**Notes – Ch13 Experimental Design and Analysis of Variance**

**Analysis of variance (ANOVA**): is a set of techniques that allow us to compare two or more sample means at the same time. Statistical studies can be classified as either experimental or observational.

In **experimental statistical study**, an experiment is conducted to generate the data. An experiment begins with identifying a variable of interest. Then one or more other variables, are identified and controlled, and data are collected about how those variables influence the variable of interest. There are 3 types of experimental design: a completely randomized design, a randomised block design, and a factorial experiment. For each design analysis of variance (ANOVA) can be used to analyse the data available.

In **observational study,** data are usually obtained through sample surveys and not a controlled experiment. Good design principles are still employed, but the rigorous controls associated with an experimental statistical study are often not possible. ANOVA can be used to analyse data obtained through observational study.

Cause-and-effect relationships can be difficult to establish in observational studies; such relationships are easier to establish in experimental studies.

**Experiment:** A study or investigation designed for the purpose of examining the effect that one variable has on the value of another variable.

**Experimental units:** The objects of interest in the experiment

**Response or dependent variable**: The variable for which a value is measured or observed. In ANOVA, the dependent variable will be a quantitative variable—for example, soft drink consumption, examination score, or the time required to type a document.

**Factor or independent variable**: A variable that is observed or controlled for the purpose of determining its effect on the value of the response variable. In ANOVA, the factor can be qualitative (e.g., marital status) or quantitative (e.g., age group). The experiment may involve different factor levels (categories).

**Treatment:** Each specific level of a factor (or, in multiple-factor experiments, the intersection of a level of one factor with a level of another factor) is referred to as a treatment. When there is only one factor in an experiment, factor levels and treatments are synonymous.

**Replication:** The number of times each experimental condition is repeated in an experiment.

Eg. A manager wants to conduct an experiment to study the mean output times of 3 types of assembly methods. The output of the assembly is the response variable. The experimental units are the workers involved in the assembly. Each assembly method is a factor and all the assembly methods are the treatments.

**Statement of hypothesis for ANOVA:**

H0: μ1 = μ2 = μ3

Ha: Not all population means are equal

If H0 is rejected, we cannot conclude that all population means are different. Rejecting H0 means that at least two population means have different values.

Eg. We have taken samples of employees who have been trained by three different methods. We want to test the effectiveness of the training methods on the productivity of the employees. If we conclude that the sample means do not differ significantly (H0 is true), we can infer that the choice of training method does not influence the productivity of the employees. If we find differences among the sample means are to large to attribute to sampling error, we can infer that the method used in training does influence the productivity of the employees.

Assumptions for analysis of variance:

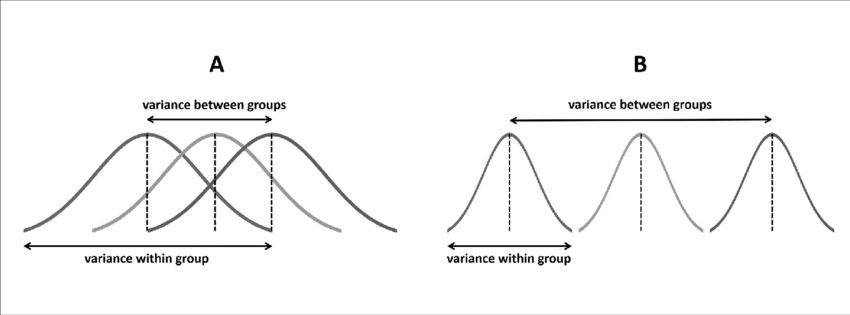
* For each population, the response variable is normally distributed.
* The variance of the response variable, denoted σ 2, is the same for all of the populations.
* The observations must be independent.

|  |  |
| --- | --- |
| Sample means are close to each other when Ho is true | Sample means come from different sampling distribution and are not as close together when Ho is false |

ANOVA is based on comparison of two different estimates of the variance of our overall population. The variation can be viewed in terms of

(1) variation between the groups, reflecting the effect of the factor levels

(2) variation within the groups, which represents random error from the sampling process.



Groups are not significantly different. Groups are significantly different.

**Completely randomized design:** An experimental design in which the treatments are randomly assigned to the experimental units

H0: μ1 = μ2 = μ3

Ha: Not all population means are equal

Where μj = mean of the jth population

A random sample of size nj has been selected from each of the k populations or treatments.

xij = value of observation i for treatment j

nj = number of observations for treatment j

= sample mean for treatment j

sj = sample standard deviation for treatment j

is the sample mean for treatment j

is sample variance for treatment j

is the overall sample mean where nT = n1 + n2 + ….. nk

If the size of each sample is n, nT = kn, then

**Between treatment Estimate of Population Variance (MSTR):** Analysis of variance indicates that we must obtain one estimate of the population variance from the variance among the three sample means. This is called Between Treatment Estimate of Population Variance or **Mean Square due to treatments (MSTR).** K – 1 represents the degrees of freedom associated with **SSTR(Sum of Square due to treatments)**.

where SSTR =

If H0 is true, MSTR provides an unbiased estimate of σ2 However, if the means of the k populations are not equal, MSTR is not an unbiased estimate of σ2 in fact, in t= hat case, MSTR should overestimate σ2.

**Within treatment Estimate of Population Variance(MSE):** Analysis of variance indicates that we must obtain another estimate of the population variance based on the variance within three samples. This is called **Within Treatment Estimate of Population Variance or Mean Square due to error (MSE).** We get a better estimate of the population variance by using a weighted average of all three sample variances. nT - k represents the degrees of freedom associated with **SSE (Sum of Square due to error)**.

where SSE =

Note that MSE is based on the variation within each of the treatments; it is not influenced by whether the null hypothesis is true. Thus, MSE always provides an unbiased estimate of σ2.

**F-test:** If the null hypothesis is true and the ANOVA assumptions are valid, the sampling distribution of MSTR/MSE is an F distribution with numerator degrees of freedom equal to k-1 and denominator degrees of freedom equal to nT - k . In other words, if the null hypothesis is true, the value of MSTR/MSE should appear to have been selected from this F distribution.

However, if the null hypothesis is false, the value of MSTR/MSE will be inflated because MSTR overestimates σ2. Hence, we will reject H0 if the resulting value of MSTR / MSE appears to be too large to have been selected from an F distribution with k-1 numerator degrees of freedom and nT - k denominator degrees of freedom. Because the decision to reject H0 is based on the value of MSTR/MSE, the test statistic used to test for the equality of k population means is as follows:

**ANOVA table for Completely Randomized Design**

The result of the preceding calculations can be displayed in a table form referred as ANOVA table. The sum of squares associated with the source of variation referred to as Total is called total sum of squares (SST).

and SST = SSTR + SSE

SST can be partitioned into two sums of squares: SSTR and SSE. The degrees of freedom corresponding to SST, can be partitioned into degrees of freedom corresponding to SSTR and SSE. The analysis of variance can be viewed as the process of **partitioning** the total sum of squares and the degrees of freedom into their corresponding sources: treatment and error.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Source of Variation** | **Sum of Squares** | **Degree of Freedom** | **Mean Squares** | **F** | **p-value** |
| **Treatments** | SSTR | k - 1 |  |  |  |
| **Error** | SSE | nT - k |  |  |  |
| **Total** | SST | nT - 1 |  |  |  |

**Test for Completely Randomised Design**

The null and alternative hypotheses are

H0: μ1 = μ2 = μ3

Ha: Not all population means are equal

Test Statistics:

Rejection Rule :

|  |  |
| --- | --- |
| p-value approach: | if then reject H0 |
| Critical value approach: | If then reject H0 |

Where the value of is based on F distribution with k – 1 numerator degrees of freedom and nT - k denominator degrees of freedom.

**Randomized block design:** An experimental design that employing blocking.

In the one-way, or completely randomized ANOVA, treatments are randomly assigned to all of the persons or other test units in the experiment. As a result, the composition of the treatment groups may be such that certain kinds of people or test units are overrepresented in some treatment groups and underrepresented in others, simply by chance. If the characteristics of the participants or test units have a strong influence on the measurements we obtain, we may be largely measuring the differing group compositions rather than the effects of the treatments.

**Blocking** is a process of using the same or similar experimental units for all treatments. The purpose of blocking is to remove a source of variation from the error term and hence provide a more powerful test for a difference in population or treatment means. In the randomized block design, persons or test units are first arranged into similar groups, or block. Each block is subjected to all the treatments. This allows us to reduce the amount of error variation. Although we are controlling, or blocking, one variable, our primary concern lies in testing whether the population means could be the same for all of the treatment groups.

A completely randomized design is useful when the experimental units are homogeneous. If the experimental units are heterogenous, blocking is often used to form homogeneous groups. Experimental studies in business often involve experimental units that are highly heterogeneous; as a result, randomized block designs are often employed.

For example, let’s assume we have randomly selected 12 citizens from a small community and these persons are to participate in an experiment intended to compare the night-vision effectiveness of four different headlamp designs. If we have treatment groups of equal size and randomly assign the treatments, it’s likely that the representation of older drivers would not be exactly the same in all four groups. This would reduce our ability to compare the headlamp designs, since night vision tends to decrease with age. In this situation, the value of the variable that we really want to measure (i.e., the distance at which a headlamp enables a suburban traffic sign to be read) is being strongly influenced by another variable (age category) that has not been considered in the experiment.

For example, in the night-vision experiment just described, use of the randomized block design would ensure that the treatment groups are comparable in terms of the age categories of their members. Exerting this control over the age-category variable (now referred to as a blocking variable) allows us to better compare the effectiveness of the headlamp designs, or treatments.