1

FOUNDATIONS OF EXPERIMENTAL DESIGN

INTRODUCTION

Empirical research provides knowledge to the researchers through direct or indirect observations or experiences. Empirical research may either involve correlational or experimental approach. In correlational research one looks to establish relationship between two variables. In such studies a premise is made that two variables may be related in some way and then values of both the variables are obtained under different conditions to test a hypothesis if indeed there is a relationship between the two. The obtained correlation is tested for its significance. The drawback of the correlational study is that it does not establish the cause and effect relationship even if the correlation is found to be statistically significant. For instance, if the observed correlation between the caffeine intake and concentration of mind is significant and positive, it cannot be said that caffeine causes concentration. The increase in the concentration due to the increase in the caffeine intake may be due to age, motivation, gender, other lifestyle parameters.

On the other hand, experimental research provides cause and effect relationship because in such experiment a treatment is deliberately administered by a researcher on a group of individuals or objects to see its impact under a controlled environment. In other words, if changes are made in the variable A that leads to changes in variable B, one can conclude that A causes B. For example, to see the impact of exercise on muscular strength a researcher may administer different intensity of exercise to different groups of individuals to see its effect. If a particular intensity of exercise improves

Repeated Measures Design for Empirical Researchers, First Edition. J. P. Verma.

muscular strength more than others, one may conclude that exercise intensity causes muscular strength. On the other hand, if there is no difference in the average muscular strength among different exercise groups, it may be inferred that the exercise intensity has nothing to do with muscular strength.

Authenticity in an experimental research is ensured only when an appropriate experimental design is used. Experimental design is a blueprint of the procedures which enables a researcher to test his hypothesis under a controlled environment. It describes the procedure of allocating treatments to the individuals in a sample. There are many ways in which an experimental design can be classified. One such classification is based on the method of allocating treatments to the subjects. On the basis of this criterion, experimental design can be classified into three categories; independent measures design, repeated measures design, and mixed design. In independent measures design each subject gets one and only one treatment, whereas in repeated measures design each subject is tested under all treatments. In mixed design each subject receives one and only treatment of first factor, but gets tested in all the treatments of second factor. This book specifically deals with some of the important repeated measures designs and mixed designs. To understand these designs and its applications, it is important to understand different aspects of experimental research such as principles of experimental design, types of statistical designs, terminologies used, and other considerations in planning an experimental research.

WHAT IS EXPERIMENTAL RESEARCH?

An *experimental research* is a process of studying the effect of manipulating independent variable on some dependent variable(s) observed on subjects in a controlled environment. For instance, in studying the effect of progressive relaxation on concentration, the progressive relaxation is an independent variable whereas the concentration is a dependent one. While conducting an experimental research, a researcher always tries to maintain control in an experiment so that valid conclusion can be drawn on the basis of findings. In experimental research the experimenter is allowed to manipulate independent variable to see its impact on the dependent variable. For instance, in the above example the experimenter can decide the duration or the intensity of the progressive relaxation program. Since the experimenter manipulates an independent variable to see its impact on dependent variable, cause and effect relationship can be explained on the basis of findings.

On the other hand in *observational study*, a researcher collects and analyzes data without manipulating independent variable. Here also the relationship is investigated between independent and dependent variable observed on the subjects. Since researcher is not allowed to manipulate an independent variable, causal interpretations cannot be efficiently made. If relationship is investigated between height and vertical jump performance of sprinters, the observed correlation may not be the strong evidence for causal relationship between them because the independent variable, height, has not been manipulated to see its impact on the vertical jump performance. This is because the experimenter cannot observe the control on the study. The subjects might have different weight, skill, motivation, and fitness level which do not allow interpreting the strong cause and effect relationship between height and the vertical jump performance. The observational study is also known as *correlational study* or *status study*.

Since validity of findings in an experiment depends upon the control observed during the experimentation, it is important to design the experiment in such a way so as to minimize the error involved in it. Using appropriate design in an experiment ensures proper allocation of treatments to the subjects so that experimental error is minimized. This ensures internal validity in the experiment. Design of experiment along with its principles has been discussed in detail in the following section.

DESIGN OF EXPERIMENT AND ITS PRINCIPLES

Design of experiment can be defined as a roadmap for organizing an experimental study for testing a research hypothesis in an efficient manner. Design of experiment facilitates an experimenter to observe control in an experiment, thereby reducing the experimental error and ensuring internal validity in findings. More specifically it provides a plan according to which treatments are allocated to the subjects in order to reduce experimental error. While planning a study a researcher needs to design an experiment in such a manner that the similarity is ensured among the experiential groups. The experimental error is controlled by controlling the effect of extraneous variables. To design an experiment a researcher must have the knowledge about homogeneity of experimental material or the subjects on which the experiment is required to be conducted. Besides, one should be able to identify those extraneous variables which may affect findings if not controlled. Depending upon the homogeneous conditions of subjects, an experimental design is identified. There are ways and means in testing the efficiency of design used in a research study. The efficiency of two different designs in the same experiment may be compared by using the error variance. This has been shown in Chapter 2. To have the control in an experiment and ensuring maximum accuracy in findings, Ronald A. Fisher has suggested the three basic principles of design of experiment, namely, Randomization, Replication, and Blocking.

Randomization

One of the main principles of design of experiment is randomization. *Randomization* refers to randomly allocating treatments to the subjects. Randomization ensures similarity in the experimental groups. It controls bias and extraneous variables which might affect findings of the study. Readers must note that the random selection of subjects and random allocation of treatments are two different things. Consider a study in which three different types of beverages, tea, coffee, and soft drink, are compared for their effect on reaction time. If 30 subjects are selected in the study let us see how randomization is done. Firstly, the initial sample of 30 subjects is selected randomly from the population of interest. Out of these subjects three subjects are randomly selected and the treatments are allocated randomly to them. Then another three samples are selected randomly from the remaining lot and treatments are again randomly allocated to them. In this all 30 subjects are assigned to three different treatment groups. In this study selecting 30 sample subjects randomly from the population of interest does not ensure that the treatment groups are similar, but helps the researcher to generalize findings about the population from which the sample has been drawn.

In other words, random selection of subjects ensures external validity in findings. Complete randomization is only possible if subjects are uniform. On the other hand, perfect random allocation of treatments to the subjects ensures that treatment groups are homogenous and do not contain any bias. This random allocation of treatments to the subjects ensures internal validity in findings. If external and internal validity are ensured, one can be quite sure in the above experiment that whatever the effect of a particular beverage on the subject's reaction time is observed, it is due to the beverage only and not due to any other reason. Besides this, randomization also provides the validity of F-test. Further, assumption of independence of observations in F-test also gets satisfied due to randomization.

Replication

Replication refers to repeating the treatment a number of times on different subjects. It is a fact that a treatment applied on single subject does not provide sufficient evidence of the effect of that treatment, so replication is needed. Single observation also does not provide the valid estimate of the parameters in the study so replication of treatments is essential. It is also a fact that standard error of sample mean or difference of sample means (group means) is inversely proportionate to the replication of treatments. So if number of replication increases the error variance or standard error decreases. If the treatment is effective the average effect of replication will reflect its experimental worth. If it is not few subjects in the sample who may have reacted to the treatment will be negated by the majority of subjects who were unaffected by it. If the above mentioned experiment of beverages is administered on three subjects only, and if the soft drink is found to improve the reaction time, the result may not be acceptable until unless this result is observed on most of the subjects in the sample. Thus, replication reduces variability in experimental results and provides confidence to the researcher in drawing conclusion about the effect of treatment on dependent variable.

Blocking

Blocking is a technique of reducing experimental error by including an extraneous variable in the experiment. Blocking refers to dividing heterogeneous experimental units into homogenous blocks so that the units in the blocks are homogeneous. In other words whole experimental material is divided into homogeneous strata. Blocks are made if experimenter has some knowledge about the experimental material or subjects prior to conducting an experiment, through pilot study or uniformity trial or some prior studies. Blocking technique is used if an experimenter knows that the variability exists among the subjects. Generally, size of the block should be equal to the number of treatments. After dividing the experimental units into blocks, the treatments are randomly allocated in each block. Blocking enhances precision in the study by reducing the experimental error. In the above example of beverages if gender is considered as blocking variable, the blocks of male and female may be made in which treatments may be allocated randomly.

STATISTICAL DESIGNS

It is important for the researchers to know about various kinds of statistical designs so as to choose the appropriate one for their study to obtain reliable findings. Selection of design depends upon many parameters such as number of factors to be investigated, variability of the experimental units, and degrees of precision required. In empirical research, studies can be classified in two categories; single factor studies and multifactor studies. In single factor study the effect of only one factor on some dependent variable is investigated, whereas, in multifactor studies effect of two or more independent factors on some dependent variable is investigated. Multifactor studies are also known as factorial experiment. Factorial experiment may have two or more independent factors, each having two or more levels. All these studies can be conducted by using any of the three basic statistical designs namely; Completely Randomized Design, Randomized Block Design and Latin Square Design. Choice of using these designs depends upon the knowledge about variability of the experimental materials or subjects. We shall discuss these designs briefly later in this chapter.

In the example of the beverages discussed above the effect of only one factor is investigated, hence the design used for analysis would be one factor design. Here the treatment factor has three levels and therefore three treatment conditions are to be compared for their effectiveness. But if along with the beverage, the effect of duration is also required to be investigated, the experiment is said to be the two factor(or multifactor) study. Similarly, more than two factors can also be simultaneously investigated and the design in such situation would be known as multifactor design. If the effect of more than two factors is investigated simultaneously, the analysis becomes very complex and therefore researchers usually investigate the effect of either one or two factors only. Whatever design a researcher uses, it is analyzed by using the group of techniques known as analysis of variance (ANOVA). In any experimental design the three basic principles of randomization, replication and blocking are used. Selection of the design of experiment depends upon the fact as to how we wish to carry out these principles. All types of designs discussed above will be explained with the help of examples in the following sections.

Completely Randomized Design

Completely randomized design is the simplest design used by a researcher to test the effectiveness of one factor (with two or more levels) on some dependent variable. This design is used when the entire experimental units (or subjects) are homogeneous in all respect and the experimenter has full control on the experiment. Under such strict control whatever the effect of the independent variable is observed on the dependent variable, can be totally attributed due to the independent variable only. This kind of experiment is also known as laboratory experiment. In completely randomized design

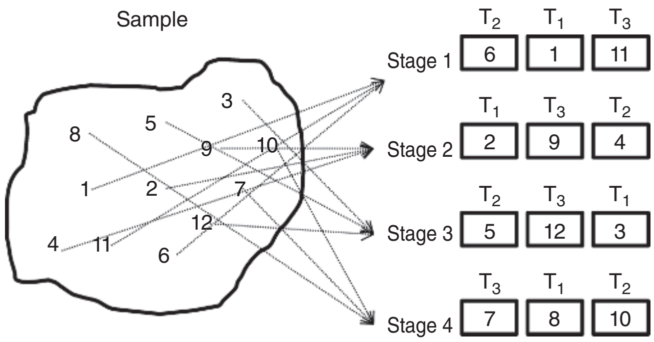

Figure 1.1 Layout of the completely randomized design

sample size in different treatment groups may differ. In this design treatments are randomly allocated to the subjects.

If the effect of three treatments, T_1, T_2, and T_3, are to be compared for their effectiveness on some dependent variable by administering them on a randomly drawn sample of size *n*, let us see how the treatments are allocated to the subjects in a completely randomized design. In the first stage, three subjects are randomly selected from the sample and all the three treatments are randomly allocated to these subjects. Then again three subjects are randomly selected from the remaining subjects in stage 2 and the treatments are again randomly allocated to these three samples. This process will go on till all the sample units are allocated in one or the other treatment groups. The readers must note that in this design each sample unit will get one and only one treatment. The layout of this design in which allocation of three treatments on 12 sample units has been done is shown in Figure 1.1. Randomization of treatments to the sample units ensures internal validity in the study.

Consider an experiment in which the effect of music is to be seen on concentration. The researcher may have three different types of music; classical, instrumental, and orchestra. This experiment can be organized by using completely randomized design in which the treatments can be randomly allocated to the samples as discussed above.

Randomized Block Design

Randomized block design (RBD) is used in a situation where the experimental units (or subjects) are not homogeneous. The reason of non-homogeneity of experimental units is known to the experimenter before applying the RBD. This reason may also be called classification or categorical variable not the treatment. Levels of these variables are also known to the experimenter. In using this design the entire sample is divided into blocks of homogeneous subjects and then the treatments are randomly allocated to these subjects within each block. Here the block is known as classificatory variable. The experimenter divides the sample into as many blocks as the levels of categorical variable. In a randomized block design, if *m* treatments are to be replicated *r* times then *r* blocks need to be created and $m \times r$ subjects need to be selected in the

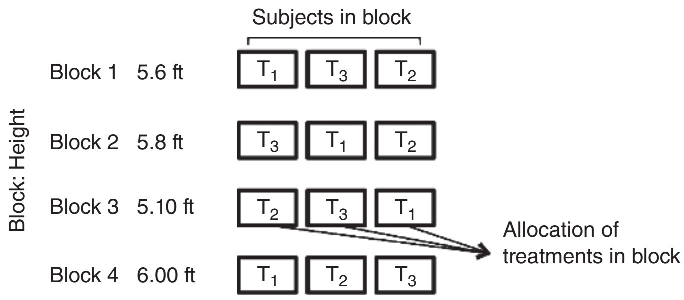

Figure 1.2 Layout of the randomized block design

study. Each of the *m* treatments is assigned randomly to only one subject in each block. The random allocation of treatment in each block is done independent of other blocks. The blocks are made on that variable which is supposed to affect dependent variable. Consider a study in which the effect of three different types of teaching methodology is compared to see their effect on the improvement of the subject's learning efficiency. It is a known fact that learning efficiency depends upon IQ level of the subjects. Thus, if the subjects vary in their intelligence level, the blocks may be made on IQ. For instance if this experiment is conducted on 12 subjects to compare the effectiveness of three teaching methodology (T_1, T_2, and T_3), the subjects may be divided into four blocks of different IQ levels, having three subjects in each. All the three treatments will then be randomly allocated in each block. The layout for 12 subjects, grouped in 4 blocks, can be shown by the Figure 1.2. Here the blocks have been made on the basis of the IQ. The reader should note that the blocks must be homogeneous within itself and heterogeneous among themselves in relation to IQ.

Another variant of the randomized block design is the one in which treatments are replicated many times within each block. For instance, if the above experiment is conducted by making the block on gender instead of IQ, all the three treatments T_1, T_2, and T_3 would be replicated many times in each gender. For instance, if a study is conducted to compare the effect of these teaching methodologies on 120 subjects, 60 male and 60 female would be taken as the subjects in the study. The treatments T_1, T_2, and T_3 will then be randomly replicated 20 times in male as well as in female group. The reader should note that in each block the number of subjects should be in the multiple of treatments. The randomized block design is generally used in a situation where the performance of the subjects is known to vary with the variation in certain variable. For instance, performance may vary with age, height, gender, and so on. Hence, blocking can be made on such variables. This design is mostly used by the researcher because of its flexibility and robustness. However, it becomes less efficient as the number of treatments increases because in that situation the block size also increases. Hence, it is difficult to maintain the homogeneity in the blocks. This design is more efficient than the completely randomized design because experimental error is reduced due to blocking. A detailed comparison of these two designs has been shown in Chapter 2.

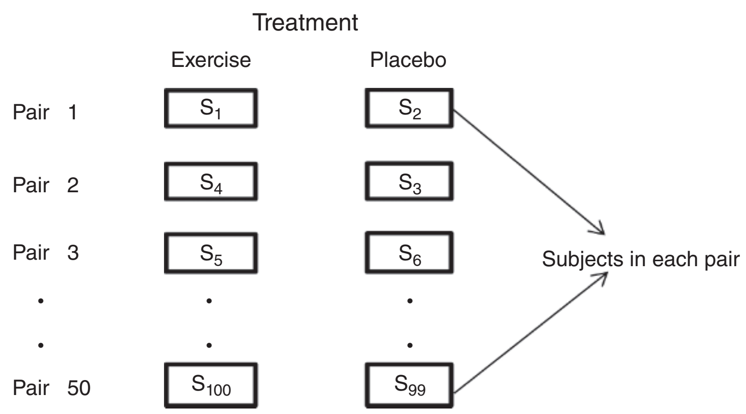

Figure 1.3 Layout of the matched pairs design

Matched Pairs Design Matched pairs design is a special case of the randomized block design. In this design the subjects are matched on some characteristics which are supposed to affect the experiment. For instance, if the effectiveness of exercise is to be investigated, matching of subjects should be done on the basis of gender and age because these variables affect the experiment, but if the matching is done on the basis of IQ no benefits would be derived from this design. This design can be used to compare only two treatments. The advantage of this design is that it can explicitly control more than one extraneous variable, for instance the age and gender in the above mentioned example. Here each matched pair is like a block. Consider an experiment in which effect of exercise on strength is to be studied on 100 students. In using this design we shall divide these 100 subjects into 50 pairs of subjects on the basis of gender and age. For instance, pair P1 may have both the subjects as female with 19 years of age and the pair P2 may have both the subjects as male with age 21 years. In each pair treatments are randomized. If one of the treatments is a low intensity exercise developed by the researcher and the other is a placebo, either of the two subjects may receive low intensity exercise and the other may receive placebo. Similarly treatments are randomized in each pair separately. The layout of this design can be understood by the Figure 1.3.

The matched pairs design is superior to the completely randomized design as well as randomized block design because here group becomes more homogeneous due to matching but the drawback is that it can compare only two treatments.

Latin Square designs

Latin square design is useful in a situation where the variability in the experimental material exist due to two extraneous variables in the experiment. The knowledge about the variability must be known to the researcher in advance for using this design. Since in RBD, only one extraneous variable affects the dependent variable hence we included one blocking variable in the experiment but in this design since effect of two extraneous variables needs to be controlled hence the experimental material is

Block: IQ	Block: Age →	Teens	Mid age	Old age
	Low	T_1	T_2	T_3
	Average	T_3	T_1	T_2
	High	T_2	T_3	T_1

Figure 1.4 Layout of Latin square design

divided in two different blocks (rows and columns). Thus in this design two blocking variables are taken. Consider a study in which the effect of three different types of teaching methodology T1, T2 and T3 on learning efficiency needs to be investigated. If it is known to the researcher that the learning efficiency in this experiment is known to be affected by the IQ and age of the subjects then the subjects would be divided in row(IQ) and column(Age) blocks. One of the restrictions in this design is that one needs to have equal number of blocks on both the extraneous variables. Number of treatments should also be equal to the number of blocks of each variable. Let us take three IQ and three Age groups in this study. Thus, three treatments can be taken in this study which can be replicated on three subjects. While allocating treatments to the subjects, each treatment should occur exactly once in each row and exactly once in each column. If T1, T2 and T3 represent traditional teaching, audio-visual teaching and flexible teaching methods respectively then one of the layout in LSD can be as shown in Figure 1.4. The main advantage of this design is that it requires less number of subjects to investigate different treatments. For instance, in this experiment 27(3×3×3) treatments can be compared by using only nine subjects.

FACTORIAL EXPERIMENT

If we consider more than one manipulated variables at different levels and their interaction effects are also to be judged then the factorial experiment is considered. Factorial experiment is used to investigate the effect of more than one factor at two or more levels on the dependent variable simultaneously. The factorial experiment is represented by m^p, where "p" represents the number of factors and m indicates the number of levels of each factor has. Thus, 2^2 factorial experiment represents that there are two factors each having two levels, and in all, have four treatment combinations. If a factorial experiment has $p+r$ independent factors so that each of the p factors has m levels and each of the r factors has n levels then the experiment is represented as $m^p \times n^r$ factorial experiment. For instance, if the factorial experiment has two factors A and B having levels 2 and 3, respectively, then it is called 2×3 factorial experiment.

Besides investigating the effect of each factor on the dependent variable, one can simultaneously investigate the interaction effect (joint effect) of the two factors on the dependent variable also. Thus, in factorial experiment all the levels of one factor may be compared in each level of the other factor. The simplest factorial experiment is 2×2, in which each factor has two levels. If the factor A has two levels (A_1, A_2) and the factor B has three levels (B_1, B_2, and B_3), this experiment is known as 2×3 factorial experiment. Thus, in this experiment six treatment combinations (A_1B_1, A_1B_2, A_1B_3, A_2B_1, A_2B_2, A_2B_3) need to be compared and therefore six groups of random samples need to be selected. If experimenter decides to replicate each treatment condition on five subjects, $30(=5 \times 2 \times 3)$ subjects need to be randomly selected from the population of interest. This experiment may be conducted in CRD, RBD or LSD designs as per the requirement of experiment. If the factorial experiment is conducted in CRD then all six treatment conditions shall be randomly allocated to these 30 subjects in sample so that each treatment is received by five subjects thus making five experimental groups. On the other hand due to heterogeneity considerations, if the factorial experiment is using RBD then these six treatment conditions shall be randomly allocated in each block. In this experiment the main effect of both the factors A and B can be investigated along with the simple effect of each factor simultaneously. The main effect can be defined as the effect of first independent variable (Factor A) on the dependent variable across all the levels of the second independent variable (Factor B). The interaction is ignored for this part. Just the rows or just the columns are used, not mixed. This is the part which is similar to one-way analysis of variance. Each of the variances calculated to analyze the main effects (Rows and Columns) is like between variances.

The advantage of factorial experiment is that one can test the significance of interaction between the two or more factors. The interaction can be defined as the joint effect of two factors or more on the dependent variable. It can also be defined as the effect that one factor has on the other factor.

In factorial experiment if the interaction effect is not significant, the main effect becomes meaningful. On the other hand, if the interaction effect is significant, the simple effects need to be investigated. The simple effect is the effect of one independent variable on the dependent variable in each level of the other factor. It is investigated by comparing the effect of all the levels of one factor in each level of the other factor.

Consider an experiment in which the effect of mental exercise (A) with three different intensities (low, medium, and high) and environment(B) with three different climatic conditions (hot, humid, and cold) are to be investigated to see their impact on the task efficiency of the subjects. This study can be conducted as a 3×3 factorial experiment organized using any of the three designs namely CRD, RBD or LSD depending upon whether the subjects in the sample have no variability, one dimensional variability or two dimensional variability respectively. Thus, in this design nine treatment combinations need to be compared. If an experimenter decides to replicate each treatment on five subjects using CRD then this experiment is called as 3^2 factorial experiment conducted in CRD with five replications. The layout design of this experiment can be shown by the Figure 1.5.

TERMINOLOGIES IN DESIGN OF EXPERIMENT

Understanding the following terminologies shall facilitate the readers to learn different designs discussed in this book in a better way.

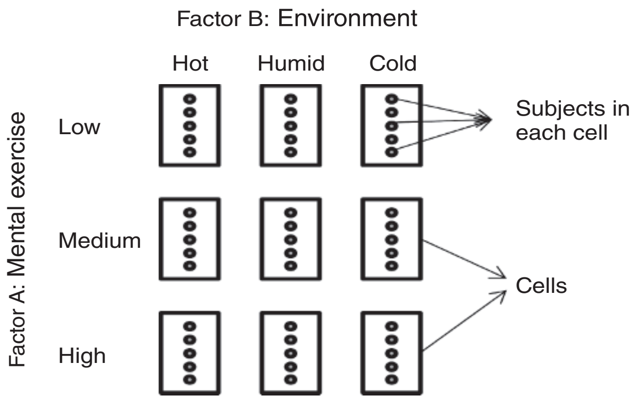

Figure 1.5 Layout of the 3×3 factorial design

Subject

Subject is usually used for an individual on whom an experiment is conducted. In investigating the effect of incentives on buying behavior, the individuals to which the incentives are offered are known as subjects.

Experimental Unit

Experimental unit can be defined as the subject/object on which the treatment is to be administered in an experiment. If the effect of beverages is investigated on concentration, the subjects consuming beverages are known as experimental units. Similarly if an impact of government policy is to be studied on the sports facilities in schools, the school would be an experimental unit.

Factor and Treatment

Factor can be defined as the independent variable whose effect is to be seen on the dependent variable whereas, different levels of factors are known as treatments. For instance, in beverages experiment, beverage is known as factor and different types of beverages are referred as treatments. In experimental research a researcher is allowed to manipulate the independent variable and therefore if one decides to compare the effect of three different types of beverages, say beer, coffee, and tea, on concentration, the beverages is said to have three levels or three *treatments*. The beverages will be known as *Factor* or *independent variable*. In other words, the *level* of a factor is the number of variation in the independent variable which an experimenter wishes to compare in relation to the response on the dependent variable.

In investigating the effect of independent variable on some dependent variable where the independent variable has not been manipulated by the experimenter is known as *classificatory variable*. For instance, to see the effect of gender on IQ among the group of students, the gender is a classificatory variable because gender has not been manipulated by the experimenter and the subjects have been classified on the basis of their possessed characteristics.

Criterion Variable Criterion variable can be defined as a variable in which the investigator is interested to see as to how it behaves due to manipulation in the independent variable. It is usually a dependent variable in the study. The criterion variable is also known as response variable. There can be many criterion variables in the study. Consider an experiment in which an investigator wishes to investigate the effect of chocolate type on its taste. Here taste is a criterion variable which is of prime importance to the researcher. The investigator manipulates chocolate by having its different variants say white, milk, and dark whose effects may be compared on taste. Thus, in this case taste depends upon different types of chocolate. Further, in this experiment if the investigator is also interested to see the effect of chocolate type on crunchiness along with the taste, there would be two criterion variables, that is, taste and crunchiness. Similarly depending upon the objectives of the study the researcher may have several criterion variables in the study. Such study is a case of multivariate analysis of variance.

Variation and Variance

Variation refers to the spread of scores around mean values, whereas the variance can be defined as a measure of variation. Variance is simply the square of standard deviation and is given by the following formula:

$$\sigma^2 = \frac{1}{n}\sum (x - \mu)^2 \tag{1.1}$$

This σ^2 is the population variance and can be estimated by the sample mean square, S^2, as given by the formula (1.2).

$$S^2 = \frac{1}{n-1}\sum (x - \bar{x})^2 \tag{1.2}$$

The estimate of variance due to independent variable(s) and variance due to error are estimated to solve different designs. Detail procedure has been shown in Chapter 2.

Experimental Error

Experimental error can be defined as an error which cannot be controlled in an experiment. Mostly this is the result of individual variation during an experiment. One cannot attribute the reason for such variation. By having control on the experiment and using proper experimental design, the researcher tries to reduce it. For instance, in comparing the effect of three different chocolates on taste, the observed effect on the subject's taste cannot be attributed to the chocolate type alone but it is partially due to the subject's characteristics such as age, gender and socioeconomic status as well. This variation which cannot be accounted for due to the chocolate is known as experimental error.

External Validity

External validity refers to the extent of generalizibility of research findings to the population from which the sample is derived. To ensure the external validity, it is important that the sample is randomly drawn from the population of interest. Thus, in a study if the sample is randomly drawn by using an appropriate probability sampling technique, external validity can be ensured. On the other hand, if the sample is drawn by using any of the nonprobability sampling method, the study lacks external validity.

Internal Validity

In an experimental research if the effect on a criterion variable can be attributed due to manipulation in an independent variable, the research study is said to have internal validity. Internal validity refers to the extent of which one can say that the variation observed in the dependent variable is due to the variation in the independent variable. To ensure the internal validity, the external variance should be controlled in the experiment. The randomization of experimental units to the subjects is one of the best ways to ensure the internal validity in the study.

CONSIDERATIONS IN DESIGNING AN EXPERIMENT

We have seen that in experimental research a researcher is interested to see the impact of some independent variable or factor on the criterion variable. A *factor* which can be manipulated by the researcher to see its impact on the criterion variable is known as treatment variable. Examples of treatment variables are exercise program, nutrition, pranayama, training intensity, incentives, and so on. In all such cases an experimenter can manipulate them. A treatment variable is basically an independent variable. If the independent variable cannot be manipulated by the researcher, it is known as classificatory variable. The *classificatory variable* can be defined as some preexisting characteristics of the subjects on the basis of which they can be classified. The variables like gender, socio economic status, sports category, geographical region, and weather conditions are all classificatory variables because these variables cannot be manipulated by the researcher. For instance, in seeing the impact of three different types of nutritional supplements on the muscular strength, the nutritional supplement is a treatment variable because it can be manipulated by the experimenter by choosing its different types. However, in investigating the effect of gender on the IQ, gender is a classificatory variable because the gender cannot be manipulated by the experimenter.

In two-factor design, in which the experimenter is interested to see the impact of two factors simultaneously on some criterion variable, one of the factors can be treatment variable whereas the other can be a classificatory variable. Consider an experiment in which the effect of sleep deprivation (with three different durations) is to be seen on the task performance in male and female subjects. In this case the two independent variables are sleep deprivation and gender, whereas the criterion variable is the task performance. Out of these two independent variables the sleep deprivation is a treatment variable because the experimenter can manipulate its duration, whereas the gender is a classificatory variable because it cannot be manipulated as the subjects are classified on the basis of their gender, a preexisting criterion among the subjects prior to conducting the experiment.

Most of the time a researcher manipulates an independent variable in his experiment because by doing this more control can be observed in the experiment resulting more accuracy in the findings. Manipulating independent variable depends upon the nature of the study. Sometime the objective of the study is such that the independent variable cannot be manipulated by the researcher, in that case the independent variable is taken as a classificatory variable. For instance, if the effect of consuming tobacco on the athlete's performance is to be investigated, the subjects can be classified into nontobacco, occasional user, and regular user categories. In this case independent variable cannot be manipulated by the researcher as the subjects cannot be forced to consume tobacco to see its impact on their athletic performance. In fact it is unethical if somebody does that.

Choosing statistical design depends upon the objective of the study and the knowledge about the heterogeneity of experimental material (subjects). An experimenter is required to manage different types of variances in such a way in the experiment so as to get the reliable findings. The purpose of each and every design is to reduce the error in the experiment for enhancing the reliability of findings. This is done by maximizing the systematic variance, controlling the extraneous variance, and minimizing the error variance in the study (Verma, 2014). By following these guidelines reliable findings in the experimental study can be achieved. Detailed discussions about these aspects have been made in the following sections.

Systematic Variance

Systematic variance refers to the measure of variation resulting from manipulating the independent variable by the researcher. The researcher must choose different levels of the independent variable in such a manner so as to have the maximum variability in the criterion variable. For instance, in the above mentioned sleep deprivation experiment the researcher must choose the duration in such a way so as to achieve the maximum variance in the task performance of the subjects in different treatment groups. Since the researcher has a freedom to manipulate this independent variable, he might choose to have three durations of sleep deprivation as 24 hours, 30 hours, and 36 hours. Of course choosing these durations must be based on some scientific theory or literature. However, if an experimenter decides to have these three durations as 22 hours, 23 hours, and 24 hours, this may not produce maximum variance in the criterion variable during experiment. Thus, a researcher must use this opportunity to maximize the systematic variance by scientifically choosing the levels of the treatment variable(s).

Extraneous Variance

While investigating the effect of independent variable on some criterion variable the researcher must ensure that whatever the variation is observed in the criterion variable, most of it can be explained due to the manipulation of the independent variable. This can be achieved by controlling the effect of extraneous variables and making the experimental groups as similar as possible. This can be done by using any one or more methods discussed below.

Randomization Method Random allocation of treatments to the subjects is the most powerful method of controlling the external variance in the experiment. It ensures the effect of all external variables to be equally distributed in the treatment groups. Randomization ensures similarity of all the treatment groups. This ensures internal validity by removing the confounding effect of the extraneous variables in the findings of a research experiment. In fact randomization is one of the main principles of the design of experiment. Randomization also controls bias in allocating treatments to the subjects. Another advantage of randomization is that it ensures normality of data in different groups which is one of the main assumptions in solving the design by using the analysis of variance. Thus, a researcher should not only draw the random sample from the population of interest, but should also allocate the treatments randomly to the subjects for controlling the external variance in the experiment.

Elimination Method Elimination method is another way of controlling the external variance in the experiment. In this method the extraneous variable which affects the criterion variable during experiment is stabilized in the study. Let us consider an experiment in which the effect of three different exercise programs is to be compared on weight reduction. Let us further assume that these programs are administered on three different groups of subjects which include male and female both. In this case weight reduction cannot be solely due to the variation in the exercise programs, but partly affected by the gender as well. In order to done away the effect of gender, the whole experiment can be organized either on male or female. In other words, effect of gender can be controlled by eliminating either of the two genders from the study. The problem with such study is that the findings cannot be generalized for the entire population. In this case if the female subjects are eliminated, the findings can only be applicable to the male subjects and not for the females; hence, elimination method lacks in external validity. Thus, if a researcher has an idea about the extraneous variable which affects criterion variable in an experiment, the elimination method may control the external variance provided objective of the study demands for it.

Matching Group Method Another method of controlling external variance is by means of matching the group on some extraneous variable. If the effect of three training programs (T_1, T_2, and T_3) is to be seen on the shooting accuracy in basket ball, the subjects may be matched in the groups on the basis of their height because height is supposed to affect the shooting accuracy besides training programs. Using this method the height of all the subjects in the sample is measured first and scores so obtained are arranged either in ascending or descending order. After arranging the subjects according to their height, the first three subjects may be selected and the three treatments (in this case) can be randomly allocated to them. Again the next three subjects are selected and all the three treatments are randomly allocated to them. This way all the subjects are allocated to some or other treatment group. The readers must note that each subject will be classified into one of the three treatment groups. If the random sample consists of 12 subjects, matching of these samples on the basis of their height can be shown by the Figure 1.6.

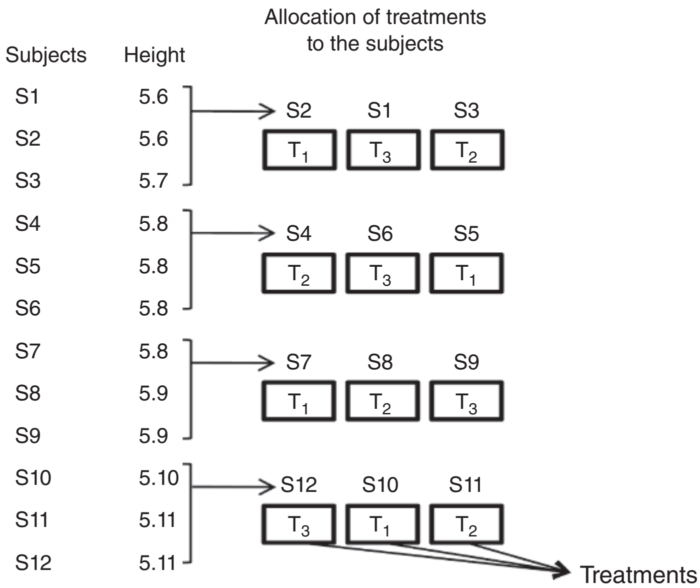

Figure 1.6 Allocation of treatments by matching the subjects

Adding Additional Independent Variable The above discussed three methods of controlling extraneous variance, that is, randomization, elimination, and matching, are known as nonstatistical methods. If the variability of an extraneous variable cannot be controlled by any of the nonstatistical methods, one may include it in the design. For instance, if an experimenter is interested to compare the effect of three different conditioning programs (T_1, T_2, and T_3) on the general fitness among the college students, in that case if the experimental groups consist of male and female both, gender can be included as an additional independent variable in the design. The gender is included in the design as an additional independent variable because it is considered to affect the criterion variable, that is, general fitness during experimentation. This variable is known as blocking variable. The purpose of adding gender as an independent variable is only to reduce the error variance in the design, as some part of the error variance can be explained by the variable gender. Besides reducing the error variance, this independent variable does not serve any other purpose in the experiment. The reader must note that by the inclusion of extraneous variable in the design the error variance is reduced but at the same time degrees of freedom of error variance also gets reduced. Thus, the design will become more efficient only when the extraneous variable is known to affect the criterion variable significantly. If the three treatments in the above mentioned experiment are different intensities of the conditioning program, that is, low, medium, and high, the layout of the design can be shown by the Figure 1.7.

Statistical Control We have seen that if nonstatistical procedures are not sufficient to control the extraneous variance, the statistical procedure can be adopted by including the extraneous variable in the design. Other statistical procedures through

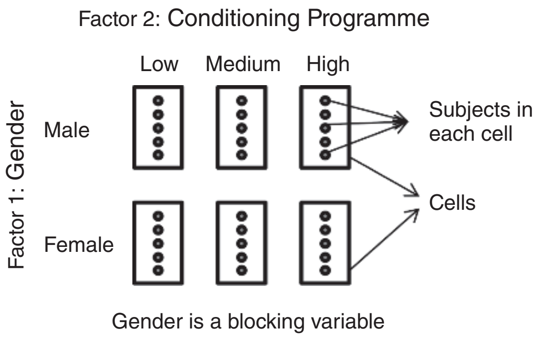

Figure 1.7 Layout of the design after including extraneous variable in the design

which control can be observed in the experiment are repeated measure designs and analysis of covariance (ANCOVA) design. In using the repeated measures design subjects serve their own control; hence, error variance is reduced to a great extent. On the other hand, in ANCOVA design the extraneous variance is controlled by adjusting the effect on the criterion variable due to treatments in relation to the extraneous variable. In ANCOVA design the extraneous variable is also known as covariate. Since this book is specifically meant for the repeated measures designs, various repeated measures designs have been discussed in different chapters along with their solutions by using the SPSS software. The ANCOVA design is out of scope of this book. The readers are advised to refer to the book Verma (2011) for this topic.

Error Variance

All nonstatistical and statistical methods discussed above to control the extraneous variance help in reducing the error variance also. Besides this by following some more guidelines the error variance can further be minimized in the design. For instance, the researcher must use the standard equipments so as to ensure the instrument reliability in data collection. Similarly the researcher must follow the proper instructions applicable in data collection of different parameters and if the study is of survey type, a clear-cut instruction must be given to the field investigators for reliable data collection.

Maximizing systematic variance, controlling extraneous variance, and minimizing error variance facilitates an investigator to compare systematic variance against error variance. Minimizing error variance gives the systematic variance a chance to show its significance. Managing all types of variances ensures that whatever the effect observed in the criterion variable is genuinely due to the treatment variable.

EXERCISE

1.1. What happens if the principles of design of experiments are not followed in conducting an experimental study?

1.2. Explain the concept of external and internal validity in an experimental study by means of examples.

1.3. What are the benefits of randomization in a research design? Explain the difference between random selection of sample and random allocation of treatments to the subjects.

1.4. What is factorial experiment? Explain its advantage in comparison to that of single factor study.

1.5. Why randomized block design is considered to be the superior design in comparison to that of completely randomized design. Discuss the layout of these two designs.

1.6. Explain the matched pairs design and discuss its advantage over completely randomized design and randomized block design.

1.7. What are the different ways and means in controlling extraneous variance in an experimental study?

1.8. In an experimental study, how different variances should be controlled? What happens if an experimenter optimally maximizes systematic variance, controls extraneous variance, and minimizes error variance?

ASSIGNMENT

1.1. A researcher decides to organize a completely randomized design to investigate the effect of different types of beverages on the concentration. He has taken three different types of beverages and a sample of 15 subjects. Describe the layout of the design in the study.

1.2. An investigator wishes to compare the effect of three different types of exercises on flexibility. He has selected nine male and nine female subjects in the study. Theory suggests that the experimental results would be affected by the gender difference. Which design you would prefer. Show the layout design in the study.

1.3. By organizing a matched pairs design an investigator wishes to see the impact of garlic on the cholesterol level of the subjects. He feels that the age and activity levels of the subjects will affect the findings in the study; hence, took these two criteria for matching the group. Suggest a plan of study and show the layout design.

BIBLIOGRAPHY

Addelman S. The generalized randomized block design. Am Stat 1969;23(4):35–36. DOI: 10.2307/2681737.JSTOR 2681737.

Addelman S. Variability of treatments and experimental units in the design and analysis of experiments. J Am Stat Assoc 1970;65(331):1095–1108. DOI: 10.2307/2284277. JSTOR 2284277.

Ader HJ, Hand DJ. *Advising on Research Methods: A Consultant's Companion*. Johannes van Kessel Publishing. 2008.

Anderson DW, Kish L, Cornell RG. On stratification, grouping and matching. Scand J Stat 1980;7(2):61–66. JSTOR 4615774.

Arceneaux K, Gerber AS, Green DP. A cautionary note on the use of matching to estimate causal effects: an empirical example comparing matching estimates to an experimental benchmark. Sociol Method Res 2010;39(2):256–282. DOI: 10.1177/0049124110378098.

Bisgaard S. Must a process be in statistical control before conducting designed experiments? Qual EngASQ 2008;20(2):143–176.

Box F. (1978) R. A. Fisher: The Life of a Scientist, New York: Wiley, http://en.wikipedia.org/wiki/Special:BookSources/0471093009 ISBN 0-471-09300-9.

Dehejia RH, Wahba S. Causal effects in nonexperimental studies: reevaluating the evaluation of training programs. J Am Stat Assoc 1999;94(448):1053–1062. DOI: 10.1080/01621459.1999.10473858.

Dehue T. Deception, efficiency, and random groups: psychology and the gradual origination of the random group design. Isis 1997;88(4):653–673. DOI: 10.1086/383850PMID 9519574.

DiNardo J. Natural experiments and quasi-natural experiments. In: Durlauf SN, Blume LE, editors. *The New Palgrave Dictionary of Economics*. 2nd ed. Palgrave Macmillan; 2008. DOI: 10.1057/9780230226203.1162.

Dunning T. *Natural Experiments in the Social Sciences: A Design-Based Approach*. Cambridge University Press; 2012.

Gates CE. What really is experimental error in block designs? Am Stat 1995;49(4):362–363. DOI: 10.2307/2684574JSTOR 2684574.

Goodwin CJ. *Research in Psychology: Methods and Design*. USA: John Wiley & Sons, Inc.; 2005.

Hacking I. Telepathy: origins of randomization in experimental design. Isis 1988; 79(3):427–451. DOI: 10.1086/354775JSTOR 234674.MR 1013489.

Heitink G. *Practical Theology: History, Theory, Action Domains: Manual for Practical Theology*. Grand Rapids, MI: Wm. B. Eerdmans Publishing; 1999. p 233. ISBN: 9780802842947.

Kempthorne O. *The Design and Analysis of Experiments (Corrected Reprint of (1952) Wiley ed.)*. Robert E. Krieger; 1979. ISBN: 0-88275-105-0.

McLeod, S. A. (2007). Experimental Design. Retrieved from http://www.simplypsychology.org/experimental-designs.html.

Montgomery D. *Design and Analysis of Experiments*. 8th ed. Hoboken, NJ: John Wiley & Sons, Inc; 2013. ISBN: 9781118146927.

Moore DS, Notz WI. *Statistics: Concepts and Controversies*. 6th ed. New York: W.H. Freemanpp. Chapter 7: Data ethics; 2006. ISBN: 9780716786368.

Moore DS, McCabe GP, Craig BA. *Introduction to the Practice of Statistics*. 6th ed. New York: W. H. Freeman and Company; 2009.

Rubin DB. Matching to remove bias in observational studies. Biometrics 1973;29(1):159–183. DOI: 10.2307/2529684. JSTOR 2529684.

Verma, J. P. (2011). Statistical methods for sports and physical education. Tata McGraw Hill Education Private Limited.

Verma, J. P. (2014). Statistics for Exercise Science and Health with Microsoft Office Excel. John Wiley & Sons.

Wilk MB. The randomization analysis of a generalized randomized block design. Biometrika 1955;42(1–2):70–79. JSTOR 2333423.