RESEARCHERID
THOMSON REUTERS

[Sharma * *et al.*, 6(7): July, 2017]
IC™ Value: 3.00

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

APPLICATIONS OF STATISTICAL TOOLS FOR DECISION MAKING UNDER UNCERTAINTY IN MANUFACTURING INDUSTRIES

Ankur Sharma[*1] & Payal[2]
[*1&2]Assistant professor, HMR Institute of Technology & Management

DOI: 10.5281/zenodo.828030

ABSTRACT

In order to survive in the present day global competitive environment, it now becomes essential for the manufacturing organizations to take prompt and correct decisions regarding effective use of their scarce resources. The content of this paper to promote the wider understanding and application of statistical methods for manufacturing decision making problems under uncertainty conditions. It is important for managers to know the statistical techniques that can be applied in industry and the ways in which these techniques can help them in their decision making. The aims of the study are the managers make decisions using Statistics. This paper will provide you with hands-on experience to promote the use of statistical thinking and techniques to apply them to make educated decisions, whenever you encounter variation in business data.

I. INTRODUCTION

Everyone is needed to be familiar with the Decision Making Process. We all rely on information, and techniques or tools, to help us in our daily lives. Operating a business also requires making decisions using information and techniques - how much inventory to maintain, what price to sell it at, what credit arrangements to offer, how many people to hire.

Manufacturing can be defined as the application of mechanical, physical and chemical processes to modify the geometry, properties and/or appearance of a given input material while making a new finished part/product. The type of manufacturing performed by an organization largely depends on the end product it produces. In the modern sense, manufacturing includes various interrelated activities, like product design, material selection, process planning, machine selection, maintenance planning and documentation, quality assurance, management and marketing of products (Rao, 2007). Today's manufacturing processes are caught between the growing needs for quality, high process safety, minimal manufacturing costs and short manufacturing times. In order to meet these demands, manufacturing processes need to be chosen in the best possible way. Selection of the manufacturing processes and optimal process parameter settings plays a pivotal role to ensure high quality of products, reduce manufacturing costs, trim down lead times and inventory levels, and increase the overall productivity of the manufacturing organizations. Decision makers in the manufacturing sector frequently face the problem of assessing a wide range of alternative options and selecting the best on e based on a set of conflicting criteria. It must be noted that in choosing the most appropriate alternative, there is not always a single definite criterion of selection, and the decision makers have to take into account a large number of criteria. Thus, there is a need for some simple, systematic and logical methods or mathematical tools to guide the decision makers in considering a number of conflicting selection criteria and their interrelations. The objective of any selection procedure is to identify the suitable evaluation criteria and obtain the most appropriate combination of criteria in conjunction with the real requirement. Thus, efforts need to be extended to identify those criteria that influence the best alternative selection for a given problem, using simple and logical methods, to eliminate the unsuitable alternatives, and select the most appropriate one to strengthen the existing selection procedures. In order to deal with those complex selection problems arising in the modern day manufacturing environment, various statistical tools that can be applied in manufacturing industries to reduce the uncertainty in decision making process.Today's good decisions are driven by data. In all aspects of our lives, and importantly in the business context, an amazing diversity of data is available for inspection and enlightenment. Moreover, business managers and professionals are increasingly encouraged to justify decisions on the basis of data.

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

II. STATISTICAL MODELING FOR DECISION-MAKING UNDER UNCERTAINTIES

Data must be collected according to a well-developed plan if valid information on a conjecture is to be obtained. The plan must identify important variables related to the conjecture, and specify how they are to be measured from the data collection plan, a statistical model can be formulated from which inferences can be drawn.

Data is known to be crude information and not knowledge by itself. The sequence from data to knowledge is: from Data to Information, from Information to Facts, and finally, from Facts to Knowledge. Data becomes information, when it becomes relevant to your decision problem. Information becomes fact, when the data can support it. Facts are what the data reveals. However the decisive instrumental (i.e., applied) knowledge is expressed together with some statistical degree of confidence. Fact becomes knowledge, when it is used in the successful completion of a decision process. Considering the uncertain environment, the chance that "good decisions" are made increases with the availability of "good information." The chance that "good information" is available increases with the level of structuring the process of Knowledge Management. The above figure also illustrates the fact that as the exactness of a statistical model increases, the level of improvements in decision-making increases.

Statistical Decision-Making Process

Unlike the deterministic decision-making process, such as linear optimization by solving systems of equations, Parametric systems of equations and in decision making under pure uncertainty, the variables are often more numerous and more difficult to measure and control. However, the steps are the same. They are:

1. Simplification
2. Building a decision model
3. Testing the model
4. Using the model to find the solution:
5. It can be used again and again for similar problems or can be modified.

Statistical methodology

Statistics is the mathematical science involving the collection, analysis and interpretation of data. A number of specialties have evolved to apply statistical theory and methods to various disciplines. Certain topics have "statistical" in their name but relate to manipulations of probability distributions rather than to statistical analysis

III. SAMPLING METHODS

Following are important methods in sampling:

Cluster sampling

With cluster sampling, every member of the population is assigned to one, and only one, group. Each group is called a cluster. A sample of clusters is chosen, using a probability method (often simple random sampling). Only individuals within sampled clusters are surveyed.

Stratified sampling

It can be used whenever the population can be partitioned into smaller sub-populations, each of which is homogeneous according to the particular characteristic of interest.

Random sampling

It is probably the most popular sampling method used in decision making today. Many decisions are made, for instance, by choosing a number out of a hat or a numbered bead from a barrel, and both of these methods are attempts to achieve a random choice from a set of items. But true random sampling must be achieved with the aid of a computer or a random number table whose values are generated by computer random number generators.

Mean: The arithmetic mean (or the average, simple mean) is computed by summing all numbers in an array of numbers (x_i) and then dividing by the number of observations (n) in the array.

Mean = = □ X_i /n, the sum is over all i's.

The mean uses all of the observations, and each observation affects the mean. The mean has valuable mathematical properties that make it convenient for use with inferential statistical analysis.

Weighted Mean: In some cases, the data in the sample or population should not be weighted equally, rather each value should be weighted according to its importance.

Median: The median is the middle value in an ordered array of observations. The median is often used to summarize the distribution of an outcome. Note that if the median is less than the mean, the data set is skewed to the right. If the median is greater than the mean, the data set is skewed to the left.

The mean has two distinct advantages over the median. It is more stable, and one can compute the mean based of two samples by combining the two means.

Mode: The mode is the most frequently occurring value in a set of observations. When the mean and the median are known, it is possible to estimate the mode for the unimodal distribution using the other two averages as follows:

$$\text{Mode} = 3(\text{median}) - 2(\text{mean})$$

This estimate is applicable to both grouped and ungrouped data sets.

IV. SHAPE OF A DISTRIBUTION FUNCTION

The Skewness-Kurtosis Chart

Skewness: It is a measure of the degree to which the sample population deviates from symmetry with the mean at the center. Skewness will take on a value of zero when the distribution is a symmetrical curve. A positive value indicates the observations are clustered more to the left of the mean with most of the extreme values to the right of the mean. A negative skewness indicates clustering to the right.. The reverse order holds for the observations with positive skewness.

Kurtosis: Kurtosis is a measure of the relative peakedness of the curve defined by the distribution of the observations. Standard normal distribution has kurtosis of +3. A kurtosis larger than 3 indicates the distribution is more peaked than the standard normal distribution. A value of less than 3 for kurtosis indicates that the distribution is flatter than the standard normal distribution.

These inequalities hold for any probability distribution having finite skewness and kurtosis.
In the Skewness-Kurtosis Chart, you notice two useful families of distributions, namely the beta and gamma families.

V. MEASURING THE QUALITY OF A SAMPLE

Average by itself is not a good indication of quality. You need to know the variance to make any educated assessment.

Statistical measures are often used for describing the nature and extent of differences among the information in the distribution. A measure of variability is generally reported together with a measure of central tendency.
Remember, quality of information and variation is inversely related. The larger the variation in the data, the lower the quality of the data. The four most common measures of variation are the range, variance, standard deviation, and coefficient of variation.

Range: The range of a set of observations is the absolute value of the difference between the largest and smallest values in the data set. It is not useful when extreme values are present. It is based solely on two values, not on the entire data set. In addition, it cannot be defined for open-ended distributions such as Normal distribution.

Notice that, when dealing with discrete random observations, some authors define the range as: Range = Largest value - Smallest value + 1.

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

Quartiles: When we order the data, for example in ascending order, we may divide the data into quarters, known as quartiles. The first Quartile (Q1) is that value where 25% of the values are smaller and 75% are larger. The second Quartile (Q2) is that value where 50% of the values are smaller and 50% are larger. The third Quartile (Q3) is that value where 75% of the values are smaller and 25% are larger.

Percentiles: Percentiles have a similar concept and therefore, are related; e.g., the 25th percentile corresponds to the first quartile Q1, etc. The advantage of percentiles is that they may be subdivided into 100 parts. The percentiles and quartiles are most conveniently read from a cumulative distribution function.

Interquartiles Range: The interquartile range (IQR) describes the extent for which the middle 50% of the observations scattered or dispersed. It is the distance between the first and the third quartiles:

IQR = Q3 - Q1,

which is twice the Quartile Deviation. For data that are skewed, the relative dispersion, similar to the coefficient of variation (C.V.) is given (provided the denominator is not zero) by the Coefficient of Quartile Variation:

CQV = (Q3-Q1) / (Q3 + Q1).

Variance: An important measure of variability is variance. Variance is the average of the squared deviations of each observation in the set from the arithmetic mean of all of the observations.

Population variance = $$\sigma^2 = \frac{\sum(X-\mu)^2}{N}$$

Sample variance = $$s^2 = \frac{\Sigma(x-\overline{x})^2}{N-1}$$

The variance is a measure of spread or dispersion among values in a data set. Therefore, the greater the variance, the lower the quality.

The variance is not expressed in the same units as the observations.

Standard Deviation: Both variance and standard deviation provide the same information; one can always be obtained from the other. In other words, the process of computing a standard deviation always involves computing a variance. Since standard deviation is the square root of the variance, it is always expressed in the same units as the raw data:

Standard Deviation = S = (Variance) Â½

For large data sets (say, more than 30), approximately 68% of the data are contained within one standard deviation of the mean, 95% contained within two standard deviations. 97.7% (or almost 100%) of the data are contained within within three standard deviations (S) from the mean.

VI. HYPOTHESIS TESTING

A hypothesis, in statistics, is a statement about a population where this statement typically is represented by some specific numerical value. In testing a hypothesis, we use a method where we gather data in an effort to gather evidence about the hypothesis. In hypothesis testing there are certain steps one must follow.

1. **Setting up two competing hypotheses** - Each hypothesis test includes two hypothesis about the population. One is the null hypothesis, notated as H_o, which is a statement of a particular parameter value. This hypothesis is assumed to be true until there is evidence to suggest otherwise. The second hypothesis is called the alternative, or research, hypothesis, notated as H_a. The alternative hypothesis is a statement of a range of alternative values in which the parameter may fall.
2. **Set some level of significance called alpha.** This value is used as a probability cutoff for making decisions about the null hypothesis. The most common alpha value is 0.05 or 5%. Other popular choices are 0.01 (1%) and 0.1 (10%).
3. **Calculate a test statistic.** Gather sample data and calculate a test statistic where the sample statistic is compared to the parameter value. The test statistic is calculated under the assumption the null hypothesis is true, and incorporates a measure of standard error and assumptions (conditions) related to

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

the sampling distribution. Such assumptions could be normality of data, independence, and number of success and failure outcomes.

4. **Calculate probability value (p-value), or find rejection region** - A p-value is found by using the test statistic to calculate the probability of the sample data producing such a test statistic or one more extreme. The rejection region is found by using alpha to find a critical value; the rejection region is the area that is more extreme than the critical value.
5. **Make a test decision about the null hypothesis** - In this step we decide to either reject the null hypothesis or decide to fail to reject the null hypothesis. Notice we do not make a decision where we will accept the null hypothesis.
6. **State an overall conclusion** - Once we have found the p-value or rejection region, and made a statistical decision about the null hypothesis (i.e. we will reject the null or fail to reject the null). Following this decision, we want to summarize our results into an overall conclusion for our test.

VII. LARGE SAMPLE TESTS FOR A POPULATION MEAN

Assumptions

1. Sample is randomly selected
2. Sample is large (n > 30) Central Limit theorem applies
3. If σ is unknown, we can use sample standard deviation s as estimate for σ .

Goal

Identify a sample result that is significantly different from the claimed value; in this case, is our sample mean statistically different from the claimed null hypothesis mean?

Steps

1. Identify the null hypothesis (specific claim to be tested) $H_0 : \mu = \mu_0$
2. Identify the alternative hypothesis that must be true when the original claim is false.

One-tailed test two -tailed test

Ha: μ or Ha: $\mu \neq \mu_0$

3. Calculate the test statistic:

$$Z = \frac{\overline{x} - \mu_0}{\sigma / \sqrt{n}}$$

1. Select the significant level α based on the seriousness of a type I error. The values of 0.05 and 0.01 are very common.
2. Reject H0 if the test statistic is in the critical region. Fail to reject H0 if the test statistic is not in the critical region.

Hypothesis Testing Of The Difference Between Two Population Means

This is a two sample z test which is used to determine if two population means are equal or unequal. There are three possibilities for formulating hypotheses.

$H_0 = \mu_1 = \mu_2$ $H_a : \mu_1 \neq \mu_2$
$H_0 :: \mu_1 \geq \mu_2$ $H_a : \mu_1 < \mu_2$
$H_0 :: \mu_1 \leq \mu_2$ $H_a : \mu_1 > \mu_2$

The same procedure is used in three different situations

- Sampling is from normally distributed populations with known variances

$$z = \frac{(\overline{x}_1 - \overline{x}_2) - (\mu_1 - \mu_2)_0}{\sqrt{\sigma_1^2 / n_1 + \sigma_2^2 / n_2}}$$

Tests with One Sample, Dichotomous Outcome

Hypothesis testing applications with a dichotomous outcome variable in a single population are also performed according to the five-step procedure. Similar to tests for means, a key component is setting up the null and

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

research hypotheses. The objective is to compare the proportion of successes in a single population to a known proportion (p_0).

In one sample tests for a dichotomous outcome, we set up our hypotheses against an appropriate comparator. We select a sample and compute descriptive statistics on the sample data.

$$\hat{p} = \frac{x}{n}$$

We then determine the appropriate test statistic (Step 2) for the hypothesis test. The formula for the test statistic is given below.

Test Statistic for Testing H_0: $p = p_0$

if $\min(np_0, n(1-p_0)) \geq 5$

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

The formula above is appropriate for large samples, defined when the smaller of np_0 and $n(1-p_0)$ is at least 5. This is similar, but not identical, to the condition required for appropriate use of the confidence interval formula for a population proportion, i.e.,

$$\min\left(n\hat{p},\, n\left(1-\hat{p}\right)\right) \geq 5$$

Here we use the proportion specified in the null hypothesis as the true proportion of successes rather than the sample proportion. If we fail to satisfy the condition, then alternative procedures, called exact methods must be used to test the hypothesis about the population proportion

Tests with Two Independent Samples, Continuous Outcome

There are many applications where it is of interest to compare two independent groups with respect to their mean scores on a continuous outcome. Here we compare means between groups, but rather than generating an estimate of the difference, we will test whether the observed difference (increase, decrease or difference) is statistically significant or not. Remember, that hypothesis testing gives an assessment of statistical significance, whereas estimation gives an estimate of effect and both are important.

for sample 1:

- n1 $\qquad \bar{X}_1 \qquad$ s1

for sample 2:

- n2 $\qquad \bar{X}_2 \qquad$ s2

In the two independent samples s1and s2 application with a continuous outcome, the parameter of interest in the test of hypothesis is the difference in population means, μ_1-μ_2. The null hypothesis is always that there is no difference between groups with respect to means, i.e.,

$$H_0: \mu_1 - \mu_2 = 0$$

The null hypothesis can also be written as follows: H_0: $\mu_1 = \mu_2$. In the research hypothesis, an investigator can hypothesize that the first mean is larger than the second (H_1: $\mu_1 > \mu_2$), that the first mean is smaller than the second (H_1: $\mu_1 < \mu_2$), or that the means are different (H_1: $\mu_1 \neq \mu_2$). The three different alternatives represent upper-, lower-, and two-tailed tests, respectively. The following test statistics are used to test these hypotheses.

Test Statistics for Testing H_0: $\mu_1 = \mu_2$

- if $n_1 \geq 30$ and $n_2 \geq 30$

$$z = \frac{\bar{X}_1 - \bar{X}_2}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

- if $n_1 < 30$ or $n_2 < 30$

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

$$t=\frac{\bar{X}_1-\bar{X}_2}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$ where df $=n_1+n_2-2$.

NOTE: The formulas above assume equal variability in the two populations (i.e., the population variances are equal, or $s_1^2 = s_2^2$). This means that the outcome is equally variable in each of the comparison populations. The test statistics include Sp, which is the pooled estimate of the common standard deviation (again assuming that the variances in the populations are similar) computed as the weighted average of the standard deviations in the samples as follows:

$$S_p=\sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}}$$

Because we are assuming equal variances between groups, we pool the information on variability (sample variances) to generate an estimate of the variability in the population.

Tests with Matched Samples, Continuous Outcome

The two comparison groups are said to be ***dependent,*** and the data can arise from a single sample of participants where each participant is measured twice (possibly before and after an intervention) or from two samples that are matched on specific characteristics (e.g., siblings). When the samples are dependent, we focus on ***difference scores*** in each participant or between members of a pair and the test of hypothesis is based on the mean difference, μ_d. The null hypothesis again reflects "no difference" and is stated as H_0: $\mu_d = 0$. Note that there are some instances where it is of interest to test whether there is a difference of a particular magnitude (e.g., $\mu_d = 5$) but in most instances the null hypothesis reflects no difference (i.e., $\mu_d=0$).

The appropriate formula for the test of hypothesis depends on the sample size.

Test Statistics for Testing H_0: $\mu_d = 0$

- if $n \geq 30$

$$z=\frac{\bar{X}_d-\mu_d}{s_d/\sqrt{n}}$$

- if $n < 30$

$$t=\frac{\bar{X}_d-\mu_d}{s_d/\sqrt{n}}$$ where df $=n-1$

Tests with Two Independent Samples, Dichotomous Outcome

Here we consider the situation where there are two independent comparison groups and the outcome of interest is dichotomous (e.g., success/failure). The goal of the analysis is to compare proportions of successes between the two groups. The relevant sample data are the sample sizes in each comparison group (n_1 and n_2) and the sample proportions ($\hat{p}_1$ **and** $\hat{p}_2$) which are computed by taking the ratios of the numbers of successes to the sample sizes in each group, i.e.,

$$\hat{p}_1=\frac{x_1}{n_1} \text{ and } \hat{p}_2=\frac{x_2}{n_2}$$

In tests of hypothesis comparing proportions between two independent groups, one test is performed and results can be interpreted to apply to a risk difference, relative risk or odds ratio.

Test Statistics for Testing H_0: p $_1$ = p

$$z=\frac{\hat{p}_1-\hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

Where $\hat{p}_1$ the proportion of successes in sample 1 is, $\hat{p}_2$is the proportion of successes in sample 2, and $\hat{p}$is the proportion of successes in the pooled sample. $\hat{p}$is computed by summing all of the successes and dividing by the total sample size, as follows:

$$p = \frac{x_1 + x_2}{n_1 + n_2}$$

(this is similar to the pooled estimate of the standard deviation, Sp, used in two independent samples tests with a continuous outcome; just as Sp is in between s_1 and s_2, $\hat{p}$will be in between $\hat{p}_1$ and $\hat{p}_2$).

The formula above is appropriate for large samples, defined as at least 5 successes ($np \geq 5$) and at least 5 failures ($n(1-p \geq 5)$) in each of the two samples. If there are fewer than 5 successes or failures in either comparison group, then alternative procedures, called exact methods must be used to estimate the difference in population proportions.

The Chi Square Statistic

A chi square (X^2) statistic is used to investigate whether distributions of categorical variables differ from one another. Basically categorical variable yield data in the categories and numerical variables yield data in numerical form.

The chi-square test is a statistical test that can be used to determine whether observed frequencies are significantly different from expected frequencies.

Chi-square is used most commonly to compare the incidence (or proportion) of a characteristic in one group to the incidence (or proportion) of a characteristic in other group(s).

$$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e}$$

where f_o = the observed frequency (the observed counts in the cells) and f_e = the expected frequency if NO relationship existed between the variables

VIII. ANOVA

Analysis of variance (ANOVA) tests the hypothesis that the means of two or more populations are equal. ANOVAs assess the importance of one or more factors by comparing the response variable means at the different factor levels. The null hypothesis states that all population means (factor level means) are equal while the alternative hypothesis states that at least one is different.

To perform an ANOVA, you must have a continuous response variable and at least one categorical factor with two or more levels. ANOVAs require data from approximately normally distributed populations with equal variances between factor levels

ANOVA type	Model and design properties
One-way	One fixed factor (levels set by investigator) which can have either an unequal (unbalanced) or equal (balanced) number of observations per treatment.
Balanced	Model may contain any number of fixed and random factors (levels are randomly selected), and crossed and nested factors, but requires a balanced design.
General linear model	Expands on Balanced ANOVAs by allowing unbalanced designs and covariates (continuous variables).

The hypotheses of interest in an ANOVA are as follows:

- H_0: $\mu_1 = \mu_2 = \mu_3 \ldots = \mu_k$
- H_1: Means are not all equal.

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

where k = the number of independent comparison groups.

Test Statistic for ANOVA

The test statistic for testing H_0: $\mu_1 = \mu_2 = \ldots = \mu_k$ is:

$$F = \frac{\Sigma n_j(\bar{X}_j - \bar{X})^2 / (k-1)}{\Sigma\Sigma(X - \bar{X}_j)^2 / (N-k)}$$

and the critical value is found in a table of probability values for the F distribution with (degrees of freedom) $df_1 = k-1$, $df_2=N-k$. The table can be found in "Other Resources" on the left side of the pages.

In the test statistic, n_j = the sample size in the j^{th} group (e.g., j =1, 2, 3, and 4 when there are 4 comparison groups), $\bar{X}_j$ is the sample mean in the j^{th} group, and $\bar{X}$ is the overall mean. k represents the number of independent groups (in this example, k=4), and N represents the total number of observations in the analysis.

NOTE: The test statistic F assumes equal variability in the k populations (i.e., the population variances are equal, or $s_1^2 = s_2^2 = \ldots = s_k^2$). This means that the outcome is equally variable in each of the comparison populations. This assumption is the same as that assumed for appropriate use of the test statistic to test equality of two independent means. It is possible to assess the likelihood that the assumption of equal variances is true and the test can be conducted in most statistical computing packages. If the variability in the k comparison groups is not similar, then alternative techniques must be used.

The decision rule for the F test in ANOVA is set up in a similar way to decision rules we established for t tests. The decision rule again depends on the level of significance and the degrees of freedom. The F statistic has two degrees of freedom. These are denoted df_1 and df_2, and called the numerator and denominator degrees of freedom, respectively. The degrees of freedom are defined as follows:

$$df_1 = k\text{-}1 \text{ and } df_2=N\text{-}k,$$

where k is the number of comparison groups and N is the total number of observations in the analysis. If the null hypothesis is true, the between treatment variation (numerator) will not exceed the residual or error variation (denominator) and the F statistic will small. If the null hypothesis is false, then the F statistic will be large.

The ANOVA Procedure

Because the computation of the test statistic is involved, the computations are often organized in an ANOVA table. The ANOVA table breaks down the components of variation in the data into variation between treatments and error or residual variation. Statistical computing packages also produce ANOVA tables as part of their standard output for ANOVA, and the ANOVA table is set up as follows:

Source of Variation	Sums of Squares (SS)	Degrees of Freedom (df)	Mean Squares (MS)	F
Between Treatments	$SSB = \Sigma n_j(\bar{X}_j - \bar{X})^2$	k-1	$MSB = \frac{SSB}{k-1}$	$F = \frac{MSB}{MSE}$
Error (or Residual)	$SSE = \Sigma\Sigma(X - \bar{X}_j)^2$	N-k	$MSE = \frac{MSE}{N-k}$	
Total	$SST = \Sigma\Sigma(X - \bar{X})^2$	N-1		

where

- X = individual observation,
- $\bar{X}_j$ = sample mean of the j^{th} treatment (or group),

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

- $\overline{X}$ = overall sample mean,
- k = the number of treatments or independent comparison groups, and
- N = total number of observations or total sample size.

IX. CONCLUSION

In competitive environment, business managers must design quality into products, and into the processes of making the products. They must facilitate a process of never-ending improvement at all stages of manufacturing and service. This is a strategy that employs statistical methods, particularly statistically designed experiments, and produces processes that provide high yield and products that seldom fail. Moreover, it facilitates development of robust products that are insensitive to changes in the environment and internal component variation. Carefully planned statistical studies remove hindrances to high quality and productivity at every stage of production. This saves time and money. It is well recognized that quality must be engineered into products as early as possible in the design process. One must know how to use carefully planned, cost-effective statistical experiments to improve, optimize and make robust products and processes..

X. REFERENCES

[1] Box, G., Tiao, G., 1992, Bayesian Inference in Statistical Analysis. John Wiley & Sons Bogacz, R., 2007 , Optimal decision making theories: Linking neurobiology with behavior, Trends in Cognitive Sciences , 11, 118 – 125.
[2] Cooper, D.R., & Emory, C.W. (1995). Business research method (5th edn). Chicago: Irwin.
[3] Dixit, A. K., Pindyck, R. S., 1994, Investment under Uncertainty. Princeton: Princeton University Press.
[4] Hannagan, T. (1995). Management concepts and practices. San Francisco: Pitman.
[5] Hargreaves, C.A. (1999). Mechanisms to assist in decision making. Unpublished MBA
[6] Dissertation, University of Wales.
[7] Huyse L.T hacker B.H., 2003 , Probabilistic assessment on the basis of interval data, In Der Kiureghian, Madanat and Pestana (Eds.), Applications of Statistics and Probability in Civil Engineering, Proceedings of ICASP9 , Mill press, Rotterdam, the Netherland s, 169-176.
[8] Juritz, J., Money, A.H., Affleck-Graves, J., & DuToit, P. (1989). A Survey of Statistical Methods used in business and industry in South Africa. South African Journal of Business Management, 20.
[9] Moolman, W.H. (1996). The involvement of statisticians in the application of statistical methods in companies - a study of the Durban-Pinetown region . Paper presented at the 1996 South African Statistical Association Conference.
[10] Schwartz, S. J., Sturr, M., & Goldberg, G. (1996) . Statistical methods in rehabilitation literature: a survey of recent publications. Archives of Physical Medicine and Rehabilitation , 77 , 497-500.

CITE AN ARTICLE

Sharme, Ankur, and Payal. "APPLICATIONS OF STATISTICAL TOOLS FOR DECISION MAKING UNDER UNCERTAINTY IN MANUFACTURING INDUSTRIES." *INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY* 6.7 (2017): 403-12. Web. 15 July 2017.