\
I_2 & 0 & . & . \\
I_3 & . & . & 0
\end{array} \qquad T_2$$

the operation T_2, this would yield

$$\begin{array}{lll} a_o\ b_o\ c_1, & a_1\ b_1\ c_1, & a_2\ b_2\ c_1 \\ a_1\ b_o\ c_o, & a_2\ b_1\ c_o, & a_o\ b_2\ c_o \\ a_2\ b_o\ c_2, & a_o\ b_1\ c_2, & a_1\ b_2\ c_2 \end{array}$$

or, more concisely

$$\begin{array}{cccc} & b_o & b_1 & b_2 \\ a_o & 1 & 2 & 0 \\ a_1 & 0 & 1 & 2 \\ a_2 & 2 & 0 & 1 \end{array} \qquad X_2$$

The twelve components can be summarized in the following tables.

Levels of a	Levels of b	W_1	W_2	W_3	X_1	X_2	X_3	Y_1	Y_2	Y_3	Z_1	Z_2	Z_3
		Levels of C											
0	0	0	2	1	0	1	2	0	2	1	0	1	2
1	0	1	0	2	2	0	1	1	0	2	2	0	1
2	0	2	1	0	1	2	0	2	1	0	1	2	0
0	1	2	1	0	1	2	0	1	0	2	2	0	1
1	1	0	2	1	0	1	2	2	1	0	1	2	0
2	1	1	0	2	2	0	1	0	2	1	0	1	2
0	2	1	0	2	2	0	1	2	1	0	1	2	0
1	2	2	1	0	1	2	0	0	2	1	0	1	2
2	2	0	2	1	0	1	2	1	0	2	2	0	1

Every combination of each pair of factors occurs in each set of 9 treatment combinations. Consequently if these sets are confounded with blocks, the main effects and two-factor interactions will be unconfounded. It is thus possible to confound part of the interaction between three factors with blocks. The part confounded will be either W, X, Y or Z. The following example serves to illustrate the plan and the analysis.

Experiment on sugar-beet

Date of sowing		Distance of rows		Sulfate of Ammonium	
d_0	18 April	s_0	10 in.	n_0	none
d_1	9 May	s_1	15 in.	n_1	.3 cwt N/acre
d_2	25 May	s_2	20 in.	n_2	.6 cwt N/acre

We are writing for $d_2\ s_1\ n_2$, simply 212.

It is important to adhere to a certain routine in the analysis of this design. The example is from F. Yates, 1937, p. 43, as is also the analysis. There are two replications in the example. In the first replicate the Y component of the interaction DSN is confounded with blocks, in the second replication it is the Z-component.

Plan and yield of sugar (cwt per acre)

	Replicate 1				Replicate 2		
	212	52.2	I	IV	220	31.3	
	111	52.7			100	36.4	
	100	47.8			001	39.4	
Y_2	220	35.2			211	35.3	Z_2
425.8	122	45.4			202	29.9	305.5
	010	44.6			121	34.4	
	201	46.0			010	33.3	
	021	51.4			112	33.6	
	002	50.5			022	31.9	
	011	49.7	II	V	002	33.6	
	000	47.8			122	31.4	
Y_1	202	44.1			020	25.7	Z_3
439.9	101	52.5			110	33.0	314.3
	112	49.3			011	36.6	
	221	46.2			212	41.4	
	210	47.1			221	37.6	
	120	47.2			200	33.2	
	022	56.0			101	41.8	
	012	50.9	III	VI	120	32.4	
	211	38.2			222	37.7	
Y_3	102	43.0			000	39.4	Z_1
359.0	121	36.5			102	43.1	317.8
	200	38.0			021	34.9	
	001	45.7			111	34.2	
	110	37.1			012	36.0	
	222	34.2			210	33.5	
	020	35.4			201	26.6	
1,224.7							937.6

2,162.3

Steps in the analysis:

(1) Determine to which group W, X, Y, Z the blocks belong, and which is its number (x_1, x_2, x_3). If this is already given, check if correct. Y in the replication 1, Z in the replication 2.

(2) Calculate the table of 27 combinations:

	n_o			n_1			n_2		
	s_o	s_1	s_2	s_o	s_1	s_2	s_o	s_1	s_2
d_o	87.2	77.9	61.1	85.1	86.3	86.3	84.1	86.9	87.9
d_1	84.2	70.1	79.6	94.3	86.9	70.9	86.1	82.9	76.8
d_2	71.2	80.6	66.5	72.6	73.5	83.8	74.0	93.6	71.9

(3) Calculate total SS [= 2,978.2], SS (blocks) [= 1,950.4] , SS (treatment combinations) [= 905.1]. At the same time check totals for blocks, replications and treatment combinations.

(4) Calculate the 3x3 tables for sd, nd, ns and nI and nT.

	n_o	n_1	n_2		s_o	s_1	s_2	
d_o	226.2	257.7	258.9		256.4	251.1	235.3	742.8
d_1	233.9	252.1	245.8		264.6	239.9	227.3	731.8
d_2	218.3	229.9	239.5		217.8	247.7	222.2	687.7
s_o	242.6	252.0	244.2	738.8				
s_1	228.6	246.7	263.4	738.7				
s_2	207.2	241.0	236.6	684.8				
I_1	223.8	255.8	238.9					
I_2	225.9	254.1	267.6					
I_3	228.7	229.8	237.7					
T_1	247.4	229.5	254.5					
T_2	228.6	264.4	244.9					
T_3	202.4	245.8	244.8					
	678.4	739.7	744.2	2,162.3				

(5) From the two bottom tables in (4) we calculate I and T which yields W, X, Y, Z, indistinctive of confounding.

	1	2	3	Total
W	715.6	694.6	752.1	2,162.3
X	721.2	719.4	721.7	2,162.3
Y	756.6	728.9	676.8	2,162.3
Blocks	439.9	425.8	359.0	1,224.7
Y'	316.7	303.1	317.8	937.6
Z	738.1	702.9	721.3	2,162.3
Blocks	317.8	305.5	314.3	937.6
Z'	420.3	397.4	407.0	1,224.7

(6) From the three upper tables in (4) we calculate the SS in the following tables:

Variance		SS		
D_1 S_1	D·S	314.517		
D_1 N_1	D·N	275.127		
S_1 N_1	S·N	329.770	946.414	
D		94.478		
S		107.800		
N		150.140	352.418	593.996
DSN (not confounded, WX)			94.227	
(partly confounded, YZ)			216.844	311.071
Treatments				905.067

The treatment SS should check with that in (3).

(7) In (3) should be entered the block totals, taking care to entering them in the right order. By subtraction we get Y' and Z'. These are the values of Y and Z for those blocks, in which they are not confounded. As a check they could be calculated directly from these blocks.

	From replicate 1				From replicate 2		
Z_1	000	47.8		Y_1	000	39.4	
	102	43.0			101	41.8	
	201	46.0			202	29.9	
	012	50.9			011	36.6	
	111	52.7			112	33.6	
	210	47.1			210	33.5	
	021	51.4			022	31.9	
	120	47.2			120	32.4	
	222	34.2	420.3		221	37.6	316.7
Z_2	001	45.7		Y_2	002	33.6	
	100	47.8			100	36.4	
	202	44.1			201	26.6	
	010	44.6			010	33.3	
	112	49.3			111	34.2	
	211	38.2			212	41.4	
	022	56.0			021	34.9	
	121	36.5			122	31.4	
	220	35.2	397.4		220	31.3	303.1
Z_3	002	50.5		Y_3	001	39.4	
	101	52.5			102	43.1	
	200	38.0			200	33.2	
	011	49.7			012	36.0	
	110	37.1			110	33.0	
	212	52.2			211	35.3	
	020	35.4			020	25.7	
	122	45.4			121	34.4	
	221	46.2	407.0		222	37.7	317.8
			1,224.7				937.6

(8)

$$SS(Y') = (316.7^2 + 303.1^2 + 317.8^2)/ - 937.6^2/27 = 14.899$$

$$SS(Z') = (420.3^2 + 397.4^2 + 407.0^2)/ - 1,224.7^2/27 = \underline{29.387}$$

$$SS(Y') + SS(Z') = \underline{\underline{44.286}}$$

(9) The analysis of variance can be set up. We may split the SS for each factor into the linear and the quadratic component.

n_2	744.2	d_2	687.7	s_2	684.8
n_o	678.4	d_o	742.8	s_o	738.8
$n_2 - n_o$	+65.8	D_ℓ	-55.1	D_ℓ	-54.0
$2n_1-(n_o+n_2)$	+56.8	D_q	+33.1	D_q	+53.8

Sum of squares:

$$
\begin{aligned}
SS(N_\ell) &= (+65.8)^2/36 = 120.268 \\
SS(N_q) &= (+56.8)^2/108 = \underline{\ \ 29.873} \\
\text{Sum} & \qquad\qquad\qquad\qquad 150.141
\end{aligned}
$$

In this way we get:

Variance	D.F.	SS	MS
Blocks	5	1950.4	390.1
N_ℓ	1	120.3*	
N_q	1	29.9	
D_ℓ	1	84.3*	
D_q	1	10.1	
S_ℓ	1	81.0*	
S_q	1	26.8	
ND	4	30.5	7.6
NS	4	71.8	18.0
DS	4	139.2	34.8
NDS unconfounded	4	94.2	17.3
" confounded	4	44.3	
Error	22	295.3	13.4
Total	53	2978.2	...

$$n_1 = 1, \quad n_2 = 22: \quad F_{0.05} = 4.30$$
$$n_1 = 4, \quad n_2 = 22: \quad F_{0.05} = 2.82$$

It is possible to calculate the 4 components $N_\ell D_\ell$, $N_\ell D_q$, $N_q D_\ell$, $N_q D_q$. The analysis shows, that none of these interaction components is significant. The interaction between three factors is not significant, nor are any of the interactions between two factors. The gain in precision through partial confounding can be evaluated again by considering what would happen in a uniformity trial. We set out the actual analysis with partial confounding and the corresponding analysis without confounding.

Variance	With confounding DF	SS	MS	Without confounding DF	SS	MS
Replications	1	1526.4	...	...	...	...
Blocks within replications	4	424.0	106.0	4	424.0	...
Between blocks	5	1950.4	...	...	...	...
Treatments	26	...	...	...		
Error	22		13.4	48	644.3	
Total	53	...		52	1068.3	20.5

The relative efficiency is measured roughly as

$$20.5/13.4 = 1.53 .$$

The gain in precision due to partial confounding is roughly equal to 53% .

A design which is particularly interesting consists in confounding NPK partially with blocks and using one single replication, the error variance being formed by the unconfounded parts of NPK, by $N_\ell P_q$, $N_\ell K_q$, $P_\ell K_q$, $N_q P_\ell$, $N_q K_\ell$, $P_q K_\ell$, $N_q P_q$, $N_q K_q$ and $P_q K_q$ (= 15 DF). The following example is given by Yates (1937).

Plan and yields of sugar-beet

Block I		Block II		Block III	
211	675	121	699	202	289
120	572	220	617	020	193
200	617	022	511	210	454
002	503	110	352	111	368
010	320	212	481	001	26
021	352	201	167	122	252
101	395	102	121	221	449
112	462	011	53	012	125
222	534	000	89	100	206
X_3		X_1		X_2	

Totals:

Block	I	4,430
Block	II	3,090
Block	III	2,362
		9,882

Analysis

(1) Total SS. = 1,005,712

(2) Set out three-way tables as follows:

	k_o			k_1			k_2		
	p_o	p_1	p_2	p_o	p_1	p_2	p_o	p_1	p_2
n_o	89	320	193	26	53	352	503	125	511
n_1	206	352	572	395	368	699	121	462	252
n_2	617	454	617	167	675	449	289	481	534

(4) From there we form the two-way tables, and those with I and T

	k_o	k_1	k_2		p_o	p_1	p_2	Total
n_o	602	431	1139		618	498	1056	2172
n_1	1130	1462	835		722	1182	1523	3427
n_2	1688	1291	1304		1073	1610	1600	4283
p_o	912	588	913	2413				
p_1	1126	1096	1068	3290				
p_2	1382	1500	1297	4179				
I_1	1058	843	1499					
I_2	853	1422	1113					
I_3	1509	919	666					
T_1	1115	1400	1236					
T_2	1143	897	780					
T_3	1162	887	1262					
Total	3420	3184	3278	9882				

(5) From the two bottom tables (I, T) we get:

	1	2	3	
W	3146	3271	3465	9882
X	3090	2362	4430	
Blocks	3090	2362	4430	
	(II)	(III)	(I)	
Y	3274	3266	3342	
Z	2782	3805	3295	

(6) Calculate SS from tables in (4) and (5). This gives

Variance	SS
N, P, N·P	478,664
N, K, N·K	441,020
P, K, PK	204,156
W	5,741
Y	387
Z	58,140
X (blocks)	244,526

(7) Calculate N_ℓ, P_ℓ, K_ℓ and N_q, P_q, K_q. The following routine is recommended:

	p_o	p_1	p_2	Total		
n_2-n_o	+455	+1112	+544	$+2111 = N_\ell$	$N_\ell P_\ell = +89$	$N_\ell P_q = +1225$
$2n\ (n_o+n_2)$	-247	+256	+390	$+399 = N_q$	$N_q P_\ell = +637$	$N_q P_q = +369$

In the same way we get the other interaction components:

Main effects:	linear	quadratic
N	2111	+399
P	1766	-12
K	-142	-330

Interactions:	linear · linear
NP	+ 89
NK	-921
PK	- 86

(8)

Analysis of variance

Variance	DF	SS	DQ	F	$F_{0.01}$	$F_{0.05}$
Blocks (X)	2	244,526	122,263	...	6.1	3.7
N_ℓ	1	247,573		14.2	8.7	4.5
N_q	1	2,948		...		
P_ℓ	1	173,264		9.9	8.7	4.5
P_q	1	3		...	...	...
K_ℓ	1	1,120		...	...	...
K_q	1	2,017		...	...	...
$N_\ell P_\ell$	1	660		...	...	...
$N_\ell K_\ell$	1	70,687		4.0	8.7	4.5
$P_\ell K_\ell$	1	616		...	...	...
Error	15	262,298	17,487	...	...	...
Total	26	1,005,712	...	...	...	

(9) It may be verified whether the assumption was correct, that only linear components of interactions are of any importance.

	Variance	DF	SS	MS
	$N_\ell P_q$	1	41684	
	$N_\ell K_q$	1	6110	
	$N_q P_\ell$	1	11271	
	$N_q K_\ell$	1	15335	
	$P_\ell K_q$	1	26136	
	$P_q K_\ell$	1	28	
	$P_q P_q$	1	1261	
	$N_q K_q$	1	95230	
	$P_q K_q$	1	972	
NPQ	Y	2	388	194
	W	2	5742	2871
	Z	2	58141	29070°
	Error	15	262298	...

In this case also, we may evaluate the gain in precision obtained through confounding. The analysis of variance for both confounded and not confounded design is set out below:

Variance	With confounding		Without confounding		
	DF	MS	DF	SS	MS
Blocks	2	122263	2	244526	...
Treatments	9	...			
Error	15	17487	24	419688	...
Total	26		26	664214	25547

The increase in precision is: $25547/17487 = 1.461$. The partial confounding used in this example resulted in a gain in precision of roughly 46 % (not taking into account the different degrees of freedom for the two designs). This efficiency calculation is only valid if we assume that the confounded interactions are not significant.

Further reduction in block size in the 2^n design.

We have seen that we can confound one interaction in a 2^3 design in order to have 4 plots per block instead of 8. If we confound the same interaction in each replication, we have complete confounding, if we confound different interactions in different replications, we have partial confounding.

In the 3^3 design we can confound part of the three-factor-interaction either completely or partially and thus reduce the block size from 27 to 9.

It is possible to reduce the block size further, if we confound not only one, but two (or more) interactions in one replication. We are investigating this in some detail for the 2^4 design.

The first object of our investigation are the relationships which exist between the 15 main effects and interactions of one replication in a 2^4 design with factors A, B, C, D. Let us set out the 16 treatment combinations and the 15 main effects and interactions. In a 2^4 design all effects and interactions are obtained by using factors +1 or -1; this permits us to write in the following table simply + or - instead of +1 or -1.

Treatment combinations	Main effects and interactions														
	A	B	AB	C	AC	BC	ABC	D	AD	BD	ABD	CD	ACD	BCD	ABCD
(1)	-	-	+	-	+	+	-	-	+	+	-	+	-	-	+
a	+	-	-	-	-	+	+	-	-	+	+	+	+	-	-
b	-	+	-	-	+	-	+	-	+	-	+	+	-	+	-
ab	+	+	+	-	-	-	-	-	-	-	-	+	+	+	+
c	-	-	+	+	-	-	+	-	+	+	-	-	+	+	-
ac	+	-	-	+	+	-	-	-	-	+	+	-	-	+	+
bc	-	+	-	+	-	+	-	-	+	-	+	-	+	-	+
abc	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-
d	-	-	+	-	+	+	-	+	-	-	+	-	+	+	-
ad	+	-	-	-	-	+	+	+	+	-	-	-	-	+	+
bd	-	+	-	-	+	-	+	+	-	+	-	-	+	-	+
abd	+	+	+	-	-	-	-	+	+	+	+	-	-	-	-
cd	-	-	+	+	-	-	+	+	-	-	+	+	-	-	+
acd	+	-	-	+	+	-	-	+	+	-	-	+	+	-	-
bcd	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-
abcd	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Total	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

We may find the following relations between the main effects and interactions; writing A·B to mean that we formally multiply effect A by effect B. If we consider the products of the coefficients for the same treatment combination, we find that the product of the coefficients for A with those for B gives the coefficients for the interaction AB. This is true for all products AB, AC, ... CD.

If we form (AB)·C in the same way, we get the coefficients for ABC. The same can be verified for all other three factor interactions.

Finally we find that (ABC)D = ABCD, which we may also get as A(BCD) = (AB)(CD) = (AC)(BD) etc. We may now ask what happens, if we multiply, say, (AB) with (BCD)?

Formally, this would give AB^2CD. Let us see, whether this makes any sense, i.e. whether we find any main effect or interaction corresponding to this product. We obtain

(1)	a	b	ab	c	ac	bc	abc	d	ad	bd	abd	cd	acd	bcd	abcd
-	+	-	+	+	-	+	-	+	-	+	-	-	+	-	+

and this is equal to ACD. So we could write

$$(AB)\ (BCD) = AB^2CD = ACD \ .$$

Let us try one other example, namely

$$(ABC)\ (BCD) = AB^2\ C^2\ D = AD.$$

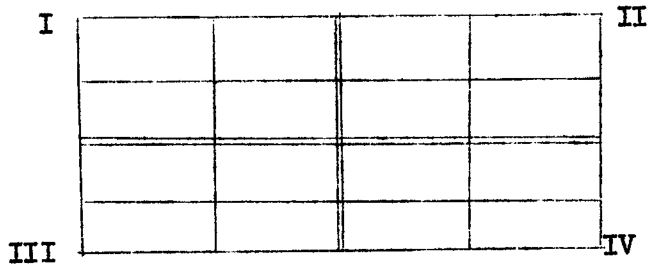
In the same way one may find

$$(ABC)\ (BC) = AB^2\ C^2 = A$$

and from this we find the simple rule that $A^2 = B^2 = C^2 = D^2 = 1$ symbolically.

In the 2^n series, any interaction between main effects or interactions can be written down immediately following these rules.

What happens, if we try to confound two interactions in one replication of $2^4 = 16$ treatment combinations, in view of getting blocks of 4 plots?



There are 3 degrees of freedom between the 4 blocks, which we may assign as follows

Block	Degree of freedom (1)	(2)	(3)
I	+	+	+
II	+	-	-
III	-	-	+
IV	-	+	-

These 3 degrees of freedom are part of the 15 independent comparisons between the 16 treatment combinations if we confound block differences with any two main effects of interactions. If the interactions are confounded with block differenees, then a third interaction (or main effect) will also be confounded with differences between blocks, because for the three comparisons we have (1)(2) = (3), (1)(3) = (2) and (2)(3) = (1).

If we confound for instance ABC with (1) and (ABCD) with (2), then D will be confounded with (3). If we confound (ABC) with (1), and (BCD) with (2), then AD will be confounded with (3).

The following data were obtained from an experiment with beans (Rothamsted Experimental Station Report 1937, p. 162). The factors and levels are

K: 0 and 1.0 cwt K_2O per acre ;

P: 0 and 0.6 cwt P_2O_5 per acre ;

N: 0 and 0.4 cwt N per acre ;

D: 0 and 10 tons of dung per acre.

Plan and yields in lb. on plots of 1/40 acre

									Block
Replication 1	p	56.3	kn	70.8	pnd	78.5	kd	94.3	I
	knd	85.4	k	68.1	pn	72.3	pd	90.7	II
	kp	90.1	kpnd	96.3	n	79.1	d	78.5	III
	kpd	98.1	kpn	77.7	(1)	63.7	nd	82.3	IV
Replication 2	kpd	94.1	pnd	86.4	kn	90.1	(1)	89.0	V
	kpnd	119.5	k	78.0	pd	84.5	n	92.1	VI
	p	109.8	kpn	89.2	nd	73.7	kd	72.5	VII
	kp	72.2	d	50.9	pn	56.7	knd	57.6	VIII

There must be three interactions (or main effects) confounded with blocks in each replication. Let us try to find out which are these interactions. If we first consider (1): I + II - III - IV, and consult the table on p. 93, we see that KP is confounded with (1). If we consider (2): I - II - III + IV, we see that KND is confounded. And finally we find that for (3): I - II + III - IV the interaction PND is confounded.

The interactions confounded in the second replication are for

(4) V + VI - VII - VIII PD

(5) V - VI - VII + VIII KPN

(6) V - VI + VII - VIII KND

From this it follows that all the main effects K, P, N, D are unconfounded, as are also the interactions

KN, KD, PN, ND, KPD and KPND .

The two-factor interactions KP and PD and the three-factor interactions KPN and PND are partially confounded, while KND is completely confounded.

The analysis of this experiment is performed in the usual way. There are 14 DF for the 14 unconfounded or partially confounded main effects and interactions, and 7 DF between blocks. This leaves 10 DF for error, as the total DF is 31. The main effects and interactions are found according to Yates' method, carrying it one step further than in a 2^3-design. To find the confounded interactions, we have to apply corrections. If we get by Yates' procedure a certain value for, say KP then we have to subtract from this (III + IV - I - II), this being the confounded part of KP. The sum of square for KP is obtained by squaring the corrected interaction and dividing by 16, because KP is estimated from replication 2 only.

It is instructive to calculate the sums of squares for each of the 10 degrees of freedom for error separately. This, as also the complete analysis is left as an exercise for the student.

4.4 Fractional factorial designs.

In factorial experiments the main effects and the interactions of two factors are mostly of interest. Higher order interactions are seldom significant. On the other hand factorial experiments are mostly used to obtain information on the overall effect of factors, preferably several factors at the same time. It may be argued, that we would often like to include more factors, even if this would imply that we would not be able to obtain higher order interactions.

Let us consider a simple case and see what happens if we try to introduce a new factor at the expense of loosing information on a certain interaction between several factors. Suppose we have a 2^3-factorial experiments with factors A, B and C. Let us assume that the interaction ABC is unimportant. In this case the contrast

$$ABC = a + b + c + abc - (1) - ab - ac - bc$$

is only part of the error fluctuations. In such a case we may be tempted to introduce a new factor D in such a way as to assign to the treatment combinations

a, b, c and abc the higher level of D

and to

(1), ab, ac and bc the lower level of D

The main effect of D corresponds in that case to the interaction ABC, or more precisely, D is confounded with ABC.

We may write

$$ABC = D$$

it being understood that the sign = means "is confounded with". From this we infer that

$$AD = A^2BC = BC$$

and similarly $BD = AC$ and $CD = AB$.

Consequently if we confound D with ABC, we will not be able to disentangle the two-factor interactions from each other. Such a design would be of interest only if we were solely interested in the main effects. For the case mentioned we would have only 8 treatment combinations, namely

ad, bd, cd, abcd
(1), ab, ac, bc

for an experiment with 4 factors. We would in fact have a half-replication, the combinations missing being

a, b, c, d, abc, acd, abd, bcd .

Note that the contrast, of which we have only taken one half of the treatment combinations, is ABCD. This is called the defining contrast. If we form the interaction between any main effect or interaction and the defining contrast, we find the effect confounded with it. E.g.

$$A\cdot(ABCD) = BCD ,$$

that is, A is confounded with BCD. Or:

$$(AB)(ABCD) = CD ;$$

which means that AB is confounded with CD.

A and BCD

AB and CD etc.

are called "aliases".

The following example will show how the analysis of a fractional factorial experiment is carried out (Linder, 1953, p. 127).

Effect of 6 factors on the optical rotation of a cocoa preparation

Duration of dilution process	A	20 and 40 minutes;
Temperature of dilution process	B	50-55 C and 80-85 C;
Temperature of clarifier	C	25-30 C and 50-55 C;
Quantity of clarifier	D	2 and 4 ml Somogy II per 10 g chocolate
Concentration of chocolate	E	5 and 10 g chocolate per 100 ml
p_H	F	7.0 and 8.1

Treatment combinations and optical rotation (x 1000)

Treatment combination	Value	Treatment combination	Value
(1)	28	af	30
ab	18	bf	16
ac	17	cf	20
bc	14	abcf	12
ad	14	df	35
bd	16	abdf	14
cd	18	acdf	15
abcd	24	bcdf	25
ae	70	ef	68
be	54	abcf	26
ce	54	acef	59
abce	30	bcef	12
de	48	adef	32
abde	54	bdef	40
acde	33	cdef	28
bcde	38	abcdef	32

The treatment combinations were carried out in a random order. There are 6 factors, and 32 treatment combinations, while $2^6 = 64$. The fractional factorial experiment is a half-replication. The defining contrast is

$$A\ B\ C\ D\ E\ F\ ,$$

from which it follows, that we have the following aliases

A = BCDEF	AB = CDEF	ABC = DEF	BCD = AEF
B = ACDEF	AC = BDEF	ABD = CEF	BCE = ADF
C = ABDEF	AD = BCEF	ABE = CDF	BDE = ACF
D = ABCEF		ACD = BEF	CDE = ABF
E = ABCDF		ACE = BDF	
F = ABCDE	EF = ABCD		

From this it may be seen that we are able to ascertain all main effects and all two factor interactions under the assumption that four-and five-factor interactions are unimportant.

If we assume furthermore that there are no real three-factor-interactions, we may consider these 10 contrasts as peaking up the experimental error. The analysis of variance would then be constructed as follows:

Variance	DF
Main effects	6
Two factor interactions	15
Error	10
Total	31

Yates' method of determining the effects breaks down in fractional factorial designs. The simplest method of determining the effects is by using the definitions and calculate them directly. The main effect A is obtained by subtracting all treatment combinations without a from those containing a .

Treatment combinations	Sum of values	
With a	480	
Without a	514	
Sum	994	
Difference	-34	(= A)

The sum (994) is calculated as a check.

Interaction AB:

Treatment combinations	Sum of values	
With a and b, or without a and b	509	
With a or b	485	
Sum	994	
Difference	+24	(= AB)

The analysis of variance is obtained in the usual way. First we calculate the total SS.

$$SS(\text{total}) = 28^2 + 18^2 + \ldots + 32^2 - 994^2/32 = 8{,}781.9$$

For each effect we get the SS with one degree of freedom as usual

$$SS(A) = (-34)^2/32 = 36.1$$
$$SS(AB) = (+24)^2/32 = 18.0$$

For the error SS we get by subtraction 819.0; the error MS is thus 81.9. The SS for error can also be obtained as the sum of each degree of freedom.

$$SS(ABC) = (-14)^2/32 = 6.1$$

and so on. By this we get

$$SS(BDE) = SS(ACF) = (+120)^2/32 = 450.0.$$

If we threw this out of the error, we would get

$$SS(\text{error}) = 369.0, \quad DF = 9, \quad MS(\text{error}) = 41.0 ,$$

instead of 81.9.

Whether the first or second MS is to be used for error, must be ascertained taking into account all technical knowledge on the process in question.

A = -34, B = -144*, C = -132*, D = -62, E = +362**, F = -66

BD = +184**, BDE = ACF = +120 [significant with MS(error) = 41.0].

5. Determination of optimum conditions.

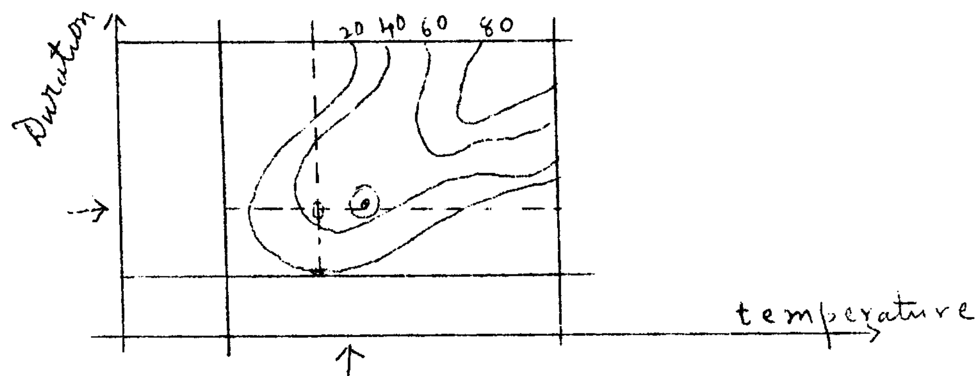
In the usual factorial experiments we are interested mainly in knowing whether the factors chosen have any effect over the range we are interested in, and also whether there are any interactions between the factors. In some cases we are also interested in knowing if the factor-response relation is linear or not in the range within which the levels of this factor are set.

In some applications, however, we are more interested to know at what levels of the factors the response is at an optimum (not necessarily a maximum). There again the "one-factor-at-a-time" approach is not suitable.

In order to discuss the different features of the problem of determining optimum conditions, let us consider a specific example, namely the influence of duration and temperature of the reaction on the yield of a chemical process.

If the chemical and physical theory of such a process were sufficiently well developed, there would be no need for experiments at all, or perhaps then only to confirm the theory. On the other hand, if the theory is not fully known, it will often be sufficient to have a rough approximation of the response surface, which is the surface of yield as a function of duration and temperature of the chemical process.

Let us first see how we would proceed using the one-factor-at-a-time approach.

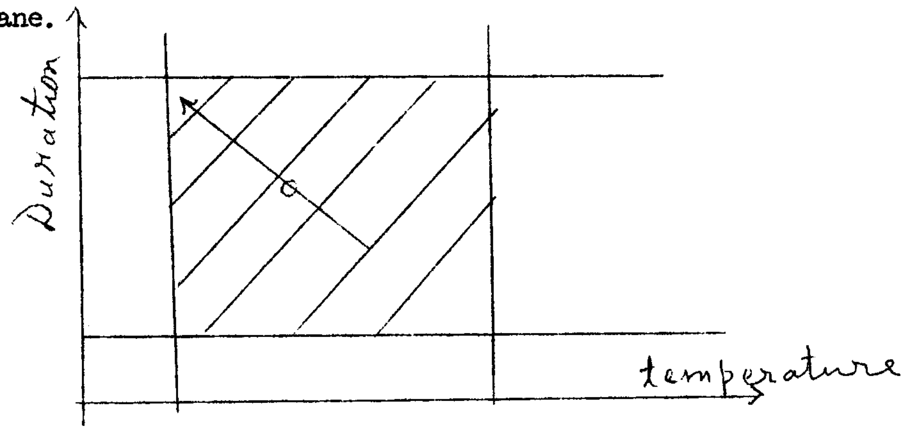


Let us assume that we know the surface of yield as a function of duration and temperature.

The contours drawn are the lines of equal yield. Let us start at a given temperature and explore the response to duration at this temperature. This would give the optimum duration. Then studying the influence of temperature for this "optimum" duration, we would get the "optimum" temperature; but as the example shows, these results would be very misleading, the point obtained being very far from the real optimum.

G.E.P. Box, working at the ICI (Imperial Chemical Industries) put forward a method in 1951, generally known as the Box-Wilson-method (Box and Wilson, 1951). His approach consists essentially of two steps.

Step 1. This is intended to lead quickly into the region of the optimum. It consists simply in making a few experiments, and to fit to the values thus obtained, a plane.



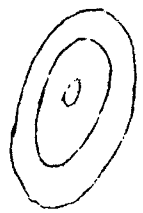
In the range of temperatures and duration which we have selected, we choose a point and "around" that we select a few values of temperature and duration, perform the necessary experiments and obtain the yields. We then fit a plane, and this gives us the direction of steepest ascent on that plane (not necessarily on the surface of yields).

If by chance the starting point is already near the optimum, then the plane will be parallel or nearly parallel to the area of temperature and duration.

Step 2. Having reached through step 1 a region near the optimum, the second step consists in exploring the configuration of the response surface near the optimum. The reason for this being that it is sometimes more economical to run a process near the optimum instead of at the optimum point, owing to technical difficulties.

The simplest way to explore the response surface statistically, is to fit a second degree surface to the observed points. According to the type of surface near the optimum we may have

(a)

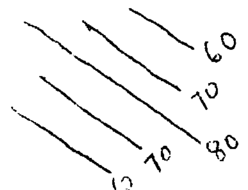


Maximum
Ellypses

(b)

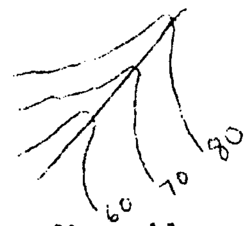
Saddle point
Hyperbolae

(c)



Ridge Parallel
straight lines

(d)



Ascending ridge
Parabolas

Statistical tools.

In both step 1 and step 2 we have to use linear regression analysis in order to fit a plane or a surface of the second degree to the points observed. Let us shortly state the essential steps in linear regression analysis.

We assume we have p independent variates

$$x_1, x_2, \dots x_p$$

and one dependent variate y.

As usual the so called "independent" variates may be correlated or even functionally related to each other, we could for instance well have

$x_5 = x_1 \cdot x_2$.

The population regression may be written as

$$\eta = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p \ .$$

We assume that the $y - \eta$ are normally distributed with mean zero and variance σ^2, and that σ^2 is the same for all sets of values of $x_1, x_2, \dots x_p$. The case $p=3$ will be sufficient to state our notations. We assume that we have N observed values of x_1, x_2, x_3 and y, the i-th being

$x_{1i}, x_{2i}, x_{3i}, y_i$.

Estimates b_o, b_1, b_2, b_3 are obtained from $\underset{i}{S}(y_i - \eta_i)^2 = \text{min.}$

This gives

$$b_o = \bar{y} - b_1 \bar{x}_1 - b_2 \bar{x}_2 - b_3 \bar{x}_3$$

and

$$b_1 S_{11} + b_2 S_{12} + b_3 S_{13} = S_{1y}$$

$$b_1 S_{12} + b_2 S_{12} + b_3 S_{23} = S_{2y}$$

$$b_1 S_{13} + b_2 S_{23} + b_3 S_{33} = S_{3y}$$

where

$$S_{11} = \underset{i}{S} (x_{1i} - \bar{x}_1)^2, \; S_{12} = \underset{i}{S} (x_{1i} - \bar{x}_1)(x_{2i} - \bar{x}_2) \ldots$$

$$S_{1y} = \underset{i}{S} (x_{1i} - \bar{x}_1)(y_i - \bar{y}), \ldots$$

And this will give

$$b_1 = \frac{1}{\Delta} \begin{vmatrix} S_{14} & S_{12} & S_{13} \\ S_{24} & S_{22} & S_{23} \\ S_{34} & S_{23} & S_{33} \end{vmatrix} \; ; \quad \Delta = \begin{vmatrix} S_{11} & S_{12} & S_{13} \\ S_{12} & S_{22} & S_{23} \\ S_{13} & S_{23} & S_{33} \end{vmatrix}$$

From this we get

$$b_1 = \frac{1}{\Delta} \left\{ S_{14} \begin{vmatrix} S_{22} & S_{23} \\ S_{23} & S_{33} \end{vmatrix} - S_{24} \begin{vmatrix} S_{12} & S_{13} \\ S_{23} & S_{33} \end{vmatrix} + S_{34} \begin{vmatrix} S_{12} & S_{13} \\ S_{22} & S_{23} \end{vmatrix} \right\}$$

or, if we define

$$C_{11} = \frac{1}{\Delta} \begin{vmatrix} S_{22} & S_{23} \\ S_{23} & S_{33} \end{vmatrix} , \quad C_{12} = -\frac{1}{\Delta} \begin{vmatrix} S_{12} & S_{13} \\ S_{23} & S_{33} \end{vmatrix} , \quad C_{13} = \frac{1}{\Delta} \begin{vmatrix} S_{12} & S_{13} \\ S_{22} & S_{23} \end{vmatrix}$$

and similarly for C_{22}, C_{23}, C_{33}, noting also that $C_{13} = C_{31}$ etc., we have

$$b_1 = C_{11} S_{14} + C_{12} S_{24} + C_{13} S_{34}$$

$$b_2 = C_{12} S_{14} + C_{22} S_{24} + C_{23} S_{34}$$

$$b_3 = C_{13} S_{14} + C_{23} S_{24} + C_{33} S_{34}$$

The empirical or observed regression will be

$$Y = \bar{y} - b_1\bar{x}_1 - x_2\bar{x}_2 - b_3\bar{x}_3 + b_1 x_1 + b_2 x_2 + b_3 x_3 .$$

$$b_o = \bar{y} - b_1\bar{x} - b_2\bar{x}_2 - b_3\bar{x}_3$$

The estimation s^2 of σ^2 is obtained as

$$s^2 = \underset{i}{S} (y_i - Y_i)^2/(N-p-1).$$

where

$$\underset{i}{S} (y_i - Y_i)^2 = S_{44} - b_1S_{14} - b_2S_{24} - b_3S_{34} .$$

From this it may be seen that we can also write an analysis of variance:

Variance	DF	SS	MS
Regression	p	$b_1S_{1y}+\ldots+b_pS_{py}$	...
Residual	N-p-1	Difference	s^2
Total	N-1	S_{44}	...

To test the hypothesis; $\beta_1 = \beta_2 = \ldots = \beta_p = 0$, we calculate

$$F = MS\,(\text{regression})/s^2 , \quad n_1 = p,\ n_2 = N-p-1 .$$

To test the hypothesis; $\beta_j = \beta_{jo}$, we calculate

$$F = (b_j - \beta_{jo})^2/c_{jj}\, s^2 , \quad n_1 = 1,\ n_2 = N-p-1 .$$

The equations for b_1, b_2, b_3 will simplify considerably when the design of the experiment, that is the levels for each factor is chosen in an appropriate way. In particular the c_{jk} may be calculated very easily. As an example of this let us consider a 2^3-factorial design. If we choose the levels at -1 and +1, which may always be attained by suitable transformations, we have

Treatment combination	x_1	x_2	x_3	y
1	-1	-1	-1	y_1
2	-1	-1	+1	y_2
3	-1	+1	-1	y_3
4	+1	-1	-1	y_4
5	-1	+1	+1	y_5
6	+1	-1	+1	y_6
7	+1	+1	-1	y_7
8	+1	+1	+1	y_8
Total	0	0	0	Sy_i
Mean	0	0	0	$\bar{y}$

From this it follows that all the $\underset{i}{S}\, x_{ji} = 0$, and due to orthogonality of x_1, x_2, x_3 also $\underset{i}{S}\, x_{ji}\, x_{ki} = 0$; furthermore, we have $\underset{i}{S}\, x_{ji}^2 = 8$. Thus

$$S_{11} = S_{22} = S_{33} = 1$$

$$S_{12} = S_{13} = S_{23} = 0$$

and $\bar{x}_1 = \bar{x}_2 = \bar{x}_3 = 0$, so that

$$\underline{b_o} = \bar{y} - b_1\bar{x}_1 - b_2\bar{x}_2 - b_3\bar{x}_3 = \underline{\bar{y}}$$

and

$$\Delta = \begin{vmatrix} 8 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 8 \end{vmatrix} = 8^3$$

$$C_{11} = 1/8, \; C_{22} = C_{33} = 1/8; \quad C_{12} = C_{13} = C_{23} = 0$$

$$b_1 = S_{14}/8, \; b_2 = S_{24}/8, \quad b_3 = S_{24}/8 \quad .$$

or

$$b_1 = (y_4+y_6+y_7+y_8-y_1-y_2-y_3-y_5)/8$$

$$b_2 = (-y_1-y_2+y_3-y_4+y_5-y_6+y_7+y_8)/8$$

$$b_3 = (-y_1+y_2-y_3-y_4+y_5+y_6-y_7+y_8)/8$$

The calculations are thus greatly simplified if we adopt a factorial design.

As an example consider the determination of optimum conditions for a chemical process carried out by B. Messikourner (1960), CIBA, Basel (Switzerland). Preliminary investigations showed that several factors influences yield. A list was prepared, in which these factors were ordered according to their presumptive importance.

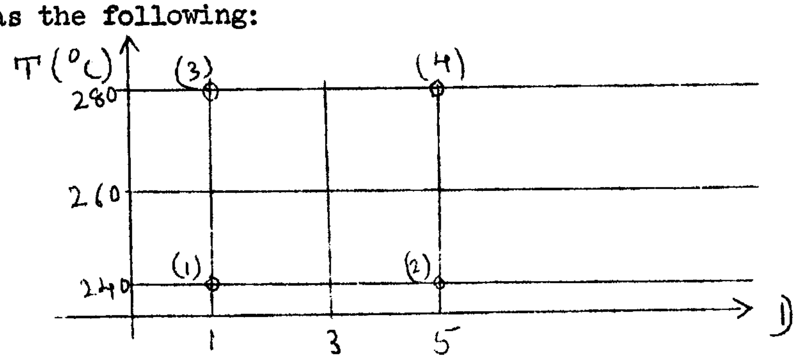
Duration of reaction ;
Temperature of reaction ;
Concentration of reacting substances ;
Velocity of stirring ;
etc. etc.

Preliminary information indicated that duration and temperature of the process were likely to be the most important factors determining the yield. It was therefore decided first to optimize for these two factors. We are here only concerned about this part of the investigations.

The starting point was a duration of 3 hours and a temperature of 260° C which had so far given the highest yield of 64 %. In what follows we will use the following notation

D = duration; T = temperature; A = yield.

Step 1. To be able to determine b_o, b_1, b_2 we have to carry out 3 determinations of the yield at 3 different points in the D-T-plane. For ease of calculation a 2^2-design was chosen symmetrical around D = 3, T = 260. The choice of the levels should be such, that the effect on A is clearly recognizable. The choice was the following:



(1) ... (4) are the numbers of the experimental units. The experiments should be carried out in a random order. There is no replication, so it is not possible to determine an experimental error. Possibly there is some preliminary knowledge on experimental errors, which may be used. Calculations are simpler if we transform T and D in such a way as to have as values for the independent variables -1 and +1.

$$x_1 = \frac{D-3}{2}, \qquad x_2 = \frac{T-260}{20}$$

The experimental results for these four points were as follows:

		D	T	A	x_1	x_2	y=A-60
	(1)	1	240	61	-1	-1	1
	(2)	5	240	66	+1	-1	6
	(3)	1	280	67	-1	+1	7
	(4)	5	280	78	+1	+1	18
Total		.12	1040	272	0	0	32
Mean		.3	260	68	0	0	8

To fit the plane

$$\eta = \beta_o + \beta_1 x_1 + \beta_2 x_2$$

we have to calculate

$$S_{11} = S_{22} = 4, \quad S_{12} = 0\ ; \quad \Delta = 16$$

$$C_{11} = C_{22} = 1/4, \quad C_{12} = 0$$

$$S_{14} = -1 + 6 - 7 + 18 = 24 - 8 = 16$$

$$S_{24} = -1 - 6 + 7 + 18 = 25 - 7 = 18$$

$$b_1 = C_{11}S_{14} + C_{12}S_{24} = 16/4 = 4.0$$

$$b_2 = C_{12}S_{14} + C_{22}S_{24} = 18/4 = 4.5$$

$$b_o = \bar{y} - b_1\bar{x}_1 - b_2\bar{x}_2 = 8.0$$

$$Y = 8.0 + 4.0\,x_1 + 4.5\,x_2$$

Calculation of standard errors for b. First determine $S(y-Y)^2$. In this case it is easy to calculate $S(y-Y)^2$ directly.

	y	Y	y-Y	$(y-Y)^2$	y^2
(1)	1	-0.5	+1.5	2.25	1
(2)	6	7.5	-1.5	2.25	36
(3)	7	8.5	-1.5	2.25	49
(4)	18	16.5	+1.5	2.25	324
Total	32	32.0	0.0	9.00	410

$$s^2 = 9.00/1 = 9.00$$

We may calculate

$$S(y-Y)^2 = S(y-\bar{y})^2 - S(Y-\bar{y})^2$$

$$S(y-y)^2 = Sy^2-\bar{y}\ Sy = 410 - 8.32 = 410-256 = 154.0$$

$$S(Y-\bar{y})^2 = b_1S_{14}+b_2S_{24}=(4.0)(16)+(4.5)(18)=64.0+81.0=145.0$$

$$S(y-Y)^2 = 154.0 - 145.0 = 9.0 \qquad \text{(check)}$$

From s^2 we get the standard errors:

	(Standard error)^2 = Error variance
$b_o = 8.0$	$s^2/N = 9.0/4 = 2.25 = (1.5)^2$
$b_1 = 4.0$	$C_{11}s^2 = 9.0/4 = 2.25 = (1.5)^2$
$b_2 = 4.5$	$C_{22}s^2 = 9.0/4 = 2.25 = (1.5)^2$

These standard errors are however of little value, based as they are on only 1 degree of freedom.

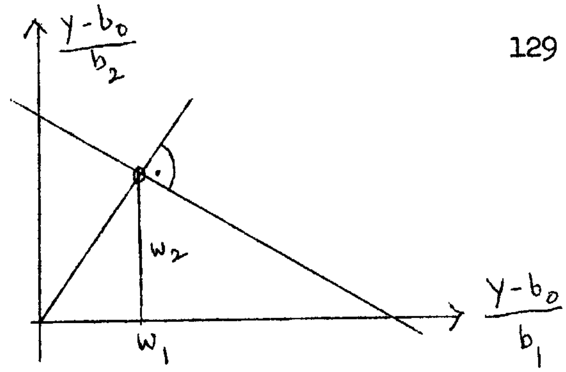
The contours for the plane

$$Y = b_o + b_1x_1 + b_2x_2$$

are easy to find. For a given value of Y we calculate the interactions with the axes x_1 and x_2.

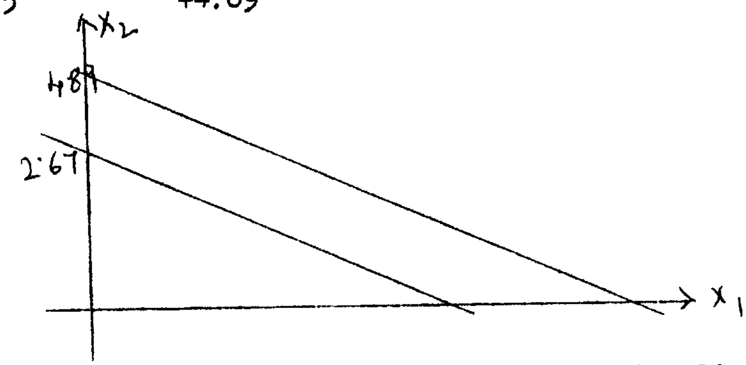
$$x_1 = 0: \quad x_2 = \frac{Y-b_o}{b_2}$$

$$x_2 = 0: \quad x_1 = \frac{Y-b_o}{b_1}$$



We find:

Y	$x_2=0{:}x_1$	$x_1=0{:}x_2$
0	-2.0	-1.78
10	+ .5	+ .44
20	+3.0	+2.67
30	+5.5	+4.89



Step 1 consists in choosing further experimental units along the line of steepest ascent. In this case we would have to go in a direction orthogonal to the contours just obtained. If w_1, w_2 are the coordinates of the point on the contours on the straight line orthogonal to the contour passing through the origin, then

$$\underline{w_1 : w_2} = \frac{Y-b_o}{b_2} : \frac{Y-b_o}{b_1} = \underline{b_1 : b_2}$$

In view of the uncertainty with which b_1 and b_2 are determined, the new points were chosen with equal coordinates x_1 and x_2 instead of in a ratio of 4.0 to 4.5, and starting from point (4), for which the yield was highest. The new points in x_1, x_2 were chosen as

no.	x_1	x_2	Y
(5)	+2	+2	25.0
(6)	+3	+3	33.5
(7)	+4	+4	42.0

From this it follows at once that the plane cannot fit the real surface of yield very closely because $Y = 42.0$ means $A = 102.0\%$, an impossible value. The values of the yield A obtained with points (5), (6) and (7) were 84, 89 and 84 respectively.

This is in itself a very remarkable result and achievement; with 7 experimental units it was possible to attain a yield of 89% , against 64% at the starting point. The duration and temperature at the yield 89% are:

$$D = 3 + 2x_1 = 3 + 2\cdot 3 = 9\text{h}$$

and

$$T = 260 + 20x_2 = 260 + 20\cdot 3 = 320^\circ\ \text{C}.$$

The question is now, whether with this point we are near the optimum value or not. Before we attempt to answer this question, we want to study the errors involved in determining b_o, b_1, b_2 somewhat more closely.

In order to do that, the same 2^2 experiment was carried out a second time, which gave the following values for y.

nr.	y
(8)	1.7
(9)	5.2
(10)	7.9
(11)	19.2

With these values and those for nr (1), (2), (3) and (4) we are in a position to determine a plane with better position; and we are also able to separate the experimental error from the lack of fit, which indicates how well (or how bad) the observations fit the plane

The calculations proceed as usual:

j	x_1	x_2	y_{ji}		$\underset{i}{S}y_{ji}$	$y_{j1}-y_{j2}$	SS	DF
1	-1	-1	1.0	1.7	2.7	-0.7	0.49/2=0.245	1
2	+1	-1	6.0	5.2	11.2	+0.8	0.64/2=0.320	1
3	-1	+1	7.0	7.9	14.9	-0.9	0.81/2=0.405	1
4	+1	+1	18.0	19.2	37.2	-1.2	1.44/2=0.720	1
Total	0	0	32.0	34.0	66.0	2.0	1.690	4

The SS between duplicate values (1.690) has 4 degrees of freedom; the error MS is 1.690/4 = .4225. Calculations to adjust a plane to the 8 points give:

$$S_{11} = S_{22} = 8, \; S_{12} = 0; \; C_{11} = C_{22} = 1/8, \; C_{12} = 0 \quad \Delta = 64 .$$

$$S_{14} = -2.7 + 11.2 - 14.9 + 37.2 = 48.4 - 17.6 = +30.8$$

$$S_{24} = -2.7 - 11.2 + 14.9 + 37.2 = 52.1 - 13.9 = +38.2$$

$$b_1 = +30.8/8 = +3.850; \quad b_2 = +38.2/8 = +4.775$$

$$b_o = \bar{y} = 66.0/8 = 8.250 .$$

$$\underline{Y = 8.250 + 3.850x_1 + 4.775x_2}$$

In the analysis of variance we have 2 degrees of freedom for regression and 5 d.f. for residual variation. This can be split up into 4 d.f. for error between duplicates and 1 d.f. for lack of fit.

Variance	D.F.	SS	MS
Regression	2	300.985	150.492
Lack of fit	1	23.805	23.805
Error	4	1.690	0.422
Total	7	326.480	...

It is instructive to see why the lack of fit is so pronounced. We only have to compute the regression values Y_i and the residuals. We get

j	Y_{ji}	$y_{ji}-Y_{ji}$		$\bar{y}_j-\bar{Y}_j$
1	- .375	+1.375	+2.075	+1.725
2	+ 7.325	-1.325	-2.125	-1.725
3	+ 9.175	-2.175	-1.275	-1.725
4	+16.875	+1.125	+2.325	+1.725

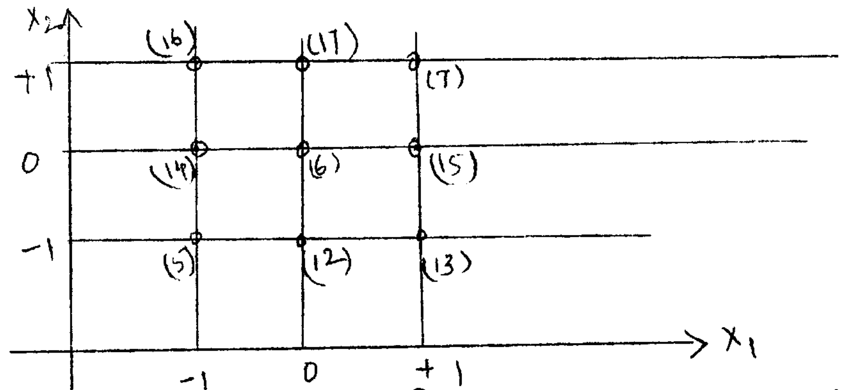
This shows that the surface of yield is higher than the plane at the points (-1, -1) and (+1, +1) and lower at the two other points. In spite of this lack of fit the plane approximation was quite useful, as the values for treatment combinations (5), (6) and (7) have shown.

Step 2.

It is required to study the surface of yield in the vicinity of the best yield attained so far, approximating this surface by a surface of the second degree. The equation of a surface of 2nd degree may be written as follows:

$$Y = b_{oo} + b_{1o}x_1 + b_{o1}x_2 + b_{20}x_1^2 + b_{11}x_1x_2 + b_{o2}x_2^2$$

This necessitates at least 6 points for the estimation of the 6 coefficients. In view of what we know, it seems appropriate to select points in the (x_1, x_2)-plane according to a 3^2 design, again transforming the (D, T)-coordinates in such a way as to get -1, 0, +1 as values of x_1 and x_2. The experimenter choose the points around the treatment combination (6), which so far gave the highest yield and retained the treatment combinations (5) and (7). The following figure shows how treatment combinations (12) to (17) were selected.



The values for x_1, x_2, x_1^2, x_1x_2 and x_2^2 are given in the following table, which also contains the values of the yields. As all the yields are higher than 80 we transform these according to the formula $y = A-80$.

Treatment combination	x_1	x_2	x_1^2	x_1x_2	x_2^2	y
(5)	-1	-1	+1	+1	+1	4
(12)	0	-1	0	0	+1	8
(13)	+1	-1	+1	-1	+1	11
(14)	-1	0	+1	0	0	5
(6)	0	0	0	0	0	9
(15)	+1	0	+1	0	0	11
(16)	-1	+1	+1	-1	+1	1
(17)	0	+1	0	0	+1	3
(7)	+1	+1	+1	+1	+1	4
Total	0	0	+6	0	+6	56

$N = 9$

$p = 5$

We may write

$$x_1 = Z_1, \; x_2 = Z_2, \; x_1^2 = Z_3, \; x_1x_2 = Z_4, \; x_2^2 = Z_5$$

and for the moment being write

$$Y = b_o + b_1Z_1 + b_2Z_2 + \ldots + b_5Z_5 \; .$$

This will allow us to use the formulae given earlier. We get

$$\bar{Z}_1 = \bar{Z}_2 = \bar{Z}_4 = 0, \qquad \bar{Z}_3 = \bar{Z}_5 = 2/3$$

$$\begin{array}{lllll}
S_{11} = 6 & S_{12} = 0 & S_{13} = 0 & S_{14} = 0 & S_{15} = 0 \\
 & S_{22} = 6 & S_{23} = 0 & S_{24} = 0 & S_{25} = 0 \\
 & & S_{33} = 2 & S_{34} = 0 & S_{35} = 0 \\
 & & & S_{44} = 4 & S_{45} = 0 \\
 & & & & S_{55} = 2
\end{array}$$

$$\begin{array}{lll}
S_{14} = 26 - 10 & & = +16 \\
S_{24} = \;\; 8 - 23 & & = -15 \\
S_{34} = 36 & -(2/3)56 & = -4/3 \\
S_{44} = \;\; 8 - 12 & & = -4 \\
S_{54} = 31 & -(2/3)56 & = -19/3
\end{array}$$

$$\bar{y} = 56/9 \quad ; \qquad \Delta = 6.6.2.4.2 = 576 \; .$$

$$C_{11} = 1/6, \; C_{22} = 1/6, \; C_{33} = 1/2, \; C_{44} = 1/4, \; C_{55} = 1/2$$

$$C_{12} = C_{13} = C_{14} = \ldots = C_{45} = 0 \; .$$

$$\begin{array}{llll}
(b_{10}) & b_1 = +16/6 = 8/3 & = 48/18 & b_o = \frac{56}{9} + \frac{2}{3} \cdot \frac{2}{3} + \frac{2}{3} \cdot \frac{19}{6} \\
(b_{01}) & b_2 = -15/6 = -5/2 & = 45/18 & \\
(b_{20}) & b_3 = -\,4/6 = -2/3 & = -12/18 & b_o = 158/18 \qquad (b_{oo}) \\
(b_{11}) & b_4 = -\,4/4 = -1 & = -18/18 & \\
(b_{02}) & b_5 = -19/6 = -19/6 & -57/18 &
\end{array}$$

$$Y = (158 + 48x_1 - 45x_2 - 12x_1^2 - 18x_1x_2 - 57x_2^2)/18$$

Furthermore, we have

$$S_{44} = 454 = \frac{56^2}{9} = 105.556$$

$$S(Y-\bar{y}) = \underset{j}{S}\, b_j\, S_{jy} = 946/9 = 105.111$$

Variance	DF	SS	MS
Regression	5	105.111	21.022
Residual	3	0.445	0.148
Total	8	105.556	...

The residual mean square 0.148 is smaller than the error mean square of 0.422 obtained previously, which shows that the fit of the equation of the 2nd degree to the actual surface of yield is quite satisfactory.

With $s^2 = 0.148$ we obtain standard errors $c_{jj}\, s^2$

j	1 and 2	3 and 5	4
$c_{jj}s^2$	.02469	.07407	.03704
	$= (.157)^2$	$(.272)^2$	$(.192)^2$

The standard errors are small, but here again we have only 3 d.f. for s^2. It is not easy to interpret the surface of the second degree represented by the regression equation as given above. We therefore transform this regression equation. First we perform a parallel translation of the coordinates in such a way as to have the new origin in the point for which Y is a maximum.

From

$$(1) \;....\quad Y = b_{00} + b_{10}x_1 + b_{01}x_2 + b_{20}x_1^2 + b_{11}x_1x_2 + b_{02}x_2^2$$

we get by partial differentiation with respect to x_1 and x_2:

$$\frac{\partial Y}{\partial x_1} = 0 \;, \qquad b_{10} + 2b_{20}x_1 + b_{11}x_2 = 0$$

$$\frac{\partial Y}{\partial x_2} = 0 \;, \qquad b_{01} + b_{11}x_1 + 2b_{02}x_2 = 0$$

If the solutions of these equations are $x_1 = d_1$, $x_2 = d_2$, then

(2a) .... $$b_{10} + 2b_{20}d_1 + b_{11}d_2 = 0$$

(2b) ... $$b_{01} + b_{11}\, d_1 + 2b_{02}d_2 = 0$$

Changing to new coordinates u_1 and u_2 by means of the transformations

$$x_1 = u_1 + d_1\,, \qquad x_2 = u_2 + d_2$$

we have from (1), taking into account (2a) and (2b)

(3) .... $$Y = b_{00} + b_{10}d_1 + b_{01}d_2 + b_{20}d_1^2 + b_{11}d_1d_2 + b_{02}d_2^2$$
$$+ b_{20}u_1^2 + b_{11}u_1u_2 + b_{02}u_2^2$$

If we put

(4) .... $$Y_M = b_{00} + b_{10}d_1 + b_{01}d_2 + b_{20}d_1^2 + b_{11}d_1d_2 + b_{02}d_2^2$$

Then (3) can be written as

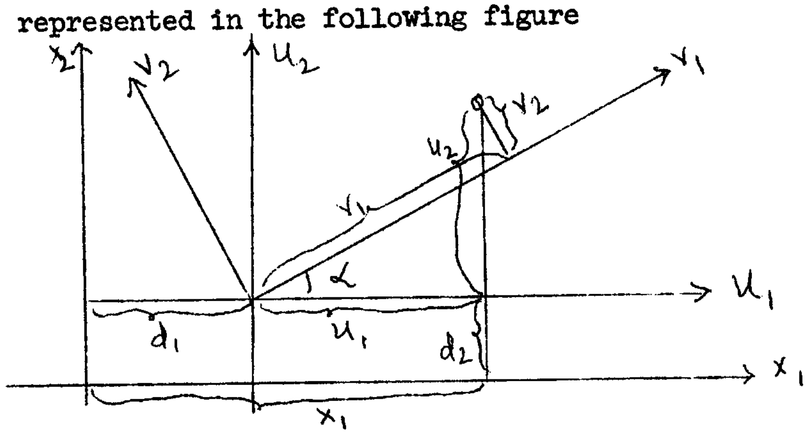
(5) .... $$Y = Y_M + b_{20}u_1^2 + b_{11}u_1u_2 + b_{02}u_2^2\ .$$

Next we change from coordinates u_1, u_2 into coordinates v_1, v_2 by means of the transformations

(6a) .... $$u_1 = v_1 \cos\ \alpha\ - v_2 \text{ sim } \alpha$$

(6b) .... $$u_2 = v_1 \sin\ \alpha\ + v_2 \cos\ \alpha$$

The two transformations from x_1, x_2 into u_1, u_2 and from u_1, u_2 into v_1, v_2 are represented in the following figure



Substitution of (6a) and (6b) in (5) gives

$$Y = Y_M + B_{20} v_1^2 + B_{11} v_1 v_2 + B_{02} v_2^2$$

where

$$B_{20} = b_{20} \cos^2 \alpha + b_{11} \sin \alpha \cos \alpha + b_{02} \sin^2 \alpha$$
$$B_{02} = b_{20} \sin^2 \alpha = b_{11} \sin \alpha \cos \alpha + b_{02} \cos^2 \alpha$$
$$B_{11} = b_{11} \cos 2\alpha + (b_{02} - b_{20}) \sin 2\alpha$$

Determining α such that $B_{11} = 0$, we het

(7) .... $\tan 2\alpha = b_{11}/(b_{20} - b_{02})$

and (5) becomes

(8) .... $Y = Y_M + B_{20} v_1^2 + B_{02} v_2^2$

By straightforward analysis it may be seen that

(9) .... $B_{20} + B_{02} = b_{20} + b_{02}$

and

(10) .... $B_{20} \cdot B_{02} = b_{20} b_{02} - b_{11}^2/4$.

This leads to

(11) .... $B_{20}^2 - (b_{20} + b_{02})B_{20} + b_{20} b_{02} - b_4^2/4 = 0$

In our example we have

$$b_{20} = -12/18 , \quad b_{11} = -18/18 , \quad b_{02} = -57/18$$
$$= -4/6 \quad = -6/6 \quad = -19/6$$

and therefore

$$b_{20}b_{02} - b_{11}^2/4 = \frac{4.19}{6.6} - \frac{1}{4} = 67/36 .$$

This expression informs us whether the contours are hyperbolae, parabolae or ellipses

$$b_{20} - b_{02} - b_{11}^2/4 \begin{cases} < 0 & \text{hyperbola} \\ = 0 & \text{parabola} \\ > 0 & \text{ellipse} \end{cases}$$

In our example the contours are ellipses. From (11) we get

$$B_{20}^2 + \frac{23}{6} B_{20} + \frac{67}{36} = 0$$

the roots of which are

$$B_{20} = -.570\,375 \quad \text{and} \quad -3.262\,958$$

and from (7) we obtain

$$\tan 2\alpha = -6/15 = -.400$$

$$2\alpha = 158^{\circ}\ 12', \quad \alpha = 79^{\circ}\ 6' \quad .$$

From (2) we obtain

$$d_1 = +2.604\,478 \,, \qquad d_2 = -.805\,970$$

Furthermore we find

$$Y_M = 13.257\,877$$

and for (8)

$$B_{20}\, v_1^2 + B_{02}\, v_2^2 = Y - 13.257\,877$$

or

$$\underline{3.262958\, v_1^2 + .570375\, v_2^2 = 13.257\,877 - Y}$$

To draw the contour-ellipses, we determine the half-axes a and b and the distance f between the focus and the center. For $Y = 5$ and 10 we obtain

	$Y = 5$	$Y = 10$	$A=80+Y$
$Y_M - Y$	8.257 877	3.257 877	
$v_2=0: v_1^2=(Y_M-Y)/B_{20}=b^2$	2.530 806	.998 447	
b	1.591	.999	
$v_1=0: v_2^2=(Y_M-Y)/B_{02}=a^2$	14.477 623	5.711 676	
a	3.805	2.390	
$f^2 = (a^2 - b^2)$	11.946 817	4.713 229	
f	3.456	2.171	

The contour ellipses for $A = 85$ and $A = 90$ may be drawn from the above data.

Remarks: The selection of suitable designs for 2, 3 or more factors must be carefully studied: rotatable designs. The method of Box and Wilson has been modified for use in actual production: EVOP = evolutionary operation. See Box and Hunter, 1959.

6. Incomplete block designs.

6.1 Introductory remarks

We have found that we can decrease the size of blocks in factorial designs thus increasing the precision, by the device of confounding.

In some experiments the number of treatments is very high. In variety trials we may have as many as hundred varieties to compare. In such cases it is clearly impossible to accomodate all varieties in one block.

In oth er cases the block size may be restricted to a very few units by the nature of things. If we try repellants to insectizides, the two arms of a person may form a block. In a similar way the two eyes of an animal may form a block when experimenting with drugs in ophthalmology.

If one is willing not to complete blocks, he will have to cope with a more complex analysis. In order to have an analysis which is not too complicated, we will retain a certain degree of symmetry in the design. It is not the purpose of this course to enter into the question of the different possible incomplete block designs and their construction. The time left to us in this semester will just be sufficient to make a few general remarks and to give some examples of incomplete block designs.

These are two broad classes of incomplete block designs, namely

(a) (totally) balanced incomplete block designs, and

(b) partially balanced incomplete block designs.

Let us first consider an example of a balanced incomplete block design. If we had to compare 7 varieties A, B, C, D, E, F and G and if we would like to have only 4 plots per block, then a suitable design would be the following:

Block	Varieties
I	A B C F
II	A B E G
III	A C D E
IV	A D F G
VI	B D E F
VII	C E F G

The design comprises 7 blocks, so these are 28 units. We may see that each variety occurs 4 times. But the fundamental property of balanced in complete blocks lies in the fact, that each pair of treatments occurs equally often together in some block.

With 7 treatments we have $7.6/2 = 21$ pairs of treatments. In the design given above each of these pairs occurs twice together in a block, e.g. A and B occur in blocks I and II, A and C in blocks I and III, and so on.

This ensures that all comparisons between varieties can be made within blocks, that is, with essentially the same precision. These comparisons may be assessed for significance using a within blocks error.

But the symmetry has yet another consequence. A occurs in blocks I, II, III and IV, B in blocks I, II, V and VI. If we compare the yields in blocks III and IV with those in V and VI, we find that this is a comparison of A with B, as the two blocks contain in both cases the varieties C, E, F and G once, D twice. We thus also have a comparison of each pair of varieties between blocks. The corresponding error must of course also be an error between blocks, and this will generally be greater than the within blocks error. The two errors (within and between blocks) have to be combined in a complete analysis of balanced incomplete block experiments.

As an example of a partially balanced incomplete block design, let us consider an experiment with 9 varieties (A, B, ... I) in blocks of 3 plots.

If we arrange the nine letters in a square of size 3, we have

A	B	C
D	E	F
G	H	I

An PBIB design is obtained by forming 6 blocks containing the varieties forming the rows and columns of the square just given.

Block	Treatments		
I	A	B	C
II	D	E	F
III	G	H	L
IV	A	D	G
V	B	E	H
VI	C	F	I

In this design we are able to compare A with B, C, D, G within blocks. On the other hand the comparisons of A with E, F, H and I are between blocks, and therefore of lower precision.

We could easily obtain a balanced design if we added to the six blocks above the following six, obtained from a corresponding graeco-latin square.

VII	A	F	H
VIII	B	D	I
IX	C	E	G
X	A	E	I
XI	B	F	G
XII	C	D	H

6.2 Balanced incomplete block designs.

Let us first consider a few examples of (completely) balanced incomplete block designs. These examples are given to show how in a given experimental situation the design can be tailored to the needs of the experimenter.

(a) Simplest possible design. In an investigation on the effect of two substances (C and D) and a control (O) on the duration of healing of an abrasion of the cornea, the following design was used, considering the two eyes of a rabbit as a block (Linder, 1953).

Rabbit	Eye 1	Eye 2
I	O	C
II	O	D
III	C	D
IV	O	C
V	O	D
VI	C	D

(b) Seven brands of tires are to be compared, measuring the wear in an appropriate way. The design on p. 120 can be used. The groups are assigned at random to cars and the treatments (tires) at random to the four wheels. In this way the variability within cars can be used as experimental error.

(c) The design mentioned under (b) is not the best possible because the position of the wheels (front, rear, right, left) is not taken into account. This drawback may be avoided, at least for 5 tires, if we use the following design due to Youden (1937, 1940).

Block (car)	Wheel front ℓ	Wheel front r	Wheel rear ℓ	Wheel rear r	
I	A	B	C	D	(E)
II	E	A	B	C	(D)
III	D	E	A	B	(C)
IV	C	D	E	A	(B)
V	B	C	D	E	(A)

This is again a balanced incomplete block design, as each pair of treatments (tire brands) occurs three times in the same block. It can be constructed from a Latin Square by deleting the last row.

Notation.

t = number of treatments
b = number of blocks
r = number of replications
k = number of units per block
$rt = bk$ = total number of units.
λ = number of blocks, in which each pair of treatments occurs together

The characteristic number λ can be determined from t, r and k. Let us consider a specific treatment, say A. According to the definition of , these are $\lambda(t-1)$ other treatments in the same blocks as A. On the other hand, A occurs in r blocks, and the number of units occuring with A in the smae blocks is r(k-1). Thus

$$\lambda(t-1) = r(k-1)$$

and

$$\underline{\lambda = r(k-1)/(t-1)}$$

Let us consider the designs given on p. 141.

(a) $t = 3, \quad b = 6, \quad r = 4, \quad k = 2$

$$= 4.1/2 = 2$$

(b) $t = 7, \ b = 7, \ r = 4, \ k = 4$

$$= 4.3/6 = 2$$

(c) $t = 5, \ b = 5, \ r = 4, \ k = 4$

$$= 4.3/4 = 3$$

<u>Analysis of balanced incomplete block designs.</u>

As an example consider the testing of 13 substances A, B, ... M with respect of their influence on the quantity of food eaten by pigs (<u>Dr. Mastrangelo</u>, Geneva). The substances were tested in 4 different containers on one day (block), the experiment being carried out on 13 days. The data are the weights of remaining food in each container (in 100 gr), the total quantity of food given was 7.5 kg.

Day		Container 1		2		3		4	Total
I	A	23	B	27	T	35	D	34	119
II	B	0	C	10	K	2	E	34	46
III	F	25	C	20	D	2	L	1	48
IV	D	28	E	32	M	21	G	18	99
V	H	17	F	22	A	20	E	31	90
VI	G	20	B	25	I	8	F	20	73
VII	C	4	T	14	H	9	G	12	39
VIII	I	0	D	21	K	19	H	20	60
IX	I	2	T	9	E	30	L	23	64
X	F	19	T	11	M	8	K	3	41
XI	G	13	K	10	A	27	L	19	69
XII	L	17	B	19	M	20	H	18	74
XIII	I	4	M	13	A	22	C	28	67
Total									889

In this design we have

$$t = 13, \quad b = 13, \quad r = 4, \quad k = 4\,, \qquad \text{so}$$

$$\lambda = 4.3/12 = 1$$

Analysis within blocks

We consider first only the error within blocks. Later we will also take into account the error between blocks. We are not giving a full treatment but we will give some indications on the way the analysis is built up.

We note that the quantity of food remaining varies considerably from day to day. It may well be that the pigs eat different quantities of food on different days. The analysis should take this into account. The model which is used for this analysis is given by

$$y_{jk} = \alpha + \tau_j + \beta_k + \delta_{jk}$$

where y_{jk} is the result for treatment j in block k.

α = general average

τ_j = effect of treatment j

β_k = effect of block k

δ_{jk} = random error, normally distributed with mean zero and variance σ_w^2.

In the model, we assume that τ and β give the effects of treatments and blocks as measured from the general mean α. Therefore

$$\tau_1 + \tau_2 + \ldots + \tau_j + \ldots + \tau_t = 0 \quad \text{(j varies over all t treatments)}$$

$$\beta_1 + \beta_2 + \ldots + \beta_k + \ldots + \beta_b = 0 \quad \text{(k varies over all b blocks)}$$

Let T be the total result for all units under treatment j. Then

$$T = y_{j1} + y_{j2} + \ldots + y_{jk} + \ldots + y_{jr} \quad ,$$

the treatment j occurring in r blocks. Substituting y_{jk} from the model we get, discarding the error term δ_{jk},

$$T = r\,\alpha + (\beta_1 + \beta_2 + \ldots + \beta_k + \ldots + \beta_r) + r\,\tau_j$$

If we consider the total B_t of all blocks in which treatment j occurs, then

$$B_t = k\,r\,\alpha + k(\beta_1 + \beta_2 + \ldots + \beta_k + \ldots + \beta_r) + r\,\tau_j +$$
$$+ \lambda\,(\tau_1 + \tau_2 + \ldots + \tau_{j-1} + \tau_{j+1} + \ldots + \tau_t)$$

Comparing T and B_t we notice that the β (and α) can be eliminated by forming

$$kT - B_t = Q = k\,r\,\tau_j - r\,\tau_j - \lambda(-\tau_j)$$

which is

$$Q = kT - B_t = [(k-1)\,r + \lambda]\,\tau_j \;.$$

Substituting for λ we get

$$(k-1)r + \lambda = (k-1)r + (k-1)r/(t-1)$$
$$= r\,t(k-1)/(t-1)$$

In order to estimate the deviation τ_j we have to form

$$(t-1)Q/rt(k-1)$$

and if a is the general mean of all units, we have to calculate.

$$a + (t-1)Q/rt\,(k-1)$$

which is the mean effect of a treatment corrected for the influence of those blocks in which the particular treatment occurs.

In our example we get

Treatment	T	B_t	$Q=4T-B_t$	$a+Q/13$
A	92	345	23	18.87
B	71	312	-28	14.94
C	62	200	48	20.79
D	85	326	14	18.17
E	127	299	209	33.17
F	86	252	92	24.17
G	63	280	-28	14.94
H	64	263	- 7	16.56
I	14	264	-208	1.10
J	69	263	13	18.10
K	34	216	-80	10.94
L	60	255	-15	15.94
M	62	281	-33	14.56
Total	889	3556	0	222.25

Checks: $4(889) = 3556$; $4(222.25) = 889$

$a = 889/52 = 17.096$, $(t-1)/rt(k-1) = 12/4.13.12 = 1/13$

The "within blocks error" is found without difficulty. From the total SS we subtract the SS between blocks, the difference being the SS within blocks. The quantities Q are obtained as comparisons within blocks, their sum of squares is therefore part of the SS within blocks. If we subtract the SS of the Q's from the SS within blocks we obtain the error SS within blocks. The total SS is obtained as

$$SS(\text{total}) = 23^2 + 27^2 + \ldots + 28^2 - 889^2/52 = 4{,}620.5 \quad ,$$

the sum of squares for blocks as

$$SS(\text{blocks}) = (119^2 + 46^2 + \ldots + 67^2)/4 - 889^2/52 = 1{,}660.3$$

The SS for the Q's which is called the "SS for treatments, adjusted for block differences", is obtained as

$$SS(\text{treatments, adjusted}) = (t-1)S(Q^2)/rtk(k-1)$$

for which we obtain in our example

$$SS(\text{treatments, adjusted}) = 12(23^2 + 28^2 + \ldots + 33^2)/4.13.4.3$$

$$= 2{,}075.7$$

The analysis of variance can be set up as follows:

Variance	DF	SS	MS
Blocks (unadjusted)	12	1,660.3	...
Treatments (adjusted)	12	2,075.7	173.0
Error, within blocks	27	884.5	$32.760 = s_w^2$
Total	51	4,620.5	...

To test the hypothesis $\tau_1 = \tau_2 = \dots = \tau_t = 0$ we may use

$$F = 173.0/32.760 = 5.280$$

With $n_1 = 12$, $n_2 = 27$ we find $F_{0.001} = 4.170$. We conclude that there are highly significant differences between some of the substances. For t-tests, we have to calculate a so-called effective error variance, using the formula

$$s_w^2\left[1 + \frac{t-k}{t(k-1)}\right] = k(t-1)\ s_w^2/t(k-1)$$

The variance of a treatment mean is obtained as

$$k(t-1)\ s_w^2/rt(k-1)$$

In our example the effective error variance of a mean is

$$(4)(12)(32.760)/(4)(13)(3) = 10.08 = (3.175)^2$$

In the analysis we have so far only used comparisons and errors within blocks. As already stated, we may also use comparisons between blocks, and compute a corresponding error between blocks.

Analysis between blocks.

The formula for B_t on p. 125

$$B_t = k\,r\,\alpha + k(\beta_1 + \beta_2 + \dots + \beta_k + \dots + \beta_r) + r\ \tau_j$$
$$+ \lambda(\tau_1 + \tau_2 + \dots + \tau_{j-1} + \tau_{j+1} + \dots + \tau_t)$$

may also be written as

$$B_t = k\,r\,\alpha + k(\beta_1 + \beta_2 + \dots + \beta_k + \dots + \beta_r) + (r-\lambda)\,\tau_j$$

from which it is seen that B_t depends only on τ_j, not on any other. B_t of course contains block parameters β_k, and an estimate of τ_j based on B_t is therefore of lower precision than one based on Q.

If $\bar{B}$ is the average of the B_t's, then

$$a + \frac{B_t - \bar{B}}{r - \lambda}$$

is an estimate of $\alpha + \tau_j$.

For treatment A we obtain

$$\begin{aligned} B_t &= & & 345 \\ -\bar{B} &= -3556/13 & = & \underline{-273.538} \\ B_t - \bar{B} &= & & 71.462 \end{aligned}$$

$$a + (B_t - \bar{B})/3 = 17.10 + 23.82 = 40.92$$

This is very much greater than the estimate obtained by

$$a + Q/13 \quad (= 18.87) .$$

To obtain the error of these estimates, we note that

$$W = (t-k)T - (t-1)B_t + (k-1)G$$

G being the general total has an expected value equal to

$$t(k-1)(\beta_{r+1} + \beta_{r+2} + \dots + \beta_b)$$

the β's being those of blocks in which treatment j does not occur. The expected value of W does not contain any τ; we may consider W as a quantity connected with block differences, adjusted for treatment effects. In our example we get W for treatment A, observing that

$$W = 9\,T - 12\,B_t + (3)(889)$$

$$
\begin{array}{rlll}
 & (9) & (92) & = 828 \\
+ & (3) & (889) & = \underline{2,667} \\
 & & & 3,495 \\
- & (12) & (345) & = \underline{-4,140} \\
 & & W & = -\ 645
\end{array}
$$

The SS of W gives

SS (blocks, adjusted for treatment differences) =

$$= S(W^2)/\ tr\ (t-k)(k-1)\ .$$

In our example we get

SS (blocks, adjusted) = 1,504.2

We may set out the following analysis of variance

Variance	DF		SS	MS
Treatments, unadjusted	(t-1)	12	2,231.8	...
Blocks, adjusted	(b-1)	12	1,504.2	$125.353=s_b^2$
Error, within blocks	(tr-t-b+1)	27	884.5	$32.760=s_w^2$
Total	tr-1	51	4,620.5	...

The effective error variance for a treatment mean $a + \frac{B_t - \bar{B}}{r-\lambda}$ is equal to

$$\frac{k}{r-\lambda}\left[\frac{kr}{\lambda t}(s_b^2 - s_w^2) + s_w^2\ \right]$$

which for our example gives

$$\frac{4}{4-1}\left[\frac{4.4}{1.13}(125.353 - 32.760) + 32.760\right]$$

$$= 195.627 = (13.99)^2\ .$$

This standard error of a mean is much larger than the standard error (3.175) obtained from the analysis within blocks.

<u>Combined analysis (within and between blocks)</u>

Consider the inverses G_w and G_b if the effective error variances of treatment means for the "within" and "between" analysis. We had

$$1/G_w = k(t-1)s_w^2/rt(k-1)$$

and

$$1/G_b = \frac{k}{r-\lambda} \left[\frac{kr}{\lambda t} (s_b^2 - s_w^2) + s_w^2 \right]$$

If t_w and t_b are the estimates obtained by the two methods, the best combined estimate will be

$$\hat{t} = \frac{G_w t_w + G_b t_b}{G_w + G_b}$$

In our example we had

$$G_w = \frac{1}{10.08} \sim 0.1; \quad G_b = \frac{1}{195.6} \sim 0.005$$

and for treatment A we obtain

$$\frac{(0.1)(18.87) + (0.005)(40.92)}{0.105} = 19.94 \quad ,$$

a value differing only to a small amount from t_w .

It may be shown that the estimates $\hat{t}$ can be calculated as

$$\hat{t} = (T + \mu \quad W)/r$$

where

$$\mu = (s_b^2 - s_w^2)/\ t(k-1)\ s_b^2 \quad .$$

If s_b^2 turns out to be smaller than s_w^2, the value $\mu = 0$ will be used. It may also be shown that the error variance of $\hat{t}$ is $1/(G_1 + G_2)$ or

$$s_w^2[1 + (t-k)\mu\ ]/r$$

In our example we have

$$\mu = 92.593/(39)(125.353) = 0.018940$$

and the error variance is

$$32.760[1 + (9)(.018\ 940)]/4 = 9.586 = (3.096)^2$$

The error variance for the combined analysis is only slightly smaller than that of the analysis "within blocks".

The results for the three analyses are summarized below.

Treatment (Substance)	W	$m+Q/13$	$m+(B_t-\bar{B})/3$	$(T+ \quad W)/4$
A	-645	18.87	40.92	19.94
B	-438	14.94	29.92	15.68
C	825	20.79	-7.42	19.41
D	-480	18.17	34.58	18.98
E	222	33.17	25.58	32.80
F	417	24.17	9.92	23.48
G	-126	14.94	19.25	15.15
H	87	16.56	13.58	16.41
I	-375	1.10	13.92	1.73
J	132	18.10	13.58	17.88
K	381	10.94	-2.08	10.30
L	147	15.94	10.92	15.70
M	-147	14.56	19.58	14.80
Total	0	222.25	222.25	222.26
Standard error	...	3.175	13.99	3.096

Steps in the complete analysis

(1) Calculate block totals, the treatment totals T and B_t.

(2) Calculate

$$W = (t-k)T - (t-1)B_t + (k-1)G$$

(3) Analysis of variance:

	DF	MS
Treatments, unadjusted	$t-1$	...
Blocks, adjusted	$b-1$	s_b^2
Error, within blocks	$tr-2t+1$	s_w^2
Total	$tr-1$	...

The SS (Blocks, adjusted) is obtained as

$$S(W^2)/tr(t-k)(k-1)$$

(4) Calculate

$$\mu = (s_b^2 - s_w^2)/t(k-1)s_b^2$$

(5) The effective error mean square is

$$s_w^2[1 + (t-k)\mu]$$

The complete or combined analysis is based on the model

$$y_{jk} = \alpha + \tau_j + \beta_k + \delta_{jk}$$

where α is again the general mean, τ_j the effect of treatment j, β_k the block effect and δ_{jk} the error within blocks, which is normally distributed with mean zero and variance σ_w^2. But now the β_k are no more fixed constants they are normally distributed with zero mean and variance σ_b^2, and they are independent of the δ_{jk}.

Balanced lattice design.

These designs are formed as mentioned in 6.1 from a latin square arrangement; they are available for 9, 16, 25, 49, 64, and 81 treatments - not for 36. They are constructed using sets of completely orthogonal latin squares.

If $k^2 = t$ is the number of treatments, we will have $r = k + 1$ replications. The steps in the analysis are as follows:

(1) Calculate the totals for blocks, replications, the grand total G and the treatment totals T.

(2) Calculate B_t; the totals of the B_t values should be k times the total of the T values.

(3) Calculate the

$$W = k\,T - (k+1)\,B_t + G \;.$$

The sum of the W should be zero.

(4) Calculate the analysis of variance

Variance	DF	MS
Replications	k	...
Treatments, unadjusted	k^2-1	...
Blocks within replications, adjusted	k^2-1	s_b^2
Error, within blocks	$(k-1)(k^2-1)$	s_w^2
Total	$k^3 + k^2 - 1$	...

The sum of squares for blocks within replications, adjusted for treatment effects, is

$$\sum(W^2)/k^3(k+1)$$

(5) The factor μ is

$$= (s_b^2 - s_w^2)/k^2 s_b^2 .$$

Calculate

$$T + \mu W$$

and

$$(T + \mu W)/(k+1)$$

the adjusted treatment means.

(6) The effective error mean square for use in t-tests is

$$s_w^2 (1 + k \quad)$$

and the standard error of the adjusted treatment means if

$$s_w^2(1 + k \mu)/(k+1)$$

(7) For an approximate test of the hypothesis

$$\tau_1 = \tau_2 = \dots = \tau_t = 0$$

calculate the sum of squares for the quantities $(T + \mu W)$. The corresponding mean square can be compared to the error mean square using F with k^2-1 and $(k-1)(k^2-1)$ degrees of freedom.

The data on page 135 can be analyzed according to the procedures just outlined. The experiment, carried out by Dr. H. Marthaler (Institute of Dentistry, University of Zurich, Switzerland) was concerned with the effect of 25 substances in preventing caries in rats. The blocks are formed by litters of 5 rats. Two such litters were matched so that each treatment (substance) was given to two rats in one cage. The treatment effects are measured by the number of fissions with dental lesions in the lower first and second molars. The highest possible number of fissions with dental lesions for one rat is 12.

The proportions 0, 1/12, 2/12, ... 11/12, 1, are transformed by the angular transformation $p = \sin^2 \phi$. In order to make the transformed values readily comparable with the actual numbers 0, 1, 2, ... 12, we use $4\phi/30$ instead of ϕ , and to get rid of decimals the values $40\phi/30$ are given.

p	0	1/12	2/12	3/12	4/12	5/12	6/12	7/12	8/12	9/12	10/12	11/12	1
$40\phi/30$	0	22	32	40	47	54	60	66	73	80	88	98	120

The randomization process is carried out in 3 steps: (1) Litters are assigned to blocks at random; (2) Within litters the animals are assigned at random

Plan of experiment and transformed number of dental lesions.

Replication 1

1	40 120	2	32 32	3	80 47	4	32 60	5	22 22
6	0 22	7	0 47	8	0 22	9	0 32	10	73 32
11	54 98	12	60 88	13	40 120	14	54 66	15	32 73
16	54 120	17	80 98	18	40 98	19	47 47	20	54 88
21	54 98	22	40 120	23	120 88	24	60 80	25	98 47

Replication 2

1	88 120	6	0 22	11	60 120	16	80 120	21	60 98
2	0 47	7	40 80	12	54 73	17	120 120	22	80 88
3	32 54	8	0 32	13	60 54	18	80 88	23	73 120
4	22 22	9	22 32	14	80 40	19	66 66	24	88 120
5	0 40	10	66 120	15	40 120	20	32 120	25	40 98

Replication 3

1	88 47	7	40 32	13	47 40	19	54 47	25	60 32
21	98 66	2	32 40	8	22 0	14	40 32	20	120 66
16	120 73	22	98 98	3	47 40	9	0 32	15	80 80
11	80 98	17	73 40	23	32 66	4	32 0	10	47 120
6	60 22	12	98 54	18	98 88	24	88 80	5	32 32

Replication 4

1	60 98	12	47 120	23	66 98	9	0 32	20	60 88
16	66 66	2	0 0	13	22 0	24	60 54	10	40 22
6	0 0	17	98 120	3	40 60	14	73 73	25	98 120
21	88 98	7	32 40	18	120 66	4	40 54	15	60 98
11	120 66	22	88 66	8	0 0	19	73 0	5	22 22

Replication 5

1	66 98	17	88 80	8	0 22	24	73 98	15	80 66
11	98 88	2	47 32	18	120 120	9	22 22	25	88 80
21	88 60	12	80 73	3	54 47	19	73 60	10	120 66
6	0 0	22	120 120	13	80 66	4	40 22	20	120 22
16	120 98	7	40 80	23	80 88	14	66 66	5	0 22

Replication 6

1	80 73	22	60 66	18	120 98	14	73 73	10	120 88
6	0 40	2	22 22	23	66 120	19	32 88	15	32 32
11	80 98	7	0 40	3	32 40	24	73 54	20	98 60
16	22 60	12	60 54	8	0 0	4	0 60	25	0 47
21	60 66	17	40 60	13	54 60	9	0 22	5	22 32

to treatments; (3) The substances are alloted to the treatment number 1, 2, ... 25 at random.

The analysis is left as an exercise to the student. It should be observed that the unit is the total of the results of the two animals treated together

6.3 Partially balanced incomplete block designs.

We discuss only one example of PBIB, namely the simple lattice. This is obtained as the rows and columns of the treatments arranged in a square, the rows and columns being used as blocks. Consequently, the number of treatments has to be a square: 9, 16, 25, 36, 64 etc The design is very suitable if only two replications of a large number of treatments are feasible

As an example consider a trial with 25 varieties of wheat carried out at the Federal Agricultural Research Station in Oerlokan (Switzerland)(Linder, 1953, p. 148).

Plan of simple lattice experiment

Blocks	Replication 1					Replication 2					Blocks
I	2	4	5	1	3	18	3	13	8	23	VI
II	16	19	17	20	18	2	7	17	22	12	VII
III	9	7	10	8	6	19	4	24	14	9	VIII
IV	11	15	13	14	12	25	15	10	5	20	IX
V	24	22	23	21	25	6	11	21	1	16	X

The randomization is again in three steps: (1) the groups of varieties in one row are assigned to the blocks at random; (2) the varieties are assigned at random to the plots within each block; (3) the varieties are assigned at random to the numbers 1, 2, ... 25.

We have

$$t = 25, \quad r = 2, \quad b = 10, \quad k = 5 .$$

Some pairs of varieties, e.g. 2, 16 and 2, 9 never occur together in one block, while other pairs, as 2,4 and 2,5 occur once together in the same block.

The analysis of the results proceeds as follows: In a first step we arrange the results according to replications and blocks in such a way as to have the numbers of the varieties in their natural order. We also set out the totals for each variety.

<u>Yields in 10g on plots of 2.25 m^2</u>

Block											B	C	μC
				Replication 1									
I	1	83	2	65	3	45	4	94	5	60	347	+53	+6.4
III	6	89	7	98	8	68	9	65	10	76	396	+ 9	+1.1
IV	11	96	12	64	13	84	14	86	15	98	428	- 1	- .1
II	16	68	17	83	18	68	19	85	20	54	358	+34	+4.1
V	21	78	22	101	23	85	24	77	25	101	442	-45	-5.4
											1971	+50	+6.1
				Replication 2									
X	1	94	6	95	11	109	16	83	21	80	461	-47	-5.7
VII	2	80	7	92	12	67	17	74	22	70	383	+28	+3.3
VI	3	52	8	71	13	72	18	65	23	79	339	+11	+1.3
VIII	4	104	9	63	14	83	19	80	24	74	404	+ 3	+ .4
IX	5	70	10	84	15	96	20	90	25	94	434	-45	-5.4
											2021	-50	-6.1

Sums for varieties (unadjusted)

										μc
1	177	2	145	3	97	4	198	5	130	+ 6.4
6	184	7	190	8	139	9	128	10	160	+ 1.1
11	205	12	131	13	156	14	169	15	194	- .1
16	151	17	157	18	133	19	165	20	144	+ 4.1
21	158	22	171	23	164	24	151	25	195	- 5.4
	-5.7		+3.3		+1.3		+ .4		-5.4	($\pm$ 6.1)

Instead of adjusting varieties for block differences, we have to adjust blocks for differences between varieties. Instead of calculating

$$Q = kT - B_t$$

we compute

$$C = T_b - rB$$

where B is the block total, T_b the total for those varieties occuring in one block. For block I we get

$$\begin{aligned} T_b &= 177 + 145 + 97 + 198 + 130 &&= 747 \\ - rB &= -(2)(347) &&= -694 \\ C & &&= +53 \end{aligned}$$

The second step consists in setting up the analysis of variance. We calculate

$$SS\ (\text{total}) = 83^2 + 65^2 + \ldots + 94^2 - 3{,}992^2/50 = 10{,}102.72$$

$$SS\ (\text{varieties, unadjusted}) = (177^2+145^2+\ldots+195^2)/2-3{,}992^2/50 = 8{,}215.72$$

$$SS\ (\text{replications}) = (2{,}021 - 1{,}971)^2/50 = 50.00$$

$$SS\ (\text{blocks, within replications, adjusted}) = S(C^2)/kr(r-1) - S(R_c^2)/k^2r(r-1)$$

where R_c is the total of the C's for each replicate. The two R_c's should be equal but with opposite signs. We obtain

$$SS\ (\text{blocks, within repl., adjusted}) = (53^2+9^2+\ldots+45^2)/10-(50^2+50^2)/50 =$$

$$= 1{,}022.00$$

The general scheme for the analysis of variance of simple lattices (with 2 replications), triple lattices (with 3 repl.) and lattices with r replications, is

Variance	D.F.	M. S.
Replications	$r-1$	...
Treatments, unadjusted	k^2-1	...
Blocks within replications, adjusted	$r(k-1)$	s_b^2
Within blocks error	$(k-1)(rk-k-1)$	s_w^2
Total	rk^2-1	

For our example we get

Variance	DF	SS	MS
Replications	1	50.00	...
Treatments, unadjusted	24	8,215.72	...
Blocks within repl., adjusted	8	1,022.00	$127.750=s_b^2$
Error within blocks	16	815.00	$50.938=s_w^2$
Total	49	10,072.72	...

In the case of the balanced lattice designs there is a simple, though approximate test of the hypothesis $\tau_1 = \tau_2 = \ldots = \tau_t = 0$. For the partially balanced lattice there is no such simple test.

To find the adjusted treatment means, we have to calculate

$$\mu = (s_b^2 - s_w^2)/k(r-1)s_b^2$$

for which we have in our example

$$\mu = (127.750 - 50.938)/(5)(1)(127.750)$$

$$\mu = 0.120524$$

From this we obtain the quantities μC, and these are used to adjust the mean totals. For variety 1 we have

$$177.0 + 6.4 - 5.7 = 177.7$$

and similarly for the other varieties.

Total of varieties, adjusted

1	177.7	2	154.7	3	104.7	4	204.8	5	132.0
6	179.4	7	194.4	8	141.4	9	129.5	10	155.9
11	199.2	12	134.2	13	157.2	14	169.3	15	188.5
16	149.4	17	164.4	18	138.4	19	169.5	20	142.7
21	146.9	22	168.9	23	159.9	24	146.0	25	184.2

To test the hypothesis $\tau_1 = \tau_2$ we have to consider three situations, namely

(a) 2 varieties in the same block

$$t = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{2(1+(r-1)\mu)s_w^2/r}}, \quad n = (k-1)(rk-k-1)$$

(b) 2 varieties in different blocks

$$t = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{2(1+r\mu)s_w^2/r}}, \quad n = (k-1)(rk-k-1)$$

(c) 2 varieties with unspecified position

$$t = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{2(1+rk\mu/(k+1))s_w^2/r}}, \quad n = (k-1)(rk-k-1)$$

The denominator in these three cases are

(a) 7.554; (b) 7.949; (c) 7.820

so that in this example formula (c) could be used without hesitation to compare any pair of treatment, whether in the same block or not.

7. Analysis of groups of experiments.

Experiments are often repeated: at different times, in different laboratories, in different localities. It has not been possible, because of lack of time, to deal with this subject of the joint analysis of a series of experiments. We are not proposing to give an account of the methods for analysing groups of experiments. We are only giving two specific examples which may be used by the student; he will find the relevant procedures in Chapter 14 of Cochran and Cox, in Chapter VIII of "The Statistics of Bioassay" by C. I. Bliss and in Chapters 11, 12 and 13 of the forthcoming, "Statistics for Biologists" by C. I. Bliss.

Example 1 shows the results of a trial on 9 varieties of wheat (A, B, C, ... I) which was carried out by the Swiss Federal Research Station, Oerlikon, in a randomized blocks design in seven localities in 1960.

Example 2 has been published by Zelinder, Weber and Linder (1951). It gives the results of 12 experiments carried out by the Swiss Federal Institute of Forest Research in 1946. One of these experiments is considered on p. 28. Six such experiments were carried out on the first, six on the second day, according to the same design. Each day two experiments were performed on logs of small diameter (approx. 25 cm), two on logs of middle diameter (approx. 35 cm) and two on logs of big diameter (approx. 40 cm). In every case one of the two experiments concerned the lower part, the other the upper part of the stem. Three species of wood were used (rows of latin square), namely spruce (s), pine (p) and larch (ℓ), with (b) and without bark. For the details of the design of individual experiments see p. 27/28.

Example 1. Grain weight of 9 varieties of wheat in 10 kg/ha .

Realta	A	B	C	D	E	F	G	H	I	Total
I	174	362	274	236	267	348	334	248	299	2 542
II	298	373	415	346	241	395	321	320	281	2 990
III	317	442	480	203	298	211	417	411	437	3 216
IV	317	363	447	368	321	328	328	336	366	3 174
Total	1106	1540	1616	1153	1127	1282	1400	1315	1383	11 922
Haag										
I	386	382	380	386	421	388	409	394	379	3 525
II	184	338	348	385	398	338	363	346	381	3 081
III	344	350	356	396	381	330	359	350	359	3 225
IV	323	315	319	377	373	328	302	356	357	3 050
Total	1237	1385	1403	1544	1573	1384	1433	1446	1476	12 881
Kalchrain										
I	508	436	401	531	529	534	497	449	486	4 371
II	450	473	425	428	448	393	469	373	457	3 916
III	489	457	434	535	527	487	479	519	475	4 402
IV	530	437	388	548	613	551	565	527	576	4 735
Total	1977	1803	1648	2042	2117	1965	2010	1868	1994	17 424
Reckenholz										
I	554	479	529	584	537	555	521	490	505	4 754
II	448	542	551	597	510	592	462	532	520	4 754
III	528	578	593	560	484	601	472	455	549	4 820
IV	506	489	496	538	600	522	565	586	572	4 874
Total	2036	2088	2169	2279	2131	2270	2020	2063	2146	19 202
Hallau										
I	468	560	530	530	495	506	478	483	496	4 546
II	444	546	551	513	486	504	491	463	493	4 491
III	467	564	583	581	505	529	516	504	527	4 776
IV	472	606	595	532	517	524	476	488	457	4 667
Total	1851	2276	2259	2156	2003	2063	1961	1938	1973	18 480
Wallierhof										
I	364	333	418	300	376	316	360	408	360	3 235
II	378	436	421	433	397	365	362	403	379	3 574
III	395	426	436	417	408	410	396	382	319	3 589
IV	394	433	487	411	414	380	379	424	382	3 704
Total	1531	1628	1762	1561	1595	1471	1497	1617	1440	14 102
Grünenmatt										
I	432	338	339	387	380	431	394	398	376	3 475
II	380	357	362	351	370	329	335	389	374	3 247
III	355	327	341	337	358	324	391	361	359	3 153
IV	346	306	303	328	318	341	331	349	330	2 952
Total	1513	1328	1345	1403	1426	1425	1451	1497	1439	12 827

Example 2. Cutting time in 1/100 minutes

Team	First day						Second day					
	1	2	3	4	5	6	1	2	3	4	5	6
Wood	Small diameter - Lower part of stem											
s	C 63	E109	D 98	B 75	A 46	F 49	C 71	E 88	D 80	B 80	A 39	F 50
p	B 68	F 62	E 78	A 60	D 40	C 42	B 76	F 56	E 70	A 55	D 37	C 43
ℓ	E127	A135	B125	F 73	C 61	D 74	E113	A116	B108	F 88	C 50	D 72
s b	F 88	D102	A125	C 86	E 61	B 60	F 84	D 95	A 96	C 60	E 47	B 56
p b	D 74	B100	C 83	E 64	F 43	A 56	D 85	B 92	C 61	E 61	F 30	A 46
ℓ b	A130	C120	F120	D113	B 60	E 98	A135	C111	F 97	D112	B 58	E 76
	Small diameter - Upper part of stem											
s	C 49	F 55	A 64	D 61	B 35	E 40	C 48	F 53	A 54	D 51	B 38	E 36
p	F 28	D 27	B 40	A 36	E 21	C 24	F 24	D 33	B 30	A 28	E 19	C 23
ℓ	A 87	B 95	E 64	F 42	C 34	D 45	A 75	B 63	E 50	F 41	C 35	D 42
s b	E 58	A 58	C 46	B 55	D 32	F 29	E 53	A 50	C 46	B 55	D 32	F 35
p b	D 35	E 35	F 30	C 24	A 22	B 24	D 34	E 29	F 26	C 22	A 21	B 26
ℓ b	B 72	C 59	D 68	E 66	F 29	A 45	B 64	C 51	D 58	E 54	F 28	A 41
	Middle diameter - Lower part of stem											
s	E150	C128	F122	B121	A 72	D 89	E152	C122	F111	B147	A 76	D 92
p	C 96	F102	A130	E 92	D 61	B 88	C102	F 94	A105	E114	D 62	B 84
ℓ	A176	D190	B177	C118	F 82	E114	A195	D166	B162	C138	F 73	E105
s b	F118	E174	D127	A108	B 68	C 82	F133	E152	D130	A140	B 84	C 77
p b	D121	B143	E113	F 62	C 55	A 81	D127	B134	E 99	F 84	C 54	A 83
ℓ b	B189	A205	C145	D135	E 93	F104	B222	A210	C147	D165	E 88	F106
	Middle diameter - Upper part of stem											
s	D 77	E 86	C 59	A 63	F 36	B 53	D 72	E 54	C 50	A 60	F 31	B 49
p	E 69	F 60	B 70	D 62	C 33	A 46	E 63	F 45	B 56	D 55	C 28	A 44
ℓ	B107	A105	F 79	C 68	D 57	E 69	B 98	A 77	F 58	C 64	D 53	E 60
s b	C 70	D 90	A 71	E 80	B 46	F 45	C 66	D 72	A 74	E 70	B 43	F 42
p b	F 58	C 63	E 78	B 76	A 39	D 56	F 55	C 54	E 60	B 70	A 40	D 51
ℓ b	A121	B127	D116	F 91	E 54	C 64	A113	B112	D 93	F 80	E 45	C 59
	Big diameter - Lower part of stem											
s	C143	F171	A176	B188	E113	D132	C168	F168	A185	B187	E106	D126
p	E316	B338	C233	D243	A152	F140	E334	B402	C202	D289	A133	F137
ℓ	D264	A280	E278	F194	B142	C118	D239	A252	E198	F174	B114	C113
s b	B259	D263	F151	A198	C 98	E127	B265	C217	F158	A175	C103	E138
p b	A330	C287	D290	E303	F130	B192	A354	C245	D232	E276	F140	B177
ℓ b	F213	E350	B249	C224	D138	A150	F208	E284	B197	C198	D126	A143
	Big diameter - Upper part of stem											
s	C 84	D130	F 93	A111	B 62	E 81	C 86	D109	F 69	A101	B 62	E 68
p	F104	C140	A150	E154	D 87	B 98	F134	C122	A148	E135n	D 80	B111
ℓ	D147	E197	B136	F110	C 76	A113	D135	E120	B127	F112	C 72	A 95
s b	B167	A141	E120	C 93	F 59	D 79	B145	A130	E115	C 90	F 53	D 98
p b	E237	F166	D186	B180	A102	C106	E142	F110	D136	B156	A 72	C 80
ℓ b	A177	B232	C144	D159	E 93	F 85	A152	B150	C119	D150	E 80	F 87

BIBLIOGRAPHY

Books

Bliss, C. I. (1952) The statistics of bioassay. Academic Press, Inc. New York.

Bliss, C. I. (?) Statistics for biologists. McGraw Hill, New York.

Brownlee, K. A. (1949) Industrial experimentation. Her Majesty's Stationery Office, London. 4th ed.

Cochran, W. G. and Cox, G. M. (1950) Experimental designs. John Wiley and Sons, New York.

Cox, D. R. (1958) Planning of experiments. John Wiley and Sons, New York.

Davies, O. L. (Editor) (1954) Design and analysis of industrial experiments. Oliver and Boyd, Edinburgh.

Federer, W. J. (1955) Experimental desing. Macmillan, New York.

Finney, D. J. (1952) Statistical method in biological assay. Charles Griffin & Co. London.

Finey, D. J. (1960) An introduction to the theory of experimental design. Chicago University Press, Chicago.

Fisher, R. A. (1935) The design of experiments. Oliver and Boyd, Edinburgh.

Fisher, R. A. and Yates, F. (1938) Statistical tables. Oliver and Boyd, Edinburgh.

Goulden, C. H. (1939) Methods of statistical analysis. John Wiley & Sons, New York.

Kempthorne, Oscar (1952) The design and analysis of experiments. John Wiley & Sons, New York.

Linder, Arthur (1953) Planen und Auwerten von Versucken. Birkhäuser, Basel.

Quenouille, M. H. (1953) The design and analysis of experiment. Charles Griffin & Co., London.

Wishart, J. (1940) Field trials: their layout and statistical analysis. Imperial Bureau of Plant Breeding and Genetics, Cambridge.

Yates, F. (1937) The design and analysis of factorial experiments. Imperial Bureau of Soil Science.

Papers

Baxter, G. P. and O. W. Lundstedt (1940) A review of the atomic wieght of iodine. J. Amer. Chemical Society, 62, p. 1829.

Bose, R. C. and K. R. Nair (1939) Partially balanced incomplete block designs. Sankhya, 4, p. 337-372.

Bose, R. C. and T. Shimamoto (1952) Classification and analysis of partially balanced incomplete block designs with two associate classes. J. Amer. Statist. Assoc., 47, p. 151-184.

Box, G. E. P. and J. S. Hunter (1959) Condensed calculations for evolutionary operations programs. Technometrics, 1, p. 77-95.

Box, G. E. P. and K. B. Wilson (1946) On the experimental attainment of optimum conditions. Jour. Roy. Stat. Soc. B, 13, p. 1-45.

Finney, D. J. (1945) The fractional replication of factorial experiments. Annals of Eugenics, 12, p. 291-301.

Messikommer, B. H. (1960) Die Anwendung der Box-Wilsonscheu Methode in der chemichen industrie. Unternehmungsforschung, 4, p. 112-137.

Nair, K. R. (1952) Analysis of partially balanced incomplete block designs illustrated on the simple square and rectangular lattices. Biometrcs, 8, p. 122-152.

Rao, C. R. (1947) General methods of analysis for incomplete block designs. J. Amer. Statist. Assoc., 42, p. 541-561.

Stevens, W. L. (1948) Statistical analysis of a non-orthogonal trifactorial experiment. Biometrika, 35, p. 345-387.

Tukey, J. R. (1949) One degree of freedom for non-additivity. Biometrics, 5, p. 232-242.

Tukey, J. W. (1955) Query 113. Biometrics, 11, p. 111-113.

Yates, F. (1935) Complex experiments. Journ. Roy. Statist. Soc. Suppl., 2, p. 181-247.

Yates, F. (1936) A new method of arranging variety trials inolving a large number of varieties. Jour. Agr. Sci., 26, p. 424-455.

Yates F. (1936) Incomplete randomized blocks. Annals of Eugenics, 7,p.121-140.

Yates, F. (1937) A further note on the arrangement of variety trials: quasi-latin squares. Annals of Eugenics, 7, p. 319-332.

Yates, F. (1939) The recovery of inter-block information in variety trials arranged in three-dimensional lattices. Annals of Eugenics, 9, p. 136-156.

Yates, F. (1940) The recovery of inter-block information in balanced incomplete block designs. Annals of Eugenics, 10, p. 317-325.

Youden, W. J. (1937) Use of incomplete block replications in estimating tobacco-mosaic virus. Contr. Boyce Thompson Instl, 9, p. 41-48.

Youden, W. J. (1940) Experimental designs to increase accuracy of greenhouse studies. Contr. Boyce Thomspon Inst., 11, p. 219-228.

Zehnder, J., Weber, A. et A. Linder (1951) Etude du rendemend des scies par les methodes statistiques. Ann. Inst. fed. recherches forestieses, 27, p. 1-18.