Experiment

In the scientific method, an **experiment** (Latin: *ex- periri*, "of (or *from*) trying") is a set of observations performed in the context of solving a particular problem or question, to support or falsify a hypothesis or research concerning phenomena. The experiment is a cornerstone in the Empirical approach to acquiring deeper knowledge about the physical world.

Design of experiments

1. That the independent variable is the only factor that varies systematically in the experiment; in other words, that the experiment is appropriately controlled - that confounding variables are eliminated; and
2. That the dependent variable truly reflects the phenomenon under study (a question of validity) and that the variable can be measured accurately (i.e., that various types of experimental error, such as measurement error can be eliminated).

In a pure application of the scientific method, hypotheses are tested by critical experiments: ones that can falsify the hypothesis in the case of a non-result (i.e., an experiment showing that the independent variable did not affect the dependent variable as predicted). Such pure applications are rare, however, in part because a result can sometimes be challenged on the basis that an experiment was not sufficiently controlled, that the dependent variable was not valid, or that various forms of error compromised the experiment. The scientific method, as a result, builds in the need for reproducibility (usually termed **replication**) and convergent evidence.

The design of experiments attempts to balance the requirements and limitations of the field of science in which one works so that the experiment can provide the best conclusion about the hypothesis being tested.

In some sciences, such as physics and chemistry, it is relatively easy to meet the requirements that all measurements be made objectively, and that all conditions can be kept controlled across experimental trials. On the other hand, in other cases such as biology, and medicine, it is often hard to ensure that the conditions of an experiment are performed consistently; and in the social sciences, it may even be difficult to determine a method for measuring the outcomes of an experiment in an objective manner.

For this reason, sciences such as physics and several other fields of natural science are sometimes informally referred to as "hard sciences", while social sciences are sometimes informally referred to as "soft sciences"; in an attempt to capture the idea that objective measurements are often far easier in the former, and far more difficult in the latter.

In addition, in the social sciences, the requirement for a "controlled situation" may actually work against the utility of the hypothesis in a more general situation. When the desire is to test a hypothesis that works "in general", an experiment may have a great deal of *internal validity*, in the sense that it is valid in a highly controlled situation, while at

the same time lack *external validity* when the results of the experiment are applied to a real world situation. One of the reasons why this may happen is the Hawthorne effect; another is that partial equilibrium effects may not persist in general equilibrium.

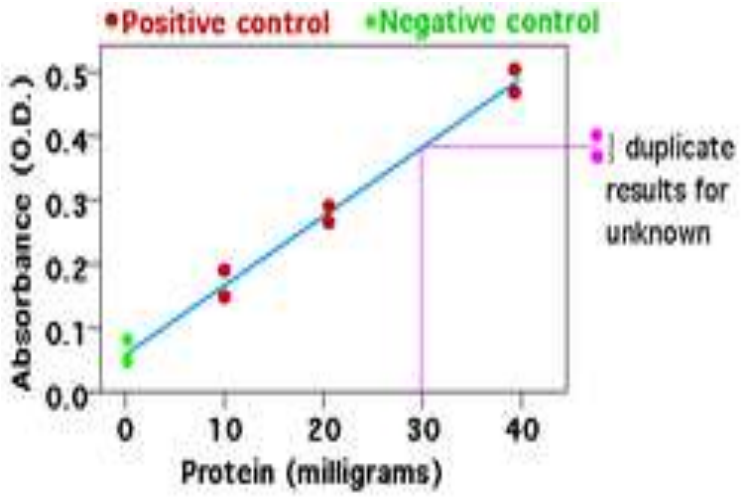
As a result of these considerations, experimental design in the "hard" sciences tends to focus on the elimination of extraneous effects, while experimental design in the "soft" sciences focuses more on the problems of external validity, often through the use of statistical methods. Occasionally events occur naturally from which scientific evidence can be drawn, which is the basis for natural experiments. In such cases the problem of the scientist is to evaluate the natural "design".

Controlled experiments

Many hypotheses in sciences such as physics can establish causality by noting that, until some phenomenon occurs, nothing happens; then when the phenomenon occurs, a second phenomenon is observed. But often in science, this situation is difficult to obtain.

For example, in the old joke, someone claims that they are snapping their fingers "to keep the tigers away"; and justifies this behavior by saying "see - its working!" While this "experiment" does not *falsify* the hypothesis "snapping fingers keeps the tigers away", it does not really support the hypothesis - *not* snapping your fingers keeps the tigers away as well.

To demonstrate a cause and effect hypothesis, an experiment must often show that, for example, a phenomenon occurs after a certain treatment is given to a subject, and that the phenomenon does *not* occur in the *absence* of the treatment.



Standard curve

A *controlled* experiment generally compares the results obtained from an experimental sample against a *control* sample, which is practically identical to the experimental sample except for the one aspect whose effect is being tested. A good example would be a drug trial. The sample or group receiving the drug would be the experimental one; and the one

receiving the placebo would be the control one. In many laboratory experiments it is good practice to have several replicate samples for the test being performed and have both a positive control and a negative control. The results from replicate samples can often be averaged, or if one of the replicates is obviously inconsistent with the results from the other samples, it can be discarded as being the result of an experimental error (some step of the test procedure may have been mistakenly omitted for that sample). Most often, tests are done in duplicate or triplicate. A positive control is a procedure that is very similar to the actual experimental test but which is known from previous experience to give a positive result. A negative control is known to give a negative result. The positive control confirms that the basic conditions of the experiment were able to produce a positive result, even if none of the actual experimental samples produce a positive result. The negative control demonstrates the base-line result obtained when a test does not produce a measurable positive result; often the value of the negative control is treated as a "background" value to be subtracted from the test sample results. Sometimes the positive control takes the form of a standard curve.

An example that is often used in teaching laboratories is a controlled protein assay. Students might be given a fluid sample containing an unknown (to the student) amount of protein. It is their job to correctly perform a controlled experiment in which they determine the concentration of protein in fluid sample (usually called the "unknown sample"). The teaching lab would be equipped with a protein standard solution with a known protein concentration. Students could make several positive control samples containing various dilutions of the protein standard. Negative control samples would contain all of the reagents for the protein assay but no protein. In this example, all samples are performed in duplicate. The assay is a colorimetric assay in which a spectrophotometer can measure the amount of protein in samples by detecting a colored complex formed by the interaction of protein molecules and molecules of an added dye. In the illustration, the results for the diluted test samples can be compared to the results of the standard curve (the blue line in the illustration) in order to determine an estimate of the amount of protein in the unknown sample.

Controlled experiments can be performed when it is difficult to exactly control all the conditions in an experiment. In this case, the experiment begins by creating two or more sample groups that are *probabilistically equivalent*, which means that measurements of traits should be similar among the groups and that the groups should respond in the same manner if given the same treatment. This equivalency is determined by statistical methods that take into account the amount of variation between individuals and the number of individuals in each group. In fields such as microbiology and chemistry, where there is very little variation between individuals and the group size is easily in the millions, these statistical methods are often bypassed and simply splitting a solution into equal parts is assumed to produce identical sample groups.

Once equivalent groups have been formed, the experimenter tries to treat them identically except for the one *variable* that he or she wishes to isolate. Human experimentation requires special safeguards against outside variables such as the *placebo effect*. Such experiments are generally *double blind*, meaning that neither the volunteer nor the

researcher knows which individuals are in the control group or the experimental group until after all of the data has been collected. This ensures that any effects on the volunteer are due to the treatment itself and are not a response to the knowledge that he is being treated.

In human experiments, a subject (person) may be given a stimulus to which he or she should respond. The goal of the experiment is to measure the response to a given stimulus.

Natural experiments

The term "experiment" usually implies a controlled experiment, but sometimes controlled experiments are prohibitively difficult or impossible. In this case researchers resort to *natural experiments*, also called *quasi-experiments*. Natural experiments rely solely on observations of the variables of the system under study, rather than manipulation of just one or a few variables as occurs in controlled experiments. To the degree possible, they attempt to collect data for the system in such a way that contribution from all variables can be determined, and where the effects of variation in certain variables remain approximately constant so that the effects of other variables can be discerned. The degree to which this is possible depends on the observed correlation between explanatory variables in the observed data. When these variables are *not* well correlated, natural experiments can approach the power of controlled experiments. Usually, however, there is some correlation between these variables, which reduces the reliability of natural experiments relative to what could be concluded if a controlled experiment were performed. Also, because natural experiments usually take place in uncontrolled environments, variables from undetected sources are neither measured nor held constant, and these may produce ilusory correlations in variables under study.

Much research in several important science disciplines, including economics, political science, geology, paleontology, ecology, meteorology, and astronomy, relies on quasi-experiments. For example, in astronomy it is clearly impossible, when testing the hypothesis "suns are collapsed clouds of hydrogen", to start out with a giant cloud of hydrogen, and then perform the experiment of waiting a few billion years for it to form a sun. However, by observing various clouds of hydrogen in various states of collapse, and other implications of the hypothesis (for example, the presence of various spectral emissions from the light of stars), we can collect data we require to support the hypothesis. An early example of this type of experiment was the first verification in the 1600s that light does not travel from place to place instantaneously, but instead has a measurable speed. Observation of the appearance of the moons of Jupiter were slightly delayed when Jupiter was farther from Earth, as opposed to when Jupiter was closer to Earth; and this phenomenon was used to demonstrate that the difference in the time of appearance of the moons was consistent with a measurable speed of light

Observational studies

Observational studies are very much like controlled experiments except that they lack probabilistic equivalency between groups. These types of experiments often arise in the area of medicine where, for ethical reasons, it is not possible to create a truly controlled group. For example, one would not want to deny all forms of treatment for a life-threatening disease from one group of patients to evaluate the effectiveness of another treatment on a different group of patients. The results of observational studies are considered much less convincing than those of designed experiments, as they are much more prone to selection bias. Researchers attempt to compensate for this with complicated statistical methods such as propensity score matching methods.

Field experiments

Field experiments are so named in order to draw a contrast with laboratory experiments. Often used in the social sciences, and especially in economic analyses of education and health interventions, field experiments have the advantage that outcomes are observed in a natural setting rather than in a contrived laboratory environment. However, like natural experiments, field experiments suffer from the possibility of contamination: experimental conditions can be controlled with m

2.3Basic Principles of an Experimental Design

Prof. Ronald A. Fisher pioneered the study of experimental designs; describe the three basic principles of the experimental design, which have been discussed below.

Replication

The repetition of the treatments under investigation is known as replication. We know that the variation in soil fertility over the field (in field experiment) and the productivity over the experimental material in other experiment is so great that all the treatments used in experiment do not get equal chance of experiencing every type of environment, if one treatment is allotted to one plot (experimental unit) only.

Some of the treatments, which are applied to plots that are more fertile, will be in an advantageous position as compared with those, which are applied to less fertile plots, and hence the yields due to the treatments may different. However, this difference in the yields of the two treatments will not be the real differential effect of the same as they are affected by varying soil fertility of the field. It may be just possible that this difference in the two yields will be due to the difference in soil fertility in the two plots and not due to the difference in the treatments.

An experimenter resorts to replication in order to average out the influence of the chance factors on different experimental units. Thus, the repetition of treatments results in more reliable estimate than is possible with a single observation.

The following are the chief advantages of replication

a. At the first instance, replication serves to reduce experimental error and thus enables us to obtain more precise estimates of the treatment effects. From statistical theory, we know that the standard error of the mean of a sample of size n is $\sigma/\sqrt{n}$ where σ the standard deviation of the population. Thus if a treatment is replicated r times, then the S.E. of its mean effect is $\sigma/\sqrt{r}$, where σ is the standard deviation of the individual plot is estimated from the 'error variance' thus the precision of the experiment is inversely proportional to the square root of the replications.
b. The most important purpose of replication is to provide an estimate of the experiment error without which we cannot text the significance of the difference between any two treatments, or determine the length of the confidence interval. The estimate the experimental error is obtained by considering the differences in the plate receiving the same treatment in different replications and these is no other alternatives of obtaining this estimate.
c. It is desirable to have as much uniformity or homogeneity as possible with in each replication but it is not important to have a great deal of uniformity between replications.
d. The adequate number of replications for various treatments in an experiment depends upon the knowledge of the variability of experimental material. e. g. fertility of soil infield experiment, which is rarely known and as such cannot be suggested in advance. A general rule is to get as many replications, which will, provide at least 12 d. f. for the error. The reason of this is that the fall of the values of F at 5% and 1% level of significance is not so rapid for the values of d. f. beyond 12 or so.

It is difficult to suggest an adequate number of replications for a particular number of treatments because that depends on the inherent variation from plot to plot for any particular character under study. Moreover, the size of an experiment is limited by lack of resources or by the conflicting claims of other experiments.
Hence, whatever the source of the experimental errors replication of the experiment steadily decreases the error associated with the difference between the average results for two treatments provided that randomization have been taken to ensure that one treatment is no more likely to be favored in any replicate than other. So that the errors affecting any treatment tend to cancel out as the number of replications is increased.
Therefore, the adequate number of replications for a particular number of treatments is very important but limited role in increasing the efficiency of the design. Number of replication, which is the key for reducing experimental error, is predictable from statistical theory. The basic quantity used to measure experimental error is the error variance per experimental unit. Which is defined as the expected value of the square of the error that affects the observation for a single experimental unit?

If σ^2 is the error variance per unit and $\overline{x_1}$ and $\overline{x_2}$ are the mean effects of two treatments both replicated times, then, variance of the difference between the meant for two treatment is given and the corresponding S.E. is given

$$\text{var}\left(\bar{x}_1 - \bar{x}_2\right) = \frac{2\sigma^2}{r} \text{ and the corresponding S.E. is given}$$

$$S.E.\left(\bar{x}_1 - \bar{x}_2\right) = \sqrt{\frac{2}{r}}\sigma$$

Then an approximate idea of the minimum number of replications required to detect the personable difference between two treatments at certain ($\propto$) level of significance is obtained by applying normal text for a large number of error degrees of freedom.

$$t = \frac{\left(\bar{x}_1 - \bar{x}_2\right)}{S.E.\left(\bar{x}_1 - \bar{x}_2\right)} = \frac{\left(\bar{x}_1 - \bar{x}_2\right)}{\sigma\sqrt{2/r}} t = \frac{d}{\sigma\sqrt{2/r}}$$

However, for small degrees of freedom, 't' is not normally distributed, but follow student's't' distribution. In that case the minimum number of replications required to detect a specified, difference'd' between the treatments at $\propto$% level if significance is given by

$$t_\alpha = \frac{d}{s\sqrt{2/r}} r = \frac{2t_\alpha{}^2 s^2}{d^2}$$

Where, s^2 is the error variance per unit and $t_\propto$ is the value of student's t at $\propto$% level of significance and corresponding to error d.f.

(For more details, see Cochran & Cox. Experimental design, Page 19)

Randomization

After the treatments and the experimental units are decided, the treatments are allotted to the experimental units at random to avoid any type of personal or subjective bias, which may be conscious or unconscious.

In other words, randomization is the process of assigning treatment of factor under investigation to experimental units in a purely chance manner. Randomization guarantees the elimination of systematic error this ensures validity of the results. It helps to have an objectives comparison among the treatments. It also ensures independence of the observations, which is so essential for drawing valid inference from the observations by applying appropriate statistical techniques.

In the principles of randomization, every experiment unit has the same chance of receiving any treatment.

Local Control

"Local Control" means control of all factor except the one about which investigation is done. Local control like replication is yet another device to control the variations due to extraneous sources and increase the precision of an experiment.

As a lower experimental error helps in detecting the smaller real differences between the treatments, it is desirable that it should be reduced as far as practically possible with our unduly increasing the number of replications or without interfering with the statistical requirement of randomness.

This is possible by making use of the principle that the adjacent areas are relatively more homogeneous than those widely separated. For instance, if an experimental field is heterogeneity with respect to soil fertility, then the field can be divided into smaller blocks such that plots with in each block tend to be more homogeneous. If the trees in a forest experiment are heterogeneous with respect to their heights, then to make any treatment comparison, the trees may have to be grouped into homogeneous groups with respect to their heights. The process of reducing the experimental error by dividing the relatively heterogeneous experimental area (field) into homogeneous blocks (due to physical contiguity as far as field experiments are concerned) is known as Local Control.

Completely Randomized Design (CRD)

- Formal, non- restriction, single factor design experimental design.
- Treatments are assigned completely at random so that each experimental unit has the same chance of receiving any one treatment.
- Here any difference among the experimental units receiving the same treatment is considered as the "*experimental error*".
- CRD is one in which all the experimental units are taken in a single group which is homogeneous as far as possible. For example, the entire field plots constituting the group having the same soil fertility, soil depth, soil texture, soil moisture etc. All the cows forming a group are of the same breed, same age, same weight, and same lactation etc.
- Commonly used when experimental units are homogeneous or an experimental area happens to be homogeneous**.**
- Involves only two principles of experimental design, they are replication & randomization.

Randomization & layout

- Whole experimental material is divided into n number of experimental units.
  $n = r \times t$, r is number of replication, t is number of treatments for equal replication
  $n = \sum r_i$, for unequal replication
- Assign the plots from 1 to n
- Assign the treatment to the experimental plots randomly

Let us take an example of CRD with four treatments A, B, C& D each replicated five times look like this.

A	D	A	B
C	B	C	A
D	A	C	B
C	B	D	C
D	B	D	A

Lay out of CRD (with equal replication)

Similarly, four treatments T1, T2, T3& T4 each are replicated 4,3,3,5 times respectively then layout of CRD have been 15 plots (units) as shown below:

T2	T2	T2
T3	T4	T3
T1	T1	T1
T4	T4	T1
T3	T4	T4

Layout of CRD (Unequal replication)

Mathematical model & statistical analysis

Its mathematical model & statistical analysis is analogous to the ANOVA of one-way classified data.

The linear model is,

$$y_{ij} = \mu + \alpha_i + e_{ij} \text{.....(1)} \quad i = 1,2,.....t.$$

$$j = 1,2.....r_i$$

y_{ij} = yield or response from the j^{th} unit receiving the i^{th} *treatment*

μ = general mean effect

α_i = effect due to i^{th} treatment

e_{ij} = error effect due to chance

$n = \sum r_i$ = total no. of experimental units

Assumptions

- All the observations are independent
- Different effects are additive in nature
- e_{ij} are identically & independently distributed N(0, σ_e^2)
- $\sum \propto_i = 0$, or $\propto_i \sim N(0, \sigma_\propto^2)$

Hypothesis

$H_0 : \mu_1 = \mu_2 = \mu_3 = \ldots\ldots\ldots\ldots\ldots = \mu_t = \mu$

i.e. $\propto_1 = \propto_2 = \propto_3 = \ldots\ldots\ldots\ldots\ldots = \propto_t = 0$

$H_1 : \mu_1 \neq \mu_2 \neq \mu_3 \neq \ldots\ldots\ldots\ldots\ldots \neq \mu_t$

i.e. $\propto_1 \neq \propto_2 \neq \propto \neq \ldots\ldots\ldots\ldots\ldots \neq \propto_t \neq 0$

Advantages

- Simple & easy layout
- Utilization of whole experimental material
- Complete flexible
- Simple analysis
- Missing data create no problem in analysis
- Specially suitable even some units are destroyed or failed to respond
- Mostly used in laboratory or green house experiment
- Gives maximum degrees of freedom for experimental error

Disadvantage

- Only suitable for small number of treatments
- Homogeneous experimental units can rarely obtained
- Less informative for heterogeneous fields
- Seldom suitable for field experiments
- All extraneous variations included in the error variation (residual variation)

Example of CRD

1. A set of data involving four tropical feed stuffs A, B, C and D tried on 20 chicks is given below. All the 20 chicks are treated alike in all respects except the feeding treatments and each feeding treatment is given to 5 chicks. Analyse the data to test the mean yield of each treatment differ significantly for =.05 using LSD and DNMRT test.

Weight gain of baby chicks fed on different feeding materials composed of tropical feed stuffs:

Observations						Ti	$\bar{Y}_{i.}$
A	55	49	42	21	52	219	43.8
B	61	112	30	89	63	355	71
C	42	97	81	95	92	407	81.4
D	169	137	169	85	154	714	142.8
					Grand total	G=1,695	

H_0: $\mu_1 = \mu_2 = \mu_3 = \mu 4 = \mu$ i.e. the treatment effects are same. In other words, all the treatments (A, B, C and D) are alike as regards their effects on increase in weight.

H_1 : $\mu_1 \neq \mu_2 \neq \mu_3 \neq \mu 4$

First of all, we have to compute C.F., TSS, SST and SSE and prepare ANOVA table.

TSS= Total sums of square=$\sum\sum Yij^2$ –C.F.=37,793.75 where C.F=correction factor=G^2/n

SST=Sums of square due to treatments=$\sum Ti^2/ri$ –C.F.=26,234.95

SSE= Sums of square due to error=TSS-SST=11,558.80

ANOVA TABLE

Source of variation	d.f.	SS	MSS=SS/d.f	Fcal
Treatments	3	26234.95	8744.98	Fcal=MST/MSE =12.105
Error	16	11558.80	722.42	
Total	19	37793.75		

F tab=F(3,16) at 5% level of significance=3.o6

Decision: Fcal>Ftab, Ho is rejected i.e. treatments A, B, C and D differ significantly. Or all the treatments (A, B, C and D) are not alike as regards their effects on increase in weight. Since Ho is rejected in this case, we proceed further to find out which of the treatments means differ significantly. For this we use LSD and DNMRT as follows:

The treatments mean effects, arranged in ascending order of magnitude as below:

Treatment	$\bar{Y}_{i.}$	$\bar{Y}_{i.}$-$\bar{Y}_{A}$	$\bar{Y}_{i.}$-$\bar{Y}_{B}$	$\bar{Y}_{i.}$-$\bar{Y}_{C}$
D	142.8	99.0	71.8	61.4
C	81.4	37.6	10.4	-
B	71.0	27.2	-	-
A	43.8	-	-	-

Using LSD,

For equal replication

LSD=r(2, α ,γ)√MSE/r =r(2,.05.16) √722.42/5=36.06

Where, r(2, α ,γ)= Special range value of rank of order 2 at .o5 level of significance with error degree of freedom(d.f)16 = r(2,.05.16) (see table)

MSE= Mean sum of squares due to error

r=Replication of each treatment

Decision: Thus, by comparing LSD value and above table values, the treatment pairs, (D, A),(C,A),(D,B) and(D,C) are judged significantly different.

Similarly, DNMRT test can be used for rank of order (p) of 2,3 and 4.

Note: Note that a non- significant F in the analysis of variance indicates the failure of the experiment to detect any differences among treatments. It doesn't, in any way, prove that all treatments are the same, because the failure to detect treatment differences based on non- significant F test, could be the result of either a very small or no difference among the treatments or due to large experimental error, or both. Thus, whenever the F test is non- significant, the researcher should examine the size of the experimental error and the numerical differences among the treatment means. If both values are large, the trial may be repeated and efforts made to reduce the experimental error so that the differences among treatments, if any, can be detected. On the other hand, if both values are small, the differences among treatments are probably too small to be of any economic value and, thus, no additional trials are needed.

Randomized Block Design (RBD)

- Most widely used experimental designs in forestry & biological research.
- Especially suitable for field experiments where the number of treatments is not large and there exists a remarkable factor based on which homogeneous sets of experimental units can be identified.
- The primary distinguished feature of the RBD is the presence of block of equal size each of which contains all the treatments.

Blocking Technique

- Grouping the experimental units into blocks such that variability within a block is minimized & variability among the blocks is maximized..

Two important points should be kept in mind while blocking

- Selection of the source of variability
- Block shape
    - ✓ Unidirectional – use long & narrow
    - ✓ Two directional – ignore the weaker one
    - ✓ Equally strong – square blocks

"**Randomization & layout**

The randomization process for RBD is applied separately & independently to each block. For example let us take a field experiments with 6 treatments A, B, C, D, E & F and 3 replications.

Block I	Block II	Block III
C	A	F
D	E	D

F	F	C
E	C	A
B	D	B
A	B	E

Gradient

Layout of RCBD

Mathematical model & statistical analysis

Its mathematical model & statistical analysis is analogous to the ANOVA of two-way classified data.

The linear model is,

$$y_{ij} = \mu + \alpha_i + \beta_j + e_{ij} \ldots\ldots\ldots\ldots (1) \quad i = 1,2,\ldots\ldots t.$$

$$j = 1,2\ldots\ldots r$$

y_{ij} = yield or response from the j^{th} block receiving the i^{th} treatment

μ = general mean effect

α_i = effect due to i^{th} treatment

β_j = effect due to j^{th} block

e_{ij} = error effect due to chance

Assumptions

- All the observations are independent
- Different effects are additive in nature
- e_{ij} are identically & independently distributed N(0, σ_e^2)
- $\sum \propto_i = 0$, $\sum \beta_j = 0$, or $\propto_i \sim N(0, \sigma_\propto^2)$,$\beta_j \sim N(0, \sigma_\beta^2)$

Hypothesis

$H_{0\propto}$: $\mu_{1.} = \mu_{2.} = \mu_{3.} = \ldots\ldots\ldots\ldots\ldots\ldots = \mu_{t.} = \mu$

i.e. $\propto_1 = \propto_2 = \propto_3 = \ldots\ldots\ldots\ldots\ldots\ldots = \propto_t = 0$

$H_{1\propto}$: $\mu_{1.} \neq \mu_{2.} \neq \mu_{3.} \neq \ldots\ldots\ldots\ldots\ldots\ldots \neq \mu_{t.} \neq \mu$

i.e. $\propto_1 \neq \propto_2 \neq \propto \neq \ldots\ldots\ldots\ldots\ldots\ldots \neq \propto_t \neq 0$

$H_{0\beta}$: $\mu_{.1} = \mu_{.2} = \mu_{.3} = \ldots\ldots\ldots\ldots\ldots\ldots = \mu_{.t} = \mu$

i.e. $\beta_1 = \beta_2 = \beta_3 = \ldots\ldots\ldots\ldots\ldots\ldots = \beta_t = 0$

$H_{1\beta}$: $\mu_{.1} \neq \mu_{.2} \neq \mu_{.3} \neq \ldots\ldots\ldots\ldots\ldots\ldots \neq \mu_{.t}$

i.e. $\beta_1 \neq \beta_2 \neq \beta_3 \neq \ldots\ldots\ldots\ldots\ldots\ldots \neq \beta_t \neq 0$

The principle advantage of RBD

- Blocking reduce error variance and provides more accurate result.

- Any number of treatments & any number of replication may be included. This is the most popular design in view of simplicity, flexibility & validity. No other design has been used as frequently as RBD
- Control treatments can easily be included without causing any complication in the analysis of the data.

Disadvantages

- If the blocks are not homogeneous the error term will be large
- It cannot accommodate large number of treatment since in this situation the homogeneity of blocks or groups is always in danger or hazard.

In many situations the criteria for blocking or grouping is not easily selectable.

Latin Square Design (LSD)

- Balanced two-way classification scheme with two superimposed blocking systems rows & columns.
- The number of rows and the number of columns must both be equal to number of treatments and each treatment occurs once in each row & once in each column.
- The principle use of LSD in forestry research is in nursery & glass house experiments
- Specially used when the variations are in two direction & perpendicular to each other. For example, if a forest researcher wants to estimate the effect of 4 different fertilizers (say) A, B, C & D in the growth of nursery seedlings of 4 different species (row) & 4 different age groups (column), he has to use the design LSD.

Randomization & layout

The process of randomization & layout for LSD is shown below for the experiment with four different treatments A, B, C & D cited above. In this experiment the researcher has to divide the total experimental land in $4\times4 = 16$ experimental units.

Step 1: The whole experimental area is divided into $4^2 = 16$ experimental units arranged in a square so that each row as well as each column contains 4 units.

Step 2: The 4 treatments are then allocated to these rows & columns in such a way that every treatment comes once & only once in each column. The randomization can be shown follows:

Gradient →

		Columns (age of seedlings)			
		1(3month)	2(4month)	3(5month)	4(6month)
Rows (species)	1	A	B	C	D
	2	B	C	D	A
	3	C	D	A	B
	4	D	A	B	C

Layout of LSD

Mathematical model & statistical analysis

$$y_{ij} = \mu + \alpha_i + \beta_j + \gamma_k + e_{ijk} \quad \ldots\ldots\ldots\ldots(1) \quad i = 1,2,\ldots\ldots t.$$
$$j = 1,2\ldots\ldots t.$$
$$k = 1,2,\ldots\ldots t$$

y_{ijk} = yield or response of the k^{th} treatment in i^{th} row & j^{th} column

μ = general mean effect

α_i = effect due to i^{th} row

β_j = effect due to j^{th} column

γ_k = effect due to k^{th} treatment

e_{ijk} = error effect due to chance

Assumptions

- All the observations are independent
- Different effects are additive in nature
- e_{ijk} are identically & independently distributed $N(0, \sigma_e^2)$
- $\sum \propto_i = 0, \sum \beta_j = 0, \sum \gamma_k = 0$, or $\propto_i \sim N(0, \sigma_\propto^2), \beta_j \sim N(0, \sigma_\beta^2), \gamma_k \sim N(0, \sigma_\gamma^2)$

Hypothesis

$H_{0\propto}$: $\mu_{1..} = \mu_{2..} = \mu_{3..} = \ldots\ldots\ldots\ldots\ldots\ldots = \mu_{t..} = \mu$

i.e. $\propto_1 = \propto_2 = \propto_3 = \ldots\ldots\ldots\ldots\ldots\ldots = \propto_t = 0$

$H_{1\propto}$: $\mu_{1..} \neq \mu_{2..} \neq \mu_{3..} \neq \ldots\ldots\ldots\ldots\ldots\ldots \neq \mu_{t..}$

i.e. $\propto_1 \neq \propto_2 \neq \propto \neq \ldots\ldots\ldots\ldots\ldots\ldots \neq \propto_t \neq 0$

$H_{0\beta}$: $\mu_{.1.} = \mu_{.2.} = \mu_{.3.} = \ldots\ldots\ldots\ldots\ldots\ldots = \mu_{.t.} = \mu$

i.e. $\beta_1 = \beta_2 = \beta_3 = \ldots\ldots\ldots\ldots\ldots\ldots = \beta_t = 0$

$H_{1\beta}$: $\mu_{.1.} \neq \mu_{.2.} \neq \mu_{.3.} \neq \ldots\ldots\ldots\ldots\ldots\ldots \neq \mu_{.t.}$

i.e. $\beta_1 \neq \beta_2 \neq \beta_3 \neq \ldots\ldots\ldots\ldots\ldots\ldots \neq \beta_t \neq 0$

$H_{0\gamma}$: $\mu_{..1} = \mu_{..2} = \mu_{..3} = \ldots\ldots\ldots\ldots\ldots\ldots = \mu_{..t} = \mu$

i.e. $\gamma_1 = \gamma_2 = \gamma_3 = \ldots\ldots\ldots\ldots\ldots\ldots = \gamma_t = 0$

$H_{1\gamma}$: $\mu_{..1} \neq \mu_{..2} \neq \mu_{..3} \neq \ldots\ldots\ldots\ldots\ldots\ldots \neq \mu_{..t}$

i.e. $\gamma_1 \neq \gamma_2 \neq \gamma_3 \neq \ldots\ldots\ldots\ldots\ldots\ldots \neq \gamma_t \neq 0$

Advantages

- Because of the two way blocking or stratification LSD controls more of the variations than CRD & RBD.
- Greater sensitivity – row & column variation is removed from error
- Easy analysis
- Several LSD of the same size may be combined and it is suitable for 5-9 no. of treatments

Disadvantages

- To obtain equal number of row, column & treatment is often difficult.

When number of treatment is large, design become impracticable because of the large number of replication required and when number of treatment is small, design gives few error degrees of freedom.

Multiple –comparison:

When an ANOVA F-test reject Ho shows that there is significance different among treatment means but it doesn't specify which pairs of treatments that differ significantly. To obtain this information, procedure for comparing treatment means like least significant difference (LSD) test and Duncan's New multiple range test (DNMRT) are needed for Biological research.

Least significant difference (LSD) test:

It is simple and most commonly used procedure for making pair comparisons. The procedure provides for a single LSD value, at a prescribed level of significance, which serves as the boundary between significant and non-significant differences between any pair of treatment means. That is two treatments are declared significantly different at a prescribed level of significance if their difference exceeds the computed LSD value, otherwise not.

LSDtest is used when i. F-test shows rejection of Ho

ii. Numbers of treatments is not too large i.e. less than 6

Formula:(i) For equal replication

$LSD = r(2, \alpha, \gamma)\sqrt{MSE/r}$

Where, $r(2, \alpha, \gamma)$= Special range value of rank 2 at α level of significance with error degree of freedom(d.f) γ ,

MSE= Mean sum of squares due to error

r=Replication of each treatment

$MSE/r = \overline{Syi}$=Standard error of single treatment mean

ii. For unequal replication

$$LSD = \frac{r(2, \alpha, \gamma)}{\sqrt{2}}\sqrt{MSE(1/ri+1/rj)}$$

Where ri and rj denote the replication of ith and jth treatments

Decision: If the observed value $(\overline{Yi}-\overline{Yj})$ is greater than computed LSD, Ho is rejected i.e. related treatment pairs are judged to be significantly different.

DNMRT test: For experiments that require the evaluation of all possible pairs of treatment means, LSD isn't suitable. It is especially true when the total no. of treatments is large i.e >6.Procedure of DNMRT is similar as that of LSD. In LSD, only single LSD value is required for any pair comparison but DNMRT requires a series of values.

This test is most widely used of several multiple range tests available. This test uses ranges of rank 2,3…….t according to the ranking of pairs of means when they are arranged.

Formula:(i) For equal replication

$$DNMRT=r(p, \alpha, \gamma)\sqrt{MSE/r}$$

Where, r(p, α ,γ)= Special range value of rank p at α level of significance with error degree of freedom(d.f) γ ,
MSE= Mean sum of squares due to error
r=Replication of each treatment

$MSE/r = \overline{Syi}$=Standard error of single treatment mean

ii. For unequal replication

$$DNMRT=\frac{r(2, \alpha, \gamma)}{\sqrt{2}}\sqrt{MSE(1/ri+1/rj)}$$

Where ri and rj denote the no.of replication of ith and jth treatments respectively.

Decision: If the observed value ($\overline{Yi}$-$\overline{Yj}$) is greater than computed DNMRT, Ho is rejected i.e. related treatment pairs are judged to be significantly different.

Factorial Experiment (multiple -factor Experiment)

An experiment in which the treatment consists of all possible combinations of the selected levels in two or more factors is referred to as factorial experiment.

They are especially important in several economic and social phenomena where usually a large no. of factors affect a particular problem. In this design the treatments consist of combination of different levels of two or more factors. A fully factorial experiment is a highly efficient way of obtaining information on each of the treatment factor & on the extent to which they interact with each other.

2^2 Factorial designs:

For example, suppose we want to find out the effect of two different fertilizer (factors) nitrogen (N) & potash (K) on the production of certain crop by using the two different amount of each fertilizer as 20 & 25kg.Here we can use factorial experiment.

Let each level of each factor is denoted as N_0, N_1, K_0, K_1 then we obtain four-treatment combination as shown below;

Factor K		Factor N	
		20kg(N_0)	25kg(N_1)
	20kg(K_0)	N_0K_0	N_1K_0
	25kg (K_1)	N_0K_1	$N_1 K_1$

These 4 treatment combinations can be compared by laying out the experiment in

(i) R.B.D., with r replicates (say), each replicate containing 4 units, or (ii) 4x4 L.S.D., and ANOVA can be carried out accordingly. In the above cases, there are 3 d.f. associated with the treatment effects. In factorial experiment our main objective is to carry out separate tests for the main effects N, K and interaction NK.

Replication I	Replication II	Replication III
N_0K_0	N_1K_0	N_0K_0
N_0K_1	$N_1 K_1$	$N_1 K_1$

N_1K_0	N_0K_1	N_1K_0
N_1 K_1	N_0K_0	N_0K_1

A sample layout of 2x2 factorial experiment in a RCBD with 3 replications

Similarly, ANOVA of a two- factor experiment on bamboo involving two levels of spacing (Factor A) and three levels of age at planting (Factor B) laid out in RCBD with three replications is given below

Age at planting (month)(FactorB) → Levels ↓	Spacing (m)(Factor A)	
	10 mx10 m (a1)	12m x 12m (a2)
6 (b1)	a1b1	a2b1
12 (b2)	a1b2	a2b2
24 (b3)	a1b3	a2b3

The 2x3 factorial treatment combinations of two levels of spacing and three levels of age

Advantages:

- Flexibility- any number of factors as well as any number of levels can be used subject to the available resources.
- Factorial treatments may be used in any experimental design
- Interaction of the treatments can be investigation
- In the absence of interaction number of replication increases.

Disadvantages

- More complex if any observation is missing
- If interaction is present, the results are more difficult to interpret
- If the number of factors & the levels are large, the size of the experiment is large

Chi-square test (χ2- test)

The tests of significance such as Z, t, F etc. tests are based on the assumption that the samples are drawn from a normal population i.e. we make assumptions about the population parameters. Such tests are called parametric tests. However, in many situations it is not possible to make dependable assumption about the parent population from which samples are drawn. To study these problems some tests called **non-parametric tests** which do not require any assumptions about the parameters are derived. Therefore, **χ2- test** is one of the most important non –parametric test that does not require any assumptions about the population parameters. Thus it is most commonly used to test the hypothesis concerning the frequency distribution of one or more classes. It is based on attribute data concerned with a finite numbers of discrete classes. The most common types of attribute data are those having two classes, which consist of the presence or absence of an attribute such as male or female, success or failure, effective or ineffective and dead or alive and there may be more than two classifications in many cases.
Chi-square tests for differences between categorical variables (i.e. nominal or ordinal). There are both ‘‘one-way” and ‘’two way” chi-square procedures.

Conditions for applying χ2 –test

- Observations recorded and used are collected on a random basis.
- The overall number of items must also be reasonably large. It should normally be at least 50 i.e. N≥50.
- Each of the observations of the sample must be independent of each other.
- The expected frequency of any item or cell should not be less than 5. If it is less than 5, the frequencies of adjacent items or cells should be pooled together in order to make it 5 or more than 5.In this case, degree of freedom decreases.
- Constraints on the cell frequencies, if any, should be linear, e.g.∑0i=∑Ei

Application of χ2 –test

1. **χ2 –test as a test of goodness of fit**

If we have a set of observed (actual) frequencies under some experiment and we want to test whether a particular distribution like uniform or other standard theoretical distribution like binomial, Poisson, normal distribution support the hypothesis (fits the data), then **χ2 –test** is used to test the goodness of fit of that distribution.. In χ2 –test as a test of goodness of fit, we used one way classification tables of observed frequencies in a single row or column.
Under this, **Ho**: The …….. distribution fits the data i.e. there is no significance difference between observed (actual or experiment) and expected (theoretical) frequencies.
H1: The ……..distribution does not fit the data i.e. there is significance difference between observed (actual or experiment) and expected (theoretical) frequencies.
Karl Pearson proved that the statistic,

$\chi 2 = \sum(Oi-Ei)^2/Ei \sim \chi 2$ with d.f = n-k (1 d.f .is lost due to $\sum Oi=\sum Ei$) gives the magnitude of the discrepancy between theory and experiment and closer the values of Oi and Ei, smaller the value of $\chi 2$ will be and the hypothesized distribution best fits the data.

Result: If $\chi 2$ cal< $\chi 2$ tab, Ho is accepted otherwise rejected.

Example: Table shows the no. of insects species collected from an undisturbed Bardia National park in different months. To test whether there is any significance different between the no. of insect species found in different month.

Month: Jan Feb March April May June July Aug Sep Octo Nov Dec

No. of species: 67 115 118 72 67 77 75 63 42 24 32 52

Solution: Ho: The diversity in terms of no. of insect species is the same in all months in the Bardia National Park. Then, the statistic

$\chi 2 = \sum(Oi-Ei)^2/Ei$ with n-1 d.f. and

Under uniform distribution, expected frequencies are calculated by, Ei $=\sum Oi/n=804/12=67$

Table:

Month	No. of species(Oi)	Ei	$(Oi- Ei)^2/Ei$
Jan	67	67	0
Feb	115	67	34.38
March	118	67	38.82
April	72	67	
May	67	67	
June	77	67	
July	75	67	
Aug	63	67	
Sep	42	67	
Octo	24	67	
Nov	32	67	
Dec	52	67	
	$=\sum Oi=804$	67	$\sum(Oi-Ei)^2/Ei= \chi 2 =113.84$

From above table, Calculated $\chi 2$ =134.84

Tabulated $\chi 2$ at 0.05 with.12-1=11 d.f=19.7

Decision: Calculated $\chi 2$> Tabulated $\chi 2$, H_0 is rejected and we conclude that the occurrence of the no. of insect species in different months is not same in Bardia NP.

2. **$\chi 2$ –test as a test of independent of attributes (Contingency table/Cross tab method)**

In this case, when the observed frequencies occupy r rows and c columns, a two way classification table called **rxc** contingency table. A two-way chi- square is used if two categorical variables are to be compared. Under this test,

Ho: There is no association between attributes or two attributes are independently distributed.

H1: There is association between attributes or two attributes are dependent.

Then, test statistic

$\chi 2=\sum\sum(Oij-Eij)^2/Eij$

Eij values are computed as follows

Let, 2x3 contingency table,

Attribute A Attribute B	A1	A2	A3	Row total (RT)
B1	a	b	c	a+b+c
B2	d	e	f	d+e+f
Column total (CT)	a+d	b+e	c+f	N=a+b+c+d+e+f

Then, expected values of Eij can be calculated by the following procedure,

Expected value of cell A1B1 or E (a)=Row total x Column total /N =RTxCT/N= (a+b+c) x (a+d)/N

Expected value of cell A2B1 or E (b) =Row total x Column total /N =RTxCT/N= (a+b+c) x (b+e)/N

Similarly others can be calculated and chi –square value can be obtained by using above formula.

Tabulated $\chi 2$ at(r-1) (c-1) d.f. =2d.f. =?

Decision: If cal $\chi 2$ > tab$\chi 2$,Ho is rejected i.e. two attributes A and B are dependent (associated).

Example: Based on information on 1000 randomly selected fields about the tenancy status of the cultivators of these fields and use of fertilizers, collected in an agro- eco survey. The following

Classification was noted.

	Owned	Rented	Row total(RT)
Using fertilizer	416 =a	184=b	600=a+b
Not using fertilizer	64=c	336=d	400=c+d
Column total(CT)	480=a+c	520=b+d	1000=a+b+c+d=N

Would you conclude that owner cultivators are more inclined towards the use of fertilizer at 5% level of significance?

Solution: Ho: Ownership of fields and the use of fertilizers are independent attributes.

Under this test,

$\chi 2 = \sum\sum (Oij-Eij)^2/Eij$

To calculate Eij values, we have to use above table as,

Expected value of a=E (a) =600x480/1000=288,

E(b)=312, E(c)=192 and E(d)=208

Now applying $\chi 2$ table

Oij	Eij	$(Oij-Eij)^2/Eij$
416	288	
64	192	
184	312	
336	208	
		=273.504

From above table, calculated $\chi 2$ =273.504

Tabulated $\chi 2$ at .05 with [(r-1) (c-1) =1] 1 d.f. =3.84

Result: Calculated $\chi 2$ is much more than tabulated $\chi 2$, Ho is rejected.

Hence, it can be concluded that the owner cultivators are more inclined towards the use of fertilizer.

Analysis of Co-variance (ANOCOVA)

8.1 Introduction

It is a well known fact that the necessary condition for the field trials (experiment) is the uniformity of the soil condition and culture (experimental material) so that some valid conclusions may be drawn about the various treatments under study. But in nature it is not possible to get an ideal experimental material as is thought of for the field experiments and therefore the treatment comparisons can not be made with sufficient precision.

In all the experiments discussed under "Design of Experiments", it has been possible to make treatment comparisons with the greatest possible precision, since the design of experiments deal with the most efficient methods of obtaining reliable data with a suitable design, and with the methods of analysis and interpretation of such data. In these designs greater experimental precision is obtained by the principle of replication, randomization and local control. It is the application of principle of local control by which it has been possible to eliminate the variation between rows and between columns from the residual error variations.

The use of the principle of confounding and the split plot technique are the other means of further increasing the precision of the estimates. But in certain experiments, the error cannot be reduced merely by above described methods.

There is yet another method by which the error affecting the treatment comparisons may be further diminished. This method is based on the fact that there are some extraneous sources of variation which also contribute to the experimental error but are not controlled. Measurements can be made for such characteristics (extraneous sources of variation) in each plot in addition to its yield data.

These additional observations are known as the ancillary or concomitant observations such as the plant number, age of the species, straw yield etc. these ancillary data may be utilized in the analysis for reducing the experimental error.

The precision of the treatment effect can thereby be increased. This form of error control is generally termed as "**Statistical control of error".**

The very logical procedure which reduces the experimental error by eliminating from it the effects of variation in the concomitant variate and thus increases the precision on the experiments of the treatment means on the basis of the regression of the main variate on the ancillary or concomitant variate is known as **"Analysis of Co-variation".**

The type of analysis, taking into account the ancillary data also along with the main data, is known as **analysis of co-variance**.

The analysis of co-variance refers to a technique in which the method of regression analysis is utilized to reduce the error of estimates of treatment effects. The method is commonly applied to the analysis of experimental data involving a dependent variable and one or more independent variable.

Just as in the analysis of variance we sort out the variance components attributable to different sources of variation like blocks, treatments, error etc. Similarly, in the analysis of co-variance we sort out the covariance effects attributable to the different sources.

In the words of Fisher, **"Analysis of co-variance combines the advantages and reconciles the requirements of the two very widely applicable procedures known as regression and analysis of variance".**

8.2 Analysis of co-variance in CR-Design

The method of analysis of co-variance may be applied for any experiment which has been properly designed. In this case, we use the method for CR design. Also here we take only one independent (ancillary) variable.

Mathematical Model

An appropriate mathematical model for this analysis for a one way classified data of k treatments each replicated r times, collected from a CR Design is

$$y_{ij} = \mu + \tau_i + bx_{ij} + e_{ij} \text{ .....(1), } i = 1, 2, ...k; j = 1, 2, ...r$$

where y_{ij} is the random observation having i^{th} treatment and j^{th} replication
μ is the fixed effect of general mean
τ_i is the fixed effect of i^{th} treatment
x_{ij} is the random observation on the ancillary variate corresponding to y_{ij}
b is the regression coefficient of y on x
e_{ij} is the error components which are assumed to be normally and independently distributed with zero mean and a constant variance σ^2 i.e $e_{ij} \sim N(0, \sigma^2)$.

Assumptions

For the valid use of the analysis of covariance the following assumptions has been made.

1. As postulated by the model, treatment, regression and error effects must be additive.
2. The error term e_{ij} must be normally and independently distributed with zero mean and a constant variance σ_e^2.
3. The values of the independent (ancillary) variable are independent of the treatments.
4. The relationship between the dependent and independent variable is linear over the range being investigated.
5. The value of regression coefficient of the linear relationship between the dependent and independent variables is assumed to be independent of the treatments.

8.3 Analysis

For the analysis of mathematical model given in the eq^n (1)

$$\text{Let, } \sum_j y_{ij} = y_{i.}, \sum_i \sum_j y_{ij} = \sum_i y_{i.} = y_{..} = G_y$$

$$\sum_j x_{ij} = x_{i.}, \sum_i \sum_j x_{ij} = \sum_i x_{i.} = x_{..} = G_x$$

By applying the technique of least squares, the estimates of the fixed effects along with the regression coefficient are as shown below from eq^n (1)

For the analysisof mathematical model given in the eq^n (1)

$$\left[\begin{array}{l} \mu = \dfrac{\sum_i \sum_j y_{ij} - b\sum_i \sum_j x_{ij}}{rk} = \dfrac{G_y - bG_x}{rk} \\ \tau_i = \dfrac{\sum_j y_{ij} - b\sum_j x_{ij}}{r} - \mu \\ b = \dfrac{\sum_i \sum_j x_{ij}\, y_{ij} - \dfrac{\sum_i x_{i.}\, y_{i.}}{r}}{\sum_i \sum_j x_{ij}^2 - \dfrac{\sum_i x_{i.}^2}{r}} \end{array}\right] \quad \text{.....(2)}$$

Substituting the estimated values of μ, τ_i, b in the eq^n (1) and solving, the error sum of squares adjusted for the variation of x is obtained as

$$S_e = \sum_i \sum_j e_{ij}^2 = \sum_i \sum_j y_{ij}^2 - \frac{\sum_i y_{i.}^2}{r} - b\left(\sum_i \sum_j x_{ij}\, y_{ij} - \frac{\sum_i x_{i.}\, y_{i.}}{r}\right) \quad \text{.....(3)}$$

The degrees of freedom for the adjusted error sum of squares is (rk-k) – 1. This has been reduced by one degree of freedom from that in the CR Design because the estimation of error variance here is subject to one more restriction imposed for estimating b.

Next, the adjusted treatment sum of squares can be obtained by subtracting the adjusted error sum of squares from adjusted total sum of squares.

The adjusted total sum of squares is given as

$$S_{t+e} = \sum_i \sum_j y_{ij}^2 - \frac{y_{..}^2}{rk} - \frac{\left(\sum_i \sum_j x_{ij}\, y_{ij} - \dfrac{x_{..}\, y_{..}}{rk}\right)^2}{\left(\sum\sum x_{ij}^2 - \dfrac{x_{..}^2}{rk}\right)} \quad \text{.....(4)}$$

$\therefore$ **The adjusted treatment sum of squares is given as**

$$S_t = S_{t+e} - S_e \quad \text{.....(5)}$$

We now give the basic calculations needed to set up the analysis of co-variance table for a CR Design. There will be a set of calculations for y, the variable of interest, a second set of calculations for the variable x, the covariate, and a third set of calculations for the cross products of x and y.

The total, treatment and error sums of squares for variable x, y and xy are given as follows respectively.

Total sum of squares for x, $S_x = \sum_i \sum_j x_{ij}^2 - \frac{x_{..}^2}{rk}$

Treatment sum of squares for x, $T_x = \sum_i \frac{x_{i.}^2}{r} - \frac{x_{..}^2}{rk}$

Error sum of squares for x, $E_x = S_x - T_x$

Total sum of squares for y, $S_y = \sum_i \sum_j y_{ij}^2 - \frac{y_{..}^2}{rk}$

Treatment sum of squares for y, $T_y = \sum_i \frac{y_{i.}^2}{r} - \frac{y_{..}^2}{rk}$

Error sum of squares for y, $E_x = S_y - T_y$

Total sum of squares for xy, $S_{xy} = \sum_i \sum_j x_{ij}\, y_{ij} - \frac{x_{..}\, y_{..}}{rk}$

Treatment sum of squares for xy, $T_{xy} = \sum_i \frac{x_{i.}\, y_{i.}}{r} - \frac{x_{..}\, y_{..}}{rk}$

Error sum of squares for xy, $E_{xy} = S_{xy} - T_{xy}$

To summarize the calculations of the sums of squares and to illustrate the covariance adjustments of covariance table denoted by part A is presented.

ANOCOVA for CR Design – Part A

Source of variation	d.f.	SS and SP			Adjustment due to regression	Adjusted sum of squares
		x	y	xy		
Treatments	k-1	T_x	T_y	T_{xy}		$S_T = S_{T+e} - S_e$
Error	(rk – k)	E_x	E_y	E_{xy}	$\frac{E_{xy}^2}{E_x}$	$S_e = E_y - \frac{E_{xy}^2}{E_x}$
Total	(rk – 1)	S_x	S_y	S_{xy}	$\frac{S_{xy}^2}{S_x}$	$S_{T+e} = S_y - \frac{S_{xy}^2}{S_x}$

Before we complete part-B of the ANOCOVA table, where tests of adjusted treatment means would be made, we need to discuss some preliminary tests useful in checking the validity of ANOCOVA assumptions.

Preliminary analysis might include tests of

i. Equal residual variances for all treatment groups.

ii. Equal slopes for all treatment groups.

iii. Common slope of zero.

iv. Equality of unadjusted x-treatment means.

v. Equality of unadjusted y-treatment means.

These tests are not made in the order listed; in fact, some are required to be made in a specific sequence, and some may not be made in a given experiment.

Recall that ANOCOVA can be described as fitting a linear regression to the data of each treatment group. The residual variation for each treatment group would be the residual variation about that groups regression lines. If there is any concern about equality of variances, this test should be made before making the tests (ii) and (iii). The test procedure for (ii) and (iii) requires equality of variances to be valid.

Next, the (ii) test about the equal slopes for all treatment groups i.e. no interaction between treatments and the covariate x can be tested under the null hypothesis

Ho: No interaction of x and treatment

Or

Ho: Equal slopes for the treatment groups.

($\beta_1 = \beta_2 = \ldots\ldots\ldots\ldots = \beta_n$)

Then by obtaining the F statistics

$$\mathbf{F = \frac{T_{xy} \big/ (k-1)}{E_{xy} \big/ (rk-k)}}\ \textbf{we could test the null hypothesis}$$

Again, if there is any concern about unequal slopes. This preliminary test should precede the test of zero common slopes.

The null hypothesis for zero common slopes is

Ho: $\beta = 0$

Then the test statistic is an F variable formed from entries on the residual line of ANOCOVA table part-A.

$$\mathbf{F = \frac{E_{xy}^2 \big/ E_x}{S_e \big/ (n-k-1)} = \frac{Residual\ adjustment}{Adjusted\ residual\ ms}}$$

Which has degrees of freedom of 1 and (n – k – 1). If we fail to reject the null hypothesis of zero common slopes, we would have no evidence of a slope.

When we arrive at such a conclusion, we would not make the covariance adjustments because adjustments to a line with zero slopes simply would be random adjustments. A test of unadjusted y treatment means would be made.

Next, a test about the equality of unadjusted x-treatment means would be considered if there is some question of treatments affected the covariate.

Simply perform a "Regular ANOVA test" on treatments using the x-data, the x-column of ANOCOVA table part-A.

The null hypothesis for this test is

Ho: All treatment means are equal and the test statistic is

$$\mathbf{F = \frac{T_x / (k-1)}{E_x / (n-k)}} \quad \textbf{which has degrees of freedom of } (k-1) \textbf{ and } (n-k).$$

If we reject Ho then we conclude that x-treatment means differ, implying that treatments have caused the x-values to change, we would not make the covariance adjustments.

When treatments are affecting the covariate x, the effects of treatments and x are interrelated; the covariance adjustment to remove the effect of x would be tantamount to removing some of the treatment effect.

A test of equality of unadjusted y-means would be made when there is evidence that treatments are affecting the covariate.

Note that the hypothesis of all x treatment means are equal would not be tested if treatments cannot possibly affect the value of the covariate, when for example; the covariate is a pre-experiment score.

The test of unadjusted y-treatment means would be made in the two situations indicated in the preceding paragraphs. Additionally this test might be made to gain insight into the treatment effects before and after adjustment for the covariate. This will be a regular ANOVA test on the y-data using the y column of ANOCOVA table part-A.

Next, we discuss about the test of adjusted y-treatment means which is presented in ANOCOVA table part-B. This table can easily be setup from part-A.

ANOCOVA for CR Design – Part - B

Source of variation	d.f.	Adjusted SS	MSS	F -ratio
Treatments	k-1	S_t		$F = \frac{s_t}{s_e}$
Error	(rk – k-1)	S_e	$MSS = s_t = \frac{S_t}{k-1}$	

Total	**(rk – 2)**	S_{t+e}	$MSE = s_e = \frac{S_e}{(n-k-1)}$	

Actually, the null hypothesis of interesting an ANOCOVA experiment is

Ho: Adjusted y-treatment means are equal.

Or

Ho: All $\tau_i = 0$

And the test statistic used is

$$F = \frac{s_t}{s_e}$$

Which under Ho and the assumptions stated earlier, has the F-distribution with (k-1) and (n-k-1) degrees of freedom.

When we reject null hypothesis, there is interest in the adjusted y treatment mean. Estimates of these means are obtained as follows,

$$\text{Adj } \bar{y}_{i.} = \bar{y}_{i.} - \hat{\beta}(\bar{x}_{i.} - \bar{x}_{..}) \text{ Where, } \hat{\beta} = \frac{E_{xy}}{E_x}$$

Also the adjusted treatment effects τ_i , whose estimates are

$$\hat{\tau}_i = \text{Adj } \bar{y}_{i.} - \bar{y}_{..}$$

Finally the estimates of differences of adjusted means are given by

$$\begin{aligned}(\text{Adj } \bar{y}_{i.} - \text{Adj } \bar{y}_{h.}) &= (\bar{y}_{i.} - \bar{y}_{h.}) - \hat{\beta}(\bar{x}_{i.} - \bar{x}_{h.}) \\ &= \hat{\tau}_i - \hat{\tau}_h\end{aligned}$$

And the estimated variance of the difference of two adjusted means is given by

$$\hat{V}(\text{Adj } \bar{y}_{i.} - \text{Adj } \bar{y}_{h.}) = MSE\left(\frac{1}{r_i} + \frac{1}{r_h} + \frac{(\bar{x}_{i.} - \bar{x}_{h.})^2}{E_x}\right)$$

Example:

A forest biometrician wished to investigate the effects of three types of fertilizer on the growth of seedlings of certain species in a nursery. Fifteen seedlings of the same age were obtained. The forester noticed that the initial heights were not the same but mo attempt was made to set up blocks according to initial height. Five seedlings were randomly assigned to each fertilizer treatment and after six months, the growth (height minus initial height) was determined for each tree giving in mm. The initial heights and growths are given below in the format (x,y)

Fertilizer		
A	**B**	**C**
(20,39)	**(13,28)**	**(25,19)**
(25,32)	**(12,40)**	**(22,21)**
(17,41)	**(15,33)**	**(27,15)**
(30,27)	**(17,34)**	**(21,23)**
(14,41)	**(14,39)**	**(25,19)**

Carry out the analysis and interpret the results.

Solution:

The first step is to carry out the analysis of variance of the growth data (y) in the usual manner to compare the growth of seedlings with out any adjustment of initial height. We know that the initial height and growth are given in the format (x,y) and hence the total and average are also given in the same format in the table given below.

	Fertilizer			**Grand** (G_x, G_y)
	A	**B**	**C**	
	(20, 39)	**(13, 38)**	**(25, 19)**	
	(25, 32)	**(12, 40)**	**(22, 21)**	
	(17, 41)	**(15, 33)**	**(27, 15)**	
	(30, 27)	**(17, 34)**	**(21, 23)**	
	(14, 41)	**(14, 39)**	**(25, 19)**	
Total	**(106, 180)**	**(71, 184)**	**(120, 97)**	**(297, 461)**
Average	**(21.2, 36)**	**(14.2, 36.8)**	**(24, 19.4)**	**(19.8, 30.73)**
$(\bar{x}_{..}, \bar{y}_{..})$				

Calculation for x-series

$$\text{Correction factor (C.F.)}_x = \frac{G_x^2}{rk} = \frac{(297)^2}{5\times 3} = 5880.6$$

$$\text{Treatment SS (SST)}_x = \frac{\sum_i x_{i.}^2}{r} - \text{C.F.} = \frac{(106)^2 + (71)^2 + (120)^2}{5} - 5880.6 = 6135.4 - 5880.6 = 254.8$$

$$\text{Total SS (TSS)}_x = \sum_{ij} x_{ij}^2 - \text{C.F.} = (20)^2 + (25)^2 ..... + (21)^2 + (25)^2 - 5880.6 = 456.4$$

$$\text{Error SS (ESS)}_x = (\text{TSS})_x - (\text{SST})_x = 456.4 - 254.8 = 201.6$$

Calculation for y-series

$$\text{Correction factor (C.F.)}_y = \frac{G_y^2}{rk} = \frac{(461)^2}{5\times 3} = 14168.06$$

$$\text{Treatment SS (SST)}_y = \frac{\sum_i y_{i.}^2}{r} - \text{C.F.} = \frac{(180)^2 + (184)^2 + (97)^2}{5} - 14168.06 = 964.93$$

$$\text{Total SS (TSS)}_y = \sum_{ij} y_{ij}^2 - \text{C.F.} = (39)^2 + (32)^2 ..... + (23)^2 + (19)^2 - 14168.06 = 1194.93$$

$$\text{Error SS (ESS)}_y = (\text{TSS})_y - (\text{SST})_y = 1194.93 - 964.93 = 230.0$$

Calculation for Cross product (xy-series)

$$\text{Correction factor for sum of products (C.F.)}_{xy} = \frac{G_x G_y}{rk} = \frac{297\times 461}{5\times 3} = 9127.8$$

$$\text{Sum of products for Treatment SPT}_{(xy)} = \frac{\sum_i x_{.i.}\, y_{i.}}{r} - 9127.8$$

$$= \frac{1}{5}[(106)(180) + (71)(184) + (120)(97)] - 9127.8$$

$$= -371.0$$

$$\text{Total Sum of products TSP}_{(xy)} = \sum_{ij} x_{ij}\, y_{ij} - \text{C.F.}$$

$$= (20)(39) + (25)(32) ..... + (21)(23) + (25)(19) - 9127.8$$

$$= -573.8$$

$$\text{Error SS (ESS)}_{xy} = (\text{TSP})_{xy} - (\text{SPT})_{xy} = -573.8 - (-371.0) = -202.8$$

ANOCOVA on Growth Part-A

Source of variation	d.f.	SS and SP			Adjustment due to regression	Adjusted sum of squares
		x	y	xy		
Treatments	k-1= 3-1= 2	T_x = 254.8	T_y = 964.9	T_{xy} = -371.0		$S_T = S_{T+e} - S_e = 473.5 - 26.0 = 447.5$
Error	(rk – k) = (15– 3) = 12	E_x = 201.6	E_y = 230.0	E_{xy} = -202.8	$\frac{E_{xy}^2}{E_x} = \frac{(-202.8)^2}{201.6} = 204.0$	$S_e = E_y - \frac{E_{xy}^2}{E_x} = 230.0 - 204.0 = 26.0$
Total	(rk – 1) = (15 – 1) = 14	S_x = 456.4	S_y = 1194.9	S_{xy} = -573.8	$\frac{S_{xy}^2}{S_x} = \frac{(-573.8)^2}{456.4} = 721.4$	$S_{T+e} = S_y - \frac{S_{xy}^2}{S_x} = 1194.9 - 721.4 = 473.5$

ANOCOVA for CR Design – Part – B

Source of variation	d.f.	Adjusted SS	MSS	F –ratio for adjusted treatment mean
Treatments	k-1=3-1=2	$S_t = 447.5$	$MSS = s_t = \frac{S_t}{k-1} = \frac{447.5}{2} = 223.75$	$F = \frac{s_t}{s_e} = \frac{223.75}{2.36} = 94.80$
Error	(rk – k-1) =15-3-1=11	$S_e = 26$	$MSE = s_e = \frac{S_e}{(n-k-1)} = \frac{26}{11} = 2.36$	
Total	(rk – 2) =15-2=13	$S_{t+e} = 473.5$		

For the test of equality of adjusted mean growths for the three fertilizers, let the null hypothesis is

H_o: Adjusted mean growths are equal for the three fertilizers.

$$Ho: \bar{y}_{1.} = \bar{y}_{2.} = \bar{y}_{3.}$$

H_1: Adjusted mean growths differ for the three fertilizers.

At the 5% level of significance, F tabulated is 3.98 for the (2, 11) degrees of freedom, and At the 1% level of significance, F tabulated is 9.65 for the (2, 11) degrees of freedom.

Hence F calculated value is greater for both 5% and 1% level of significance, we reject null hypothesis. That means there is sufficient evidence that adjusted growth means for the three fertilizers are unequal.

When we reject null hypothesis, there is interest in the adjusted y treatment mean. Estimates of these means are obtained as follows,

$$\textbf{Adj } \bar{y}_{i.} = \bar{y}_{i.} - \hat{\beta}(\bar{x}_{i.} - \bar{x}_{..}) \textbf{ Where, } \hat{\beta} = \frac{E_{xy}}{E_x}$$

$$\therefore \textbf{Adj } \bar{y}_{1.} = \bar{y}_{1.} - \hat{\beta}(\bar{x}_{1.} - \bar{x}_{..}) = 36 - \frac{-202.8}{201.6}(21.2 - 19.8) = 37.40$$

$$\therefore \textbf{Adj } \bar{y}_{2.} = \bar{y}_{2.} - \hat{\beta}(\bar{x}_{2.} - \bar{x}_{..}) = 36.8 - \frac{-202.8}{201.6}(14.2 - 19.8) = 31.17$$

$$\therefore \textbf{Adj } \bar{y}_{3.} = \bar{y}_{3.} - \hat{\beta}(\bar{x}_{3.} - \bar{x}_{..}) = 19.4 - \frac{-202.8}{201.6}(24.0 - 19.8) = 23.62$$

Also the adjusted treatment effects τ_i , whose estimates are

$$\hat{\tau}_i = \textbf{Adj } \bar{y}_{i.} - \bar{y}_{..}$$

$$\therefore \hat{\tau}_1 = \textbf{Adj } \bar{y}_{1.} - \bar{y}_{..} = 37.40 - 30.73 = 6.67$$

$$\therefore \hat{\tau}_2 = \textbf{Adj } \bar{y}_{2.} - \bar{y}_{..} = 31.17 - 30.73 = 0.44$$

$$\therefore \hat{\tau}_3 = \textbf{Adj } \bar{y}_{3.} - \bar{y}_{..} = 23.62 - 30.73 = -7.11$$

Finally the estimates of differences of adjusted means are given by

$$(\textbf{Adj } \bar{y}_{i.} - \textbf{Adj } \bar{y}_{h.}) = (\bar{y}_{i.} - \bar{y}_{h.}) - \hat{\beta}(\bar{x}_{i.} - \bar{x}_{h.}) = \hat{\tau}_i - \hat{\tau}_h$$

$$\therefore (\textbf{Adj } \bar{y}_{1.} - \textbf{Adj } \bar{y}_{2.}) = 6.67 - 0.44 = 6.23$$

$$\therefore (\textbf{Adj } \bar{y}_{1.} - \textbf{Adj } \bar{y}_{3.}) = 6.67 - (-7.11) = 13.78$$

$$\therefore (\textbf{Adj } \bar{y}_{2.} - \textbf{Adj } \bar{y}_{3.}) = 0.44 - (-7.11) = 7.55$$

And the estimated variance of the difference of two adjusted means is given by

$$\hat{V}(\textbf{Adj } \bar{y}_{i.} - \textbf{Adj } \bar{y}_{h.}) = MSE\left(\frac{1}{r_i} + \frac{1}{r_h} + \frac{(\bar{x}_{i.} - \bar{x}_{h.})^2}{E_x}\right)$$

And the standard error of the difference of two adjusted means is given by

$$S.E.(\textbf{Adj } \bar{y}_{i.} - \textbf{Adj } \bar{y}_{h.}) = \sqrt{MSE\left(\frac{1}{r_i} + \frac{1}{r_h} + \frac{(\bar{x}_{i.} - \bar{x}_{h.})^2}{E_x}\right)}$$

$$\therefore S.E.(\textbf{Adj } \bar{y}_{1.} - \textbf{Adj } \bar{y}_{2.}) = \sqrt{2.36\left(\frac{1}{5} + \frac{1}{5} + \frac{(21.2 - 14.2)^2}{201.6}\right)} = 1.23$$

$$\therefore S.E.(\textbf{Adj } \bar{y}_{1.} - \textbf{Adj } \bar{y}_{3.}) = \sqrt{2.36\left(\frac{1}{5} + \frac{1}{5} + \frac{(21.2 - 24)^2}{201.6}\right)} = 1.01$$

$$\therefore \text{S.E.}\left(\text{Adj } \bar{y}_{2.} - \text{Adj } \bar{y}_{3.}\right) = \sqrt{2.36\left(\frac{1}{5}+\frac{1}{5}+\frac{(14.2-24)^2}{201.6}\right)} = 1.43$$

Exercise

1. A corn breeder tested four newly developed varieties in a CRD experiment. In addition to recording the yield per plot, Y he also recorded the number of plants harvested per plot, X. The data were as follows in the format (x,y):

Variety				
I	II	III	IV	Totals
(40,320)	(37,282)	(32,290)	(41,273)	(150,1165)
(32,300)	(34,278)	(32,283)	(42,271)	(140,1132)
(38,325)	(41,290)	(39,310)	(40,283)	(158,1208)
(42,341)	(30,270)	(33,265)	(36,266)	(141,1142)
(35,316)	(45,293)	(37,296)	(37,280)	(154,1185)
(187,1602)	(187,1413)	(173,1444)	(196,1373)	(743,5832)

   Carry out the analysis and interpret the results.

2. An animal scientist tested four Feeding stuff in a CRD experiment. The gain in weight of rats resulted are presented in the table below. In addition to recording the gain in weight, Y he also recorded the quantity of feed, X. The data were as follows in the format (x,y):

Feeding stuff			
A	B	C	D
(96,98)	(109,64)	(179,71)	(127,72)
(108,102)	(125,86)	(132,84)	(100,54)
(94,102)	(85,51)	(163,71)	(151,109)
(128,108)	(82,72)	(143,62)	(116,93)

   Carry out the analysis and interpret the results.

3.

Example : (ANOCOVA in RBD)

A varietal trial on maize with yield data (gm/plot) of three varieties and the crop stand per plot (in bracket) are shown below:

Varieties	Blocks					Total	Average
	I	II	III	IV	V		
1	6 (10)	3 (6)	4 (7)	7 (13)	4 (17)	24 (53)	4.8 (8.6)
2	10 (12)	6 (18)	8 (14)	4 (10)	5 (15)	33 (69)	6.6 (13.8)
3	9 (13)	8 (12)	6 (16)	11 (8)	5 (10)	39 (59)	7.8 (11.8)
Total yield	25	17	18	22	14	96	6.4
Total stand	(35)	(36)	(37)	(31)	(32)	(171)	(11.0)

Let the variable y denote the yield and x denote the crop stand per plot

The first step is to carry out the analysis of variance of the yield data (y) in the usual manner to compare the yield of the varieties with out any adjustment of various plant numbers.

$$\textbf{Correction factor (C.F.)} = \frac{G_y^2}{rk} = \frac{(96)^2}{5\times 3} = 614.4$$

$$\textbf{Block SS (BSS)} = \frac{\sum_j y_{.j}^2}{k} - \text{C.F.} = \frac{(25)^2+(17)^2+(18)^2+(22)^2+(14)^2}{3} - 614.4 = 24.93$$

$$\textbf{Variety SS (VSS)} = \frac{\sum_i y_{i.}^2}{r} - \text{C.F.} = \frac{(24)^2+(33)^2+(39)^2}{5} - 614.4 = 22.8$$

$$\textbf{Total SS (TSS)} = \sum_{ij} y_{ij}^2 - \text{C.F.} = (6)^2+(3)^2 \ldots\ldots\ldots\ldots + (11)^2+(5)^2 - 614.4 = 79.6$$

$$\textbf{Error SS (ESS)} = \text{TSS} - \text{BSS} - \text{VSS} = 79.6 - 24.93 - 22.8 = 31.87$$

ANOVA for plot yield (y)

Source of variation	d.f.	SS	MSS	F-ratio	F-tabulated
Block	5-1=4	24.93	6.23	$F_c = 2.86$	$F_t = 4.40$
Variety	3-1=2	22.8	11.4		
Error	8	31.87	3.98		
Total	15-1=14	79.6			

Since, our null hypothesis Ho is all unadjusted y-treatment means are equal and from the analysis we accept the null hypothesis and we conclude that the difference in yield is not significant. This may be due to wide variation in plant numbers in different plots as may be seen from the tabulation of data. It is this influence of variate (plant number) called the

concomitant or ancillary on the main variate y (yield); to contribute partly to the experimental error.

Secondly, the analysis of data on plant stand (x) will be carried out to test the null hypothesis that all unadjusted x-treatment means are equal.

For this,

$$\textbf{Correction factor (C.F.)} = \frac{G_x^2}{rk} = \frac{(171)^2}{5\times 3} = 1949.4$$

$$\textbf{Block SS (BSS)} = \frac{\sum_j x_{.j}^2}{k} - \text{C.F.} = \frac{(35)^2 + \ldots\ldots + (32)^2}{3} - 1949.4 = 59.6$$

$$\textbf{Variety SS (VSS)} = \frac{\sum_i x_{i.}^2}{r} - \text{C.F.} = \frac{(43)^2 + (69)^2 + \ldots\ldots (59)^2}{5} - 1949.4 = 68.8$$

$$\textbf{Total SS (TSS)} = \sum_{ij} x_{ij}^2 - \text{C.F.} = (10)^2 + (6)^2 \ldots\ldots\ldots\ldots + (8)^2 + (10)^2 - 1949.4 = 175.6$$

$$\textbf{Error SS (ESS)} = \text{TSS} - \text{BSS} - \text{VSS} = 175.6 - 59.6 - 68.8 = 47.2$$

ANOVA for plot stand (x)

Source of variation	d.f.	SS	MSS	F-ratio	F-tabulated
Block	5-1=4	59.6	14.9	$F_c = 5.83$	$F_t = 4.40$
Variety	3-1=2	68.8	34.4		
Error	8	47.2	5.9		
Total	15-1=14	175.6			

From the analysis, the null hypothesis is rejected and we conclude that variation in plant number in different plot is significant.

Now, one way is to carry out the statistical analysis not on the plot yields, but on the values of average yield per plant obtained in each plot, by dividing the plot yield by plant number. Such an adjustment might exaggerate the yield rate for plot with fewer plants. Hence the systematic procedure for adjustment of the data is known as the analysis of covariance.

Computation for the analysis of Co-variance

A third table i.e. the analysis of covariance of the plot yield (y) and the plant number (x) is constructed by using at every state products of x and y observations.

$$\textbf{Correction factor for sum of products (C.F.)} = \frac{G_x G_y}{rk} = \frac{171\times 96}{5\times 3} = 1094.4$$

$$\textbf{Sum of products for Block, } BSP_{(xy)} = \frac{\sum_i x_{i.}\, y_{i.}}{k} - \text{C.F.} = 1133.3 - 1094.4 = 38.9$$

$$\textbf{Sum of products for Variety } VSP_{(xy)} = \frac{\sum_j x_{.j}\, y_{.j}}{r} - \text{C.F.} = 1222.0 - 1094.4 = 27.6$$

$$\textbf{Total Sum of products } TSP_{(xy)} = \sum_{ij} x_{ij}\, y_{ij} - \text{C.F.} = (10)^2 + (6)^2 ..... + (8)^2 + (10)^2 - 1949.4 = 175.6$$

$$\textbf{Error SS(ESS)} = TSS - BSS - VSS = 175.6 - 59.6 - 68.8 = 47.2$$

3. χ2 test as a test of homogeneity

The χ2 test as a test of homogeneity is an extension of the χ2 test of independence. Tests of homogeneity are designed to determine whether two or more independent random samples are drawn from the same population or from different populations. Instead of one sample as we use with independent problem we shall now have 2 or more samples.
For e.g., we may be interested in finding out whether or not university students of various levels i.e. I.SC., B.SC,M.SC feel the same in regard to the amount of work required by their Professors i.e. too much work, right amount of work , too little work required by their professors.
We shall take the hypothesis that there exists no difference in opinion among the three classes of pupil on the issue.

Examples:

1. Hypothesis Ho: there is no significance difference between male and female farmers in their discussion with males.

Discussion patterns of male respondents:

Discussion by	Discussion with males		
	Yes	No	Row total
Male	112	23	135
Female	143	24	167
Column total	255	47	302

Chi-square value=0.4078 NS, df=1

χ2- test for discussion patterns showed that there is no significance difference between male and female farmers in their discussion with males.

2. Hypothesis Ho: there is no significance difference between male and female farmers in their discussion with females.

Discussion patterns of female respondents:

Discussion by	Discussion with females		
	Yes	No	Row total
Male	79	88	167
Female	82	53	135
Column total	161	141	302

Chi-square value=5.3279 significant, df=1, $p<0.05$

χ2- test for discussion patterns showed that significance difference between male and female farmers in their discussion with females.

9.1 Confounding

In factorial experiments when the number of factors and the levels at which they are employed increases, the total number of treatment combinations increases rather rapidly and consequently the blocks size has to be enlarged. In many experiments even a single replication of each treatment combination may require too many experimental units. For example, for a 2^5 factorial experiment, a complete factorial would require 32 units and it is not advisable to adopt a randomized block design for it, because blocks of 32 plots are too big to ensure homogeneity with in them. The heterogeneity introduced as a consequence of the size of the experiment results in extraneous variation which will add to experimental error. As a consequence of increase in the block size or handling such a huge experiment, the purpose of local control is defeated due to the following reasons.

1. It is sometimes impracticable to get one complete replicate units which are relatively homogeneous.
2. The greater heterogeneity is introduced in the experimental error and reduces the discriminating power of the tests of significance (t, F) thus vitiating the conclusion to be drawn.

Hence, the precision of a factorial experiment is adversely affected if the treatment combinations are large in number. In order to maintain homogeneity with in the blocks, the experimenter must either cut down the number of factors (which of course will mean loss of information) or use an incomplete factorial which investigates the main effect of the factors and their more important interactions under uniform conditions by suitably subdividing the experimental material into smaller homogeneous blocks. The heterogeneity of blocks is allowed to affect only interactions which are likely to be unimportant.

In a factorial experiment, the estimates of main effects and interactions of the lower order can be obtained more precisely, if each replication is further subdivided into equal smaller blocks. By so doing incomplete blocks of small size are obtained, making the experimental material more homogeneous within the blocks, and thus increasing the precision. But at the same time the main effect or the interaction which is obtained by comparing the treatment combinations of one block with those of other is entangled or confounded with the block effect.

In this way we see that by dividing the replicates into more homogeneous smaller blocks, certain effects or interactions can be confounded with the block effects and the precision on the remaining ones can be increased. "Thus confounding can be defined as a device by which the precision on the main effects and certain interactions is increased by sacrificing the precision on certain high order interactions."

"The process by which unimportant (high order interactions) comparisons are deliberately mixed up or entangled with the block comparisons, for the purpose of assessing more important (main or low order interactions) comparisons with greater precision is called confounding."

"Confounding may also be defined as the technique of reducing the size of a replication over a number of blocks at the cost of losing some information on some effect which is not of much practical importance."

The device of confounding consists in subdividing the replicate into two or more equal subgroups (blocks) and the various treatment combinations into two or more groups of equal size following certain rules by which we sacrifice some information on certain high order interactions and allocating the treatment combinations of any group to any block at random.

In a confounded factorial experiment, preferably only high order interactions, that is, interactions with three or more factors are confounded, because their loss is immaterial. If, an experimenter is more interested in main effects and two order interactions, these should not be confounded as far as possible. The designs for such confounded factorials can be called incomplete randomized block designs. The treatments groups are first allotted at random to the different blocks, the treatment allotted to a block are then distributed at random to its different units.

Confounding in a 2^3 Factorial Experiment

Let us consider a 2^3 factorial experiment with factors **A, B** and **C** each at two levels, so that there are 8 treatment combinations which require 8 units of homogeneous material each to form a block. If we decide to use blocks of 4 units (plots) each then a full replication will require only 2 block. In this case 8 treatment combinations are divided into 2 groups of 4 treatments each in a special way so as to confound any one of the less important interactions with blocks and these groups are allocated at random in the 2 blocks.

For example let us consider confounding the highest order interaction **ABC**, which is given by

$$\mathbf{ABC} = \frac{1}{4}\left[(abc)\ (bc)\ (ac) + (c)\ (ab) + (b) + (a) - (1)\right]$$

$$= \frac{1}{4}\left[(abc) + (a) + (b) + (c)\ (ab)\ (ac)\ (bc)\ (1)\right]$$

If we decide to confound the highest order interaction **ABC** with blocks all the treatment combinations with +ve sign are allocated at random in one block and those with –ve sign in the other block. Thus the following arrangement gives **ABC** confounded with blocks and hence we lose information on **ABC**.

Replication	
Block 1	**Block 2**
abc	**ab**
a	**ac**
b	**bc**
c	**1**

If we use complete confounding, i.e. confounding the same interaction effect in all replication, for experiment with replication 3, the plan would be as follows.

Replication 1		Replication 2		Replication 3	
Block 1	**Block 2**	**Block 3**	**Block 4**	**Block 5**	**Block 6**
abc	**ab**	**a**	**ab**	**b**	**bc**
a	**ac**	**c**	**1**	**a**	**ac**
b	**bc**	**b**	**bc**	**abc**	**1**
c	**1**	**abc**	**ac**	**c**	**ab**

There are two important properties of this plan. The total of blocks 1, 3 and 5 minus the total of blocks 2, 4 and 6 is the **ABC** interaction total. Thus the **ABC** interaction is one of the components of the comparisons amongst block. It is said to be completely confounded with blocks. The other six factorial effects **A, B, C, AB, AC** and **BC** will be found to be orthogonal with the block totals, so these are not influenced by differences among blocks and can thus be estimated and tested as usual with out any difficulty.

Advantages and Disadvantages

The only and the greatest advantage of confounding scheme lies in the fact, that it reduces the experimental error considerably by stratifying the experimental material into homogeneous subsets or subgroups. The removal of the variation among incomplete blocks (freed from treatments) within replication often results in smaller error mean square as compared with a randomized complete block design, thus making the comparisons among some treatments more precise.

The following are the disadvantages of confounding.

1. The confounded contrasts are replicated fewer times than are the other contrasts and as such there is loss of information on them and they can be estimated with a lower degree of precisions as the number of replications for them is reduced. In the confounding scheme, the increased precision is obtained at the cost of sacrifice of information (partial or complete) on certain relatively non-important interactions. It may be pointed out here that an indiscriminate use of confounding may result in complete or partial loss of information on the contrasts or comparisons of greatest importance. As such the experimenter should confound only those treatment combinations or contrasts which are of relatively less or no importance at all.
2. The algebraic calculations are usually more difficult and the statistical analysis is complex. Especially when some of the units are missing.

A number of problems arise if the treatments interact

9.2 Split Plot Design

General confounding scheme for factorial experiments were discussed in previous unit. There we indicated that the researcher should use a confounding system which will allow the estimation of the most important treatment components. Generally, high order interaction components would be confounded with incomplete blocks unless knowledge or information suggest otherwise.

We now consider a broad class of factorial experiments where at least one main effect is confounded with incomplete blocks. Split plot is a term applied to such experiments. There are many variations of this basic design; we will discuss the simple split plot design here.

A split plot design is a special form of factorial design, in which there are at least two different degrees of replication will be used. One set o f factors is allocated to the main plots of the experiment, so that the number of replications of those factors is equal to number of blocks, or rows. Each of these main plots is then split for the remaining set 9 or sets) of factors, so that the number of replications of these factors is equal to the number of main plots.

The main plots may be arranged in randomized blocks, in Latin square, or in any other suitable design. The allocation of the sub plots must be made random with in each main plot.

In some of the factorial experiments it may be conceived that some of the factors like ploughing, irrigation, sowing (planting or spreading or scattering) dates etc. may produce much larger differences than others and can not be applied conveniently to the smaller units and therefore require larger units for the purpose where as, there are some which can be very conveniently applied to the smaller units like manures, varieties etc. To conduct the experiment related to above described condition each block will be divided into a number of plots known as main plots equal to the number of levels of the first factor and then sub-dividing each main plot into a number of sub-plots equal to the number of levels of the second factor. The levels of the first factor are randomized in the main plots of each block and the levels of the second factor are randomized in the sub-plots of each main plot.

If there is a third factor also which is to be studied with still more precision as compared with the second factor, the sub-plots will again be further split into smaller ultimate plots to be assigned to the levels of the third factor.

In split plot experiments, the total variability may be split up into different components Viz. the variability due to main factor, sub-factor, main x sub-interaction and the error variability.