College of Natural Sciences

Department of Statistics

Course title*:* Statistical Quality Control Lecture Note

Course code: Stat3142

Credit hours: 3

Instructor's Name: Reta Habtamu Bacha

Email: *habtamuretab@yahoo.com*

April: 2020/2021

Chapter 1

1 Introduction to Statistical Quality Control

1.1 Introduction

Quality can be defined in many ways, ranging from "satisfying customers' requirements" to "fitness for use" to "conformance to requirements." It is obvious that any definition of quality should include customers, satisfying who must be the primary goal of any business.

Statistical quality control abbreviated as "SQC" is one of the most important application of the statistical techniques in industry. These technique is based on the theory of probability and sampling and are being extensively used in all most all industries of our country as well as other world , such as air craft, armament, automobile, textiles , plastic, rubber, petroleum, electrical equipments, telephone, transportation, chemical and medicine and so on. In fact it is impossible to think any of industry field were statistical techniques are not used.

The basic problems in any production process are not the quantum of the product but the quality of the product. The product is basically interested to see that the product is of acceptable quality. i.e. it conforms to certain prescribed standards or specification the quality of the manufactured product depends on the number of factors, starting with its design and specification the production process, the raw materials, machines and equipments, the expertise and skill of the person who handle them, and inspection of the final product. For successful marketing of the product, it is imperative that the end product must confirm to the standards or expectation lay down by the customers. Statistical quality control aims at achieving this target by keeping the various steps of the process (from arrival of the material) through each of their processing to the final delivery of goods with in control. The principle of quality control by statistical techniques covers almost all aspects of production, i.e. quality of materials, quality of man power, quality of machines and quality of management.

The purpose of statistical quality control is to ensure, in a cost efficient manner, that the product shipped to customers meets their specifications. Inspecting every product is costly and inefficient, but the consequences of shipping non conforming product can be significant in terms of customer dissatisfaction. Statistical Quality Control is the process of inspecting enough product from given lots to probabilistically ensure a specified quality level. Statistical quality control provides the statistical techniques necessary to assure and improve the quality of products.

1.2 Definition of Statistical Quality Control (SQC)

Statistical quality control or SQC are also called statistical quality assurance (SQA) refers to the area of statistics concerned with applying statistical methods to improving the quality of products.

Quality control emphasizes testing of products to uncover defects, and reporting to management who make the decision to allow or deny the release. Whereas quality assurance attempts to improve and stabilize production, and associated processes, to avoid, or at least minimize, issues that led to the defects in the first place.

To prevent mistakes from arising, several QA methodologies are used. However, QA does not necessarily eliminate the need for QC: some product parameters are so critical that testing is still necessary. QC activities are treated as an integral part of the overall QA processes.

Quality assurance, or **QA** for short, refers to a program for the systematic monitoring and evaluation of the various aspects of a project, service, or facility to ensure that standards of quality are being met.

It is important to realize also that quality is determined by the program sponsor. QA cannot absolutely guarantee the production of *quality* products, unfortunately, but makes this more likely.

Two key principles characterized QA: "fit for purpose" (the product should be suitable for the intended purpose) and "right first time" (mistakes should be eliminated). QA includes regulation of the quality of raw materials, assemblies, products and components; services related to production; and management, production and inspection processes.

It is important to realize also that *quality* is determined by the intended users, clients or customers, not by society in general: it is not the same as 'expensive' or 'high quality'. Even goods with low prices can be considered quality items if they meet a market need.

What meant by a quality of a product?

Every article or product is required for specific purpose. If it is fully serves purpose, it is of good quality otherwise was not. It means that if an article or materials meets the specification required for its rightful use, it is good quality, and if not then the quality of article is considered to be poor. Statistical quality control plays a great role in maintaining the quality of product. There is hardly any control on the quality of products produced by the nature and hence the statistical quality control remained confined to articles produced by the industry. Variations in items of produced in any manner are inevitable. These variations occur due to two types of causes namely (i) chance factors and (ii) assignable cause.

Chance and Assignable cause of variation

No production process is good enough to produce all items exactly alike in spite of sophisticated technology and precession of modern scientific equipments; no two articles produced by the same machine are perfectly identical in measurable characteristics. Some amount of variation, through of an infinitesimal nature, is inherit and inevitable in every repetitive process in industry. These variation are broadly termed as resulting from

- Chance cause
- Assignable cause

Chance cause: - some stable patterns of variation or constant cause system are inherent in every scheme of production and inspection. The variation due to these chance causes can be regard as independent normal variation zero means and same error variance says σ^2.

These causes are natural to any production process and the variation due to them is known as natural variation or chance of variation. The natural variation is beyond the control of human hand and can not detected or prevented. Natural variation also some times known as allowable variation as it can not be eliminated and one has to allow such variation in process. Some deviation from the desired specification is bound to occur in the articles produced, howsoever efficient, the production process maybe. If the variation occur due to some inherit pattern of variation and no cause can be assign to it, it is called chance or random variation. Chance variation is tolerable and does not martially affect the quality of product. In such a situation, the process is said to be under statistical quality control.

Assignable cause: - an assignable cause system is one in which the cause of variation are not distributed normally but not sporadically. Assignable cause also known as non – random cause and the variation due to these cause is termed as chaotic or erratic or preventive variation. Some of the important factors of assignable cause of variation are substandard or defective raw materials, new techniques or operations, negligence of operation, wrong or improper handling of the machines , faulty equipment, unskilled or in experienced technical staff, and so on. These cause can be identified and eliminated and have to be traced in any production process before the production becomes defective. If the article show marked deviation from the given specification of product, the utility of articles is jeopardy. In that situation, one has to make a search for the cause of responsible for the large variation in the product. The cause due to faulty process and procedure are known as assignable causes. The variation due to assignable causes is of non-random nature. Hence, the role of statistical quality control is to collect and analyses relevant data for the purpose of detecting weather the process is under control on not. If not, what can possibly be the reason for faults?

Example 1. A MOHA company produced bottles of various soft drinks. Bottles of Pepsi are supposed to have 300ml of drink. Bottles are filled by a machine so some will have more than 300ml, some less if the machine is set incorrectly then the bottle will filled too much the company loose money or the bottle will filled too little customer complain. A

company will take a sample of bottles. If the mean volume is significantly different from 300ml it includes some thing wrong in the process.

Example 2. An electrical company makes light bulbs that are suppose to have an average life time of 1000 hours. A buyer intends to buy a batch of 10,000 bulbs. The buyer first test a sample of 100 light bulbs if they have a mean life time of $\geq$ 1000 hours he accepts the batch other wise the batch is rejected.

Example 3. A company makes cups and plates. Some of cups might be chipped or cracked. A sample will be taken to cheek the proportion of cracked cups. If this is above a certain value the process must be adjusted. In all above three examples

- ✓ There is a cost involving in sampling.
- ✓ There is a decision to be made that can result an error. For these case wrongly rejects a good products it results supplier or producer risk as a result of good products are rejected. Where as wrongly accepted a bad products a consumers or buyers at risk at a result of bad or defective products are distributed to in a certain market. the main aims of statistical quality control (SQC) is that to develop and control statistical procedures that have

- Low cost i.e. the ability to minimize cost.
- Low producer risk. By taking remedial action and by controlling the production process, it reduces supplier or generally company risk.
- Low buyers risk. This implies by checking and controlling the production process it minimize or reduces the consumer risk, by taking a carefully designed hypothesis testing.

1.3 History of Statistical Quality Control

Quality Control has been with us for a long time. How long? It is safe to say that when manufacturing began and competition accompanied manufacturing, consumers would compare and choose the most attractive product. If manufacturer A discovered that manufacturer B's profits soared, the former tried to improve his/her offerings, probably by improving the quality of the output, and/or lowering the price. Improvement of quality

did not necessarily stop with the product - but also included the *process* used for making the product.

The process was held in high esteem, as manifested by the medieval guilds of the Middle Ages. These guilds mandated long periods of training for apprentices, and those who were aiming to become master craftsmen had to demonstrate evidence of their ability. Such procedures were, in general, aimed at the maintenance and improvement of the quality of the process.

In modern times we have professional societies, governmental regulatory bodies such as the Food and Drug Administration, factory inspection, etc., aimed at assuring the quality of products sold to consumers. *Quality Control* has thus had a long history.

Science of statistics in fairly recent

On the other hand, *statistical* quality control is comparatively new. The science of statistics itself goes back only two to three centuries. And its greatest developments have taken place during the 20th century. The earlier applications were made in astronomy and physics and in the biological and social sciences. It was not until the 1920s that statistical theory began to be applied effectively to quality control as a result of the development of sampling theory

The concept of quality control in manufacturing was first advanced by Walter Shewhart

The first to apply the newly discovered statistical methods to the problem of quality control was Walter A. Shewhart of the Bell Telephone Laboratories. He issued a memorandum on May 16, 1924 that featured a sketch of a modern control chart.

Shewhart kept improving and working on this scheme, and in 1931 he published a book on statistical quality control, *"Economic Control of Quality of*

Manufactured Product", published by Van Nostrand in New York. This book set the tone for subsequent applications of statistical methods to process control.

Contributions of Dodge and Roming to sampling inspections

Two other Bell Labs statisticians, H.F. Dodge and H.G. Romig spearheaded efforts in applying statistical theory to sampling inspection. The work of these three pioneers constitutes much of what nowadays comprises the theory of statistical quality and control. There is much more to say about the history of statistical quality control and the interested reader is invited to peruse one or more of the references.

Classical techniques of hypothesis testing, or function of type I error and type II error etc. from the bases of acceptance sampling. This theory developed in the early 1900, by Neymon- Pearson, Fischer's and others. In 1920s Shewhart introduces the control charts, this controls the quality of a product as it is being made rather than after it is made. This case is key components of statistical process control (SPC). The growth of statistical process control coincides with the growth of factors as away of manipulating goods. In 1940s because of a war there were a lot of researchers and researches in to statistical process control would introduce sequential sampling. In non –sequential sampling we have a fixed sample size. **Example** we might decide to examine 100 items and rejected the batch if ≥ 10 items are bad this can be wistful. i.e. if we have sampled 20 items and found 8 are bad. It is clear that the batch is bad we still sample another 80 items. Similarly if we have examine 70 items all are bad includes it is clear that the batch is good. But we still need sample to another 70 items. This implies this process is inefficient. But sequential sampling is can be more efficient. In the 1950s, Japan's industry has been destroyed by wars, they invited American statistician, "Deming" to introduce statistical methods to industry. Deming introduced "14 point" of good management and partiality. As a result Japan's industry has grown. Now statistical quality control and Deming methods are used in many industries's world wide.

- ***Example for sequential sampling***: - take a sample of 100 items if > 20 items are bad or defective we will reject the batch. This implies non–sequential sampling; in non–sequential sampling we need try to sample a fixed number of items. In sequential sampling after we have sampled the "i^{th}" item there are three possibilities.
- *Accept the batch*
- *Reject the batch*
- *Sample another item.*

1.4 Uses of Statistical Quality Control

Statistical quality control is a very important technique which is used to assess the cause of variations in the quality of manufactured product. It enables us to determine weather the quality standards are being met with out inspecting every unit produced in the process. It primary aims at the isolation of the chance and assignable cause of variation and consequently helps in the detection, identification and elimination of the assignable causes of erratic fluctuations when ever they are present. A production process is said to be in a state of statistical control if it is operating in the presence of chance causes only and is free from assignable cause of variation.

The following are some of the principal uses when a manufacturing process is operating in state of statistical quality control.

1. An obvious uses of statistical quality control is the control, maintenance, and improvement in the quality standard.
2. The technique of brining a process in good statistical control involves the detection and removal of assignable cause of variation. This ultimately results in the diagnosis and elimination of many production troubles, and in the improvement quality.
3. Before the development of statistical quality control techniques, the methods of controlling quality were confined only to finished products. But the statistical quality control techniques based on the probability and sampling techniques enable us to predict the quality of manufactured product. A process in good statistical control is predictable. It implies that there are no apparent chaotic or assignable causes and it is believe that it will remain in provided the conditions of manufacture remain same.

Hence, if a process is working in a state of statistical control, we can more safely guarantee the product. This gives better quality assurance and hence more satisfaction to the customer.

4. Statistical quality control techniques enable us to decide when to take remedial action and when to leave the process alone. If the process is in statistical control, we know that it is going to do and hence we can safely dispense with frequent and unwarranted inspection of the manufactured units and adjustment in the machines. If the product is not standard quality, i.e. the product does not conform to specifications, in spite of the fact that the process is in statistical control, then some radical change in the production process, raw material, machines, equipment, etc... has to be done.

   The diagnosis of assignable causes of variation gives us an early and timely warning about the occurrence of defects. this is ultimately help in

   - ✓ Reduction in waste and scrap.
   - ✓ Reduction in cost per unit
   - ✓ Reduction in inspection

All these factors result in less cost of production, enhanced productivity and more profits for the manufacturing concerns.

5. The mere presence of a statistical quality control scheme in any manufacturing concern has a very healthy effect as it creates quality consciousness among their personnel. Such a scheme keeps the staff and the workers on their alert theory by increasing their efficiency.

1.4 Quality Improvement in Modern Business Environment

Invariably, the Quality of output is directly dependent upon that of the participating constituents. Some of which are sustainable and effectively controlled while others are [not]..

The major problem which leads to a decrease in sales was that the specifications did not include the most important factor, “What the specifications have to state in order to satisfy the customer requirements?”

The major characteristics, ignored during the search to improve manufacture and overall business performance were:

- Reliability
- Maintainability
- Safety
- Strength.

As the most important factor had been ignored, a few refinements had to be introduced:

1. Marketing had to carry out their work properly and define the customer's specifications.
2. Specifications had to be defined to conform to these requirements.
3. Conformance to specifications i.e. drawings, standards and other relevant documents, were introduced during manufacturing, planning and control.
4. Management had to confirm all operators are equal to the work imposed on them and holidays, celebrations and disputes did not affect any of the quality levels.
5. Inspections and tests were carried out, and all components and materials, bought in or otherwise, conformed to the specifications, and the measuring equipment was accurate, this is the responsibility of the QA/QC department.
6. Any complaints received from the customers were satisfactorily dealt with in a timely manner.
7. Feedback from the user/customer is used to review designs.
8. Consistent data recording and assessment and documentation integrity.
9. Product and/or process change management and notification.

If the specification does not reflect the true quality requirements, the product's quality cannot be guaranteed. For instance, the parameters for a pressure vessel should cover not only the material and dimensions but operating, environmental, safety, reliability and maintainability requirements.

During the 1980s, the concept of “company quality” with the focus on management and people came to the fore. It was realized that, if all departments approached quality with an open mind, success was possible if the management led the quality improvement process.

The company-wide quality approach places an emphasis on four aspects:-

1. Elements such as controls, job management, adequate processes, performance and integrity criteria and identification of records
2. Competence such as knowledge, skills, experience, qualifications
3. Soft elements, such as personnel integrity, confidence, organizational culture, motivation, team spirit and quality relationships.
4. Infrastructure (as it enhances or limits functionality)

The quality of the outputs is at risk if any of these aspects is deficient in any way.

The approach to quality management given here is therefore not limited to the manufacturing theatre only but can be applied to any business or non-business activity:

- Design work
- Administrative services
- Consulting
- Banking
- Insurance
- Computer software development
- Retailing
- Transportation
- Education

It comprises a quality improvement process, which is generic in the sense it can be applied to any of these activities and it establishes a behaviors pattern, which supports the achievement of quality.

This in turn is supported by quality management practices which can include a number of business systems and which are usually specific to the activities of the business unit concerned.

It has become customary to use consultants and contractors when introducing new quality practices and methods, particularly where the relevant skills and expertise are not available within the organization. In addition, when new initiatives and improvements are required to boost the current quality system, or perhaps improve upon current manufacturing systems, the use of temporary consultants becomes a viable solution when allocating valuable resources.

- There are various types of consultants and contractors available in the market; most will have the skills needed to facilitate improvement activities such as Quality Management Systems (QMS) auditing and procedural documentation writing. More experienced consultants are likely to have knowledge of specialized quality improvement activities such as Six Sigma, Measurement Systems Analysis (MSA), and Quality Function.

Good manufacturing practice

"**Good manufacturing practice**" or "**GMP**" refers to the quality control of manufacturing for foods, pharmaceutical products, and medical devices. GMPs are guidelines that outline the aspects of production that would affect the quality of a product. Many countries have legislated that pharmaceutical and medical device companies must follow GMP procedures, and have created their own GMP guidelines that correspond with their legislation.

Other good practices

Other good-practice systems, along the same lines as GMP, exist:

- Good laboratory practice (GLP), for laboratories conducting non-clinical studies (toxicology and pharmacology studies in animals);

- Good clinical practice (GCP), for hospitals and clinicians conducting clinical studies on new drugs in humans;
- Good regulatory practice (GRP), for the management of regulatory commitments, procedures and documentation.

Collectively, these and other good-practice requirements are referred to as "GxP" requirements, all of which follow similar philosophies.

Quality Management

The role of a quality manager is one that is constantly evolving. With increased demands on manufacturing operations quality managers are finding themselves subject to stricter guidelines from both their customers and regulatory bodies. Quality managers are now required to take proactive approaches to ensure they are meeting the variety of requirements that are imposed on them. Historically individual programs were developed at the plant level to address quality needs; however, quality managers are now taking an approach that includes a complete enterprise wide view of the business.

IQS (International Quality Standard) provides the following capabilities to quality managers:

- Document operating procedures, policies, and work instructions
- Manage nonconformance's and corrective actions
- Manage and investigate customer complaints
- Manage supplier related issues including supplier performance
- Implement closed-loop processes that support continuous improvement efforts .

1.6 Modeling Process Quality

A **process** is a unique combination of tools, materials, methods, and people engaged in producing a measurable output; for example a manufacturing line for machine parts. All processes have inherent statistical variability which can be evaluated by statistical methods.

The **Process Capability** is a measurable property of a process to the specification, expressed as a process capability index (e.g., C_{pk} or C_{pm}) or as a process performance index (e.g., P_{pk} or P_{pm}). The output of this measurement is usually illustrated by a histogram and calculations that predict how many parts will be produced out of specification.

Process capability is also defined as the capability of a process to meet its purpose as managed by an organization's management and process definition structures.

Two parts of process capability are: 1) Measure the variability of the output of a process, and 2) Compare that variability with a proposed specification or product tolerance.

Measure the Process

The input of a process usually has at least one or more measurable characteristics that are used to specify outputs. These can be analyzed statistically; where the output data shows a normal distribution the process can be described by the process mean (average) and the standard deviation.

A process needs to be established with appropriate process controls in place. A control chart analysis is used to determine whether the process is "in statistical control". If the process is not in statistical control then capability has no meaning. Therefore the process capability involves only common cause variation and not special cause variation.

A batch of data needs to be obtained from the measured output of the process. The more data that is included the more precise the result, however an estimate can be achieved with as few as 17 data points. This should include the normal variety of production conditions, materials, and people in the process. With a manufactured product, it is common to include at least three different production runs, including start-ups.

The process mean (average) and standard deviation are calculated. With a normal distribution, the "tails" can extend well beyond plus and minus three standard deviations, but this interval should contain about 99.73% of production output. Therefore for a

normal distribution of data the process capability is often described as the relationship between six standard deviations and the required specification.

Capability study

The input of a process is expected to meet customer requirements, specifications, or product tolerances. Engineering can conduct a process capability study to determine the extent to which the process can meet these expectations.

The ability of a process to meet specifications can be expressed as a single number using a process capability index or it can be assessed using control charts. Either case requires running the process to obtain enough measurable output so that engineering is confident that the process is stable and so that the process mean and variability can be reliably estimated. Statistical process control defines techniques to properly differentiate between stable processes, processes that are drifting (experiencing a long-term change in the mean of the output), and processes that are growing more variable. Process capability indices are only meaningful for processes that are stable (in a state of statistical control).

Chapter Two

2.Methods of Statistical Process Control and Capability Analysis

*2.*1 Introduction

Statistical process control (SPC) is an effective method of monitoring a process through the use of control charts. Control charts enable the use of objective criteria for distinguishing background variation from events of significance based on statistical techniques. Much of its power lies in the ability to monitor both process center and its variation about that center, by collecting data from samples at various points within the process. Variations in the process that may affect the quality of the end product or service can be detected and corrected, thus reducing waste as well as the likelihood that problems will be passed on to the customer. With its emphasis on early detection and prevention of problems, SPC has a distinct advantage over quality methods, such as inspection, that apply resources to detecting and correcting problems in the end product or service.

The underlying concept of statistical process control is based on a comparison of what is happening today with what happened previously. We take a snapshot of how the process typically performs or build a model of how we think the process will perform and calculate control limits for the expected measurements of the output of the process. Then we collect data from the process and compare the data to the control limits. The majority of measurements should fall within the control limits. Measurements that fall outside the control limits are examined to see if they belong to the same population as our initial snapshot or model. Stated differently, we use historical data to compute the initial control limits. Then the data are compared against these initial limits. Points that fall outside of the limits are investigated and, perhaps, some will later be discarded. If so, the limits would be recomputed and the process repeated.

In addition to reducing waste, SPC can lead to a reduction in the time required to produce the product or service from end to end. This is partially due to a diminished likelihood that the final product will have to be reworked, but it may also result from using SPC data

to identify bottlenecks, wait times, and other sources of delays within the process. Process cycle time reductions coupled with improvements in yield have made SPC a valuable tool from both a cost reduction and a customer satisfaction standpoint.

2.2 History of Statistical Process Control

Statistical process control was pioneered by Walter A. Shewhart in the early 1920s. W. Edwards Deming later applied SPC methods in the United States during World War II, thereby successfully improving quality in the manufacture of munitions and other strategically important products. Deming was also instrumental in introducing SPC methods to Japanese industry after the war had ended. Shewhart created the basis for the control chart and the concept of a state of statistical control by carefully designed experiments. While Dr. Shewhart drew from pure mathematical statistical theories, he understood that data from physical processes seldom produces a "normal distribution curve" (a Gaussian distribution, also commonly referred to as a "bell curve"). He discovered that observed variation in manufacturing data did not always behave the same way as data in nature (for example, Brownian motion of particles). Dr. Shewhart concluded that while every process displays variation, some processes display controlled variation that is natural to the process (common causes of variation), while others display uncontrolled variation that is not present in the process causal system at all times (special causes of variation).

In 1989, the Software Engineering Institute introduced the notion that SPC can be usefully applied to non-manufacturing processes, such as software engineering processes, in the **Capability Maturity Model** (CMM). This notion that SPC is a useful tool when applied to non-repetitive, knowledge-intensive processes such as engineering processes has encountered much skepticism, and remains controversial today.

The following description relates to manufacturing rather than to the service industry, although the principles of SPC can be successfully applied to either. For a description and example of how SPC applies to a service environment, refer to Roberts (2005.SPC has

also been successfully applied to detecting changes in organizational behavior with Social Network Change Detection introduced by McCulloh (2007). Selden describes how to use SPC in the fields of sales, marketing, and customer service, using Deming's famous Red Bead Experiment as an easy to follow demonstration.

In mass-manufacturing, the quality of the finished article was traditionally achieved through post-manufacturing inspection of the product; accepting or rejecting each article (or samples from a production lot) based on how well it met its design specifications. In contrast, Statistical Process Control uses statistical tools to observe the performance of the production process in order to predict significant deviations that may later result in rejected product.

Two kinds of variation occur in all manufacturing processes: both these types of process variation cause subsequent variation in the final product. The first is known as natural or common cause variation and may be variation in temperature, properties of raw materials, strength of an electrical current etc. This variation is small, the observed values generally being quite close to the average value. The pattern of variation will be similar to those found in nature, and the distribution forms the bell-shaped normal distribution curve. The second kind of variation is known as special cause variation, and happens less frequently than the first.

For example, a breakfast cereal packaging line may be designed to fill each cereal box with 500 grams of product, but some boxes will have slightly more than 500 grams, and some will have slightly less, in accordance with a distribution of net weights. If the production process, its inputs, or its environment changes (for example, the machines doing the manufacture begin to wear) this distribution can change. For example, as its cams and pulleys wear out, the cereal filling machine may start putting more cereal into each box than specified. If this change is allowed to continue unchecked, more and more product will be produced that fall outside the tolerances of the manufacturer or consumer, resulting in waste. While in this case, the waste is in the form of "free" product for the consumer, typically waste consists of rework or scrap.

By observing at the right time what happened in the process that led to a change, the quality engineer or any member of the team responsible for the production line can troubleshoot the root cause of the variation that has crept in to the process and correct the problem.

SPC indicates when an action should be taken in a process, but it also indicates when NO action should be taken. An example is a person who would like to maintain a constant body weight and takes weight measurements weekly. A person who does not understand SPC concepts might start dieting every time his or her weight increased, or eat more every time his or her weight decreased. This type of action could be harmful and possibly generate even more variation in body weight. SPC would account for normal weight variation and better indicate when the person is in fact gaining or losing weight.

War time production

Around the time of World War I, manufacturing processes typically became more complex with larger numbers of workers being supervised. This period saw the widespread introduction of mass production and piecework, which created problems as workmen could now earn more money by the production of extra products, which in turn led to bad workmanship being passed on to the assembly lines.

To counter bad workmanship, full time inspectors were introduced into the factory to identify quarantine and ideally correct product quality failures. Quality control by inspection in the 1920s and 1930s led to the growth of quality inspection functions, separately organized from production and big enough to be headed by superintendents.

The systematic approach to quality started in industrial manufacture during the 1930s, mostly in the USA, when some attention was given to the cost of scrap and rework. With the impact of mass production, which was required during the Second World War, it became necessary to introduce a more appropriate form of quality control which can be identified as Statistical Quality Control, or SQC. Some of the initial work for SQC is credited to Walter A. Shewhart of Bell Labs, starting with his famous one-page memorandum of 1924.

.

SQC came about with the realization that quality cannot be fully inspected into an important batch of items. By extending the inspection phase and making inspection organizations more efficient, it provides inspectors with control tools such as sampling and control charts, even where 100 per cent inspection is not practicable. Standard statistical techniques allow the producer to sample and test a certain proportion of the products for quality to achieve the desired level of confidence in the quality of the entire batch or production run.

Postwar

In the period following World War II, many countries' manufacturing capabilities that had been destroyed during the war were rebuilt. The U.S. sent General Douglas MacArthur to oversee the re-building of Japan. During this time, General MacArthur involved two key individuals in the development of modern quality concepts: W. Edwards Deming and Joseph Juran. Both individuals promoted the collaborative concepts of quality to Japanese business and technical groups, and these groups utilized these concepts in the redevelopment of the Japanese economy.

Although there were many individuals trying to lead United States industries towards a more comprehensive approach to quality, the U.S. continued to apply the QC concepts of inspection and sampling to remove defective product from production lines, essentially ignoring advances in QA for decades.

2.3 Definitions of Statistical Process Control

Statistical Process Control (SPC) is an effective method of monitoring a process through the use of control charts. Much of its power lies in the ability to monitor both process center and its variation about that center. By collecting data from samples at various points within the process, variations in the process that may affect the quality of the end product or service can be detected and corrected, thus reducing waste as well as the likelihood that problems will be passed on to the customer. It has an emphasis on early detection and prevention of problems.

.

As already stated the main objective in any production process is to control and maintain a satisfactory quality level for its product. In other words, it should be insured that the product conformed to specified quality standards, i.e. it should not contain a large number of defective items. This is termed *as "process control*" and is achieved through the technique of control charts invented by W.A.Shewhart in 1924.on other hand, by product control we mean controlling the quality of the product by critical examination at strategic points and this archived through '*sampling inspection plans'* invented by Dodge and Roming. Product control limits aims at guarantying a certain quality level is being maintained by the producer. In other words, it attempts to ensure that the product marketed for sale does not contain a large number of defective items.

Examples

A thermostat is a simple example for a closed control loop: It constantly measures the current temperature and controls the heater's valve setting to increase or decrease the room temperature according to the user-defined setting. A simple method switches the heater either completely on, or completely off, and an overshoot and undershoot of the controlled temperature must be expected. A more expensive method varies the amount of heat provided by the heater depending on the difference between the required temperature (the "set point") and the actual temperature. This minimizes over/undershoots.

Many organizations use statistical process control to bring the organization to Six Sigma levels of quality, in other words, so that the likelihood of an unexpected failure is confined to six standard deviations on the normal distribution. This probability is less than four one-millionths. Items controlled often include clerical tasks such as order-entry as well as conventional manufacturing tasks. Traditional statistical process controls in manufacturing operations usually proceed by randomly sampling and testing a fraction of the output. Variances in critical tolerances are continuously tracked and where necessary corrected before bad parts are produced.

.

2.3. Control Charts

One of the most important tools of production management and control of quality in the manufactured product is the '*control charts'* technique. The discovery and development of control charts were made by a young physicist Walter A. Shewhart of bell telephone laboratories in 1924 and the following years and is regards as a land mark in the history of industrial quality control. Based on the theory of probability and sampling, it enables us to detect the presence of assignable causes of erratic variation in the process. These causes are then identified and eliminated and the process is established and controlled at desired performances. In other words, control chart is in an indispensable tool for bringing a process under statistical control. As already pointed out, a process in statistical control implies that there are no apparent chaotic and assignable causes of variation and it will remain in control provided the conditions of manufacture the same. Such the process does not call for any corrective actions on the part of management because the entire variation in the product is due to chance causes which are beyond the control of human hand. The Shewhart's control charts provide us a very simple but powerful graphic method of finding if a process is in statistical control or not. Its construction is based plotting 3-σ limits and sequence of suitable sample statistic example, mean, range, standard deviation, fraction defective etc…computed from independent sample drawn at random from the product of the process. These sample points depict the frequency extent of variations from specified standards. Any sample point is going out side 3-σ control limits is an indication of the lack of statistical control. i.e. presence of some assignable cause of variation, which must be traced, identified and eliminated. A typical control charts consists of the following three horizontal lines.

- ✓ Upper control limit (UCL)
- ✓ Lower control limit (LCL)
- ✓ Central line (CL)

Control charts in this section are classified and described according to three general types: variables, attributes and multivariate.

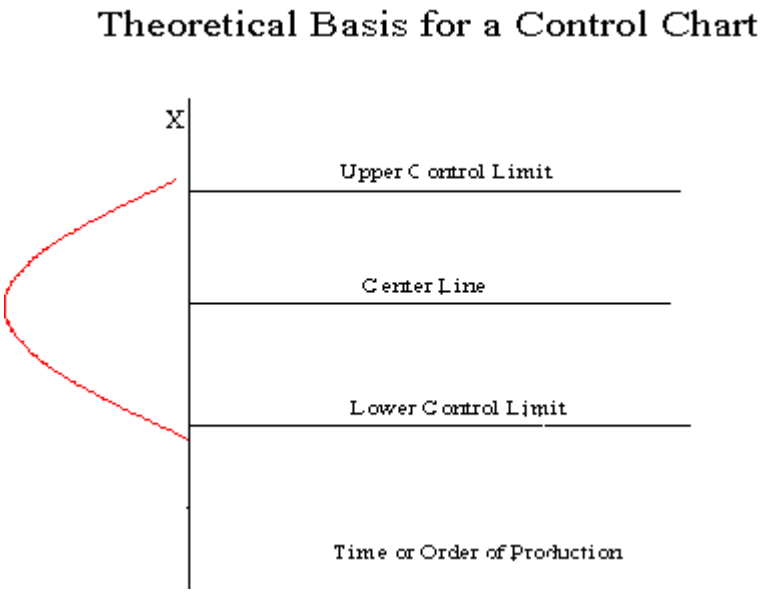
Comparison of univariate and multivariate control data

Control charts are used to routinely monitor quality. Depending on the number of process characteristics to be monitored, there are two basic types of control charts. The first, referred to as a univariate control chart, is a graphical display (chart) of one quality characteristic. The second, referred to as a multivariate control chart, is a graphical display of a statistic that summarizes or represents more than one quality characteristic.

Characteristics of control charts

If a single quality characteristic has been measured or computed from a sample, the control chart shows the value of the quality characteristic versus the sample number or versus time. In general, the chart contains a center line that represents the mean value for the in-control process. Two other horizontal lines, called the upper control limit (UCL) and the lower control limit (LCL) are also shown on the chart. These control limits are chosen so that almost all of the data points will fall within these limits as long as the process remains in-control. The figure below illustrates this.

Chart demonstrating basis of control chart



Why control charts "work"

.

The control limits as pictured in the graph might be .001 *probability* limits. If so, and if chance causes alone were present, the probability of a point falling above the upper limit would be one out of a thousand, and similarly, a point falling below the lower limit would be one out of a thousand. We would be searching for an assignable cause if a point would fall outside these limits. Where we put these limits will determine the risk of undertaking such a search when in reality there is no assignable cause for variation.

Since two out of a thousand is a very small risk, the 0.001 limits may be said to give practical assurances that, if a point falls outside these limits, the variation was caused be an assignable cause. It must be noted that two out of one thousand is a purely arbitrary number. There is no reason why it could not have been set to one out a hundred or even larger. The decision would depend on the amount of risk the management of the quality control program is willing to take. In general (in the world of quality control) it is customary to use limits that approximate the 0.002 standard.

Letting X denote the value of a process characteristic, if the system of chance causes generates a variation in X that follows the normal distribution, the 0.001 probability limits will be very close to the 3σ limits. From normal tables we glean that the 3σ in one direction is 0.00135, or in both directions 0.0027. For normal distributions, therefore, the 3σ limits are the practical equivalent of 0.001 probability limits.

Plus or minus "3 sigma" limits are typical

In the U.S., whether X is normally distributed or not, it is an acceptable practice to base the control limits upon a multiple of the standard deviation. Usually this multiple is 3 and thus the limits are called 3-sigma limits. This term is used whether the standard deviation is the universe or population parameter, or some estimate thereof, or simply a "standard value" for control chart purposes. It should be inferred from the context what standard deviation is involved. (Note that in the U.K., statisticians generally prefer to adhere to probability limits.)

If the underlying distribution is skewed, say in the positive direction, the 3-sigma limit will fall short of the upper 0.001 limit, while the lower 3-sigma limit will fall below the

.

0.001 limit. This situation means that the risk of looking for assignable causes of positive variation when none exists will be greater than one out of a thousand. But the risk of searching for an assignable cause of negative variation, when none exists, will be reduced. The net result, however, will be an increase in the risk of a chance variation beyond the control limits. How much this risk will be increased will depend on the degree of skew ness.

If variation in quality follows a Poisson distribution, for example, for which *np* = .8, the risk of exceeding the upper limit by chance would be raised by the use of 3-sigma limits from 0.001 to 0.009 and the lower limit reduces from 0.001 to 0. For a Poisson distribution the mean and variance both equal *np*. Hence the upper 3-sigma limit is 0.8 + 3 *sqrt*(.8) = 3.48 and the lower limit = 0 (here *sqrt* denotes "square root"). For *np* = .8 the probability of getting more than 3 successes = 0.009.

Strategies for dealing with out-of-control findings

If a data point falls outside the control limits, we assume that the process is probably out of control and that an investigation is warranted to find and eliminate the cause or causes.

Does this mean that when all points fall within the limits, the process is in control? Not necessarily. If the plot looks non-random, that is, if the points exhibit some form of systematic behavior, there is still something wrong. For example, if the first 25 of 30 points fall above the center line and the last 5 fall below the center line, we would wish to know why this is so. Statistical methods to detect sequences or nonrandom patterns can be applied to the interpretation of control charts. To be sure, "in control" implies that all points are between the controls limits and they form a random pattern.

2.3.1 Tools for Statistical Quality Control

During the 1920's, Dr. Walter A. Shewhart proposed a general model for control charts as follows:

Shewhart Control Charts for variables

Let *w* be a sample statistic that measures some continuously varying quality characteristic of interest (e.g., thickness), and suppose that the mean of *w* is μ_w, with a standard deviation of σ_w. Then the center line, the UCL and the LCL are

UCL = $\mu_w + k\sigma_w$
Center Line = μ_w
LCL = $\mu_w - k\sigma_w$

Where *k* is the distance of the control limits from the center line, expressed in terms of standard deviation units. When *k* is set to 3, we speak of 3-sigma control charts.

Historically, k = 3 has become an accepted standard in industry.

The centerline is the process mean, which in general is unknown. We replace it with a *target* or the average of all the data. The quantity that we plot is the sample average, $\bar{X}$. The chart is called the $\bar{X}$ chart.

We also have to deal with the fact that σ is, in general, unknown. Here we replace σ_w with a given standard value, or we estimate it by a function of the *average standard deviation*. This is obtained by averaging the individual standard deviations that we calculated from each of *m* preliminary (or present) samples, each of size *n.* This function will be discussed shortly.

It is equally important to examine the standard deviations in ascertaining whether the process is in control. There is, unfortunately, a slight problem involved when we work with the usual estimator of σ. The following discussion will illustrate this.

Sample Variance If σ^2 is the unknown variance of a probability distribution, then an unbiased estimator of σ^2 is the sample variance

$$s^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$$

However, *s*, the sample standard deviation is *not* an unbiased estimator of σ. If the underlying distribution is normal, then *s* actually estimates $c_4\,\sigma$, where c_4 is a constant that depends on the sample size *n*. This constant is tabulated in most text books on statistical quality control and may be calculated using

C_4 factor

$$c_4 = \sqrt{\frac{2}{n-1}}\,\frac{\left(\frac{n}{2}-1\right)!}{\left(\frac{n-1}{2}-1\right)!}$$

To compute this we need a *non-integer factorial,* which is defined for *n*/2 as follows:

Fractional Factorials

$$\left(\frac{n}{2}\right)! = \left(\frac{n}{2}\right)\left(\frac{n}{2}-1\right)\left(\frac{n}{2}-2\right)\cdots\left(\frac{1}{2}\right)\sqrt{\pi}$$

For example, let $n = 3.5 = 7/2$. Then

$$\left(\frac{7}{2}\right)! = (3.5)! = (3.5)(2.5)(1.5)(.5)(1.77246) = 11.632$$

With this definition the reader should have no problem verifying that the c_4 factor for $n = 10$ is .9727.

Mean and standard deviation of the estimators

So the mean or expected value of the sample standard deviation is $c_4\,\sigma$.

The standard deviation of the sample standard deviation is

$$\sigma_s = \sigma\sqrt{1-c_4^2}$$

What are the differences between control limits and specification limits?

Control limits vs. specifications

Control Limits are used to determine if the process is in a state of statistical control (i.e., is producing consistent output).

Specification Limits are used to determine if the product will function in the intended fashion.

How many data points are needed to set up a control chart?
How many samples are needed?

Shewhart gave the following rule of thumb:

"It has also been observed that a person would seldom if ever be justified in concluding that a state of statistical control of a given repetitive operation or production process has been reached until he had obtained, under presumably the same essential conditions, a sequence of not less than twenty five samples of size four that are in control."

It is important to note that control chart properties, such as false alarm probabilities, are generally given under the assumption that the parameters, such as μ and σ, are known. When the control limits are not computed from a large amount of data, the actual properties might be quite different from what is assumed.

.

When do we recalculate control limits?

When do we recalculate control limits?

Since a control chart "compares" the current performance of the process characteristic to the past performance of this characteristic, changing the control limits frequently would negate any usefulness.

So, only change your control limits if you have a valid, compelling reason for doing so. Some examples of reasons:

- When you have at least 30 more data points to add to the chart and there have been no known changes to the process - you get a better estimate of the variability
- If a major process change occurs and affects the way your process runs.
- If a known, preventable act changes the way the tool or process would behave (power goes out, consumable is corrupted or bad quality, etc.)

.

As already pointed out the control of quality in manufactured product can be broadly in to two heads;

1. Control charts for variable (mean, range standard deviation charts)

2. Control chats for attributes (p, np and c chats)

2.3.2 Variable Control Charts

As pointed out above, 'control charts' is the technique of process control. A number of control charts have been developed depending on the ways of assessing the quality of manufactured product. Control charts for variables are designed to achieved and maintain a satisfactory quality level for a process whose product is amenable to quantitative measurements like the thickness, length or diameter of screw or nut, weight of the bolts, tensile strength of yarn or steel pipes, resistance of a wire, etc. the observation on such units can be expressed in specific unit of measurements. In such cases the quality control involves the control of variation both in measures of central tendency and depression of the characteristics. Variable control charts are used when quality is measured as variables (length, weight, tensile strength, etc.). The main purpose of the variable control charts is to monitor the process mean and the standard deviation.The variables under consideration are of continuous character and are assumed to be distributed normally. Control charts for variables are:

1. Control chart for mean

2. Control charts for range(R)

3. Control charts for standard deviation(S)

2.3.2.1 Shewhart $\bar{X}$ and R and S Control Charts

$\bar{X}$ and S Shewhart Control Charts

We begin with $\bar{X}$and *S* charts. We should use the *s* chart first to determine if the distribution for the process characteristic is stable.

.

Let us consider the case where we have to estimate σ by analyzing past data. Suppose we have *m* preliminary samples at our disposition, each of size *n*, and let s_i be the standard deviation of the *i*th sample. Then the average of the *m* standard deviations is

$$\bar{s} = \frac{1}{m}\sum_{i=1}^{m} s_i$$

Control Limits for $\overline{X}$ *and S Control Charts*

We make use of the factor c_4 described on the previous page.

The statistic $\bar{s}/c_4$ is an unbiased estimator of σ. Therefore, the parameters of the *S* chart would be

$$UCL = \bar{s} + 3\frac{\bar{s}}{c_4}\sqrt{1 - c_4^2}$$

$$\text{Center Line} = \bar{s}$$

$$LCL = \bar{s} - 3\frac{\bar{s}}{c_4}\sqrt{1 - c_4^2}$$

Similarly, the parameters of the $\overline{X}$ chart would be

$$UCL = \bar{\bar{x}} + 3\frac{\bar{s}}{c_4\sqrt{n}}$$

$$\text{Center Line} = \bar{\bar{x}}$$

$$LCL = \bar{\bar{x}} - 3\frac{\bar{s}}{c_4\sqrt{n}}$$

$\bar{\bar{x}}$, the "grand" mean is the average of all the observations.

It is often convenient to plot the $\overline{X}$ and *s* charts on one page.

Control Limits for $\overline{X}$ *and R Control Charts*

$\overline{X}$ *and R control charts*

If the sample size is relatively small (say equal to or less than 10), we can use the range

instead of the standard deviation of a sample to construct control charts on $\bar{X}$ and *the range, R*. The range of a sample is simply the difference between the largest and smallest observation.

There is a statistical relationship between the mean range for data from a normal distribution and σ, the standard deviation of that distribution. This relationship depends only on the sample size, *n*. The mean of *R* is $d_2\,\sigma$, where the value of d_2 is also a function of *n*. *An estimator of* σ is therefore R/d_2.

Armed with this background we can now develop the $\bar{X}$ and *R* control chart.

Let $R_1, R_2 \ldots R_k$, be the range of *k* samples. The average range is

$$\bar{R} = \frac{R_1 + R_2 + \ldots + R_k}{k}$$

Then an estimate of σ can be computed as

$$\hat{\sigma} = \frac{\bar{R}}{d_2}$$

This chart controls the process variability since the sample range is related to the process standard deviation. *The center line of the R chart is the average range.*

To compute the control limits we need an estimate of the true, but unknown standard deviation $W = R/\sigma$. This can be found from the distribution of $W = R/\sigma$ (assuming that the items that we measure follow a normal distribution). The standard deviation of *W* is d_3, and is a known function of the sample size, *n*. It is tabulated in many textbooks on statistical quality control.

Therefore since $R = W\sigma$, the standard deviation of *R* is $\sigma_R = d_3\,\sigma$. But since the true σ is unknown, we may estimate σ_R by

$$\hat{\sigma}_R = d_3 \frac{\bar{R}}{d_2}$$

As a result, the parameters of the *R* chart with the customary 3-sigma control limits are

$$UCL = \bar{R} + 3\hat{\sigma}_R = \bar{R} + 3d_3\frac{\bar{R}}{d_2}$$

$$\text{Center Line} = \bar{R}$$

$$LCL = \bar{R} - 3\hat{\sigma}_R = \bar{R} - 3d_3\frac{\bar{R}}{d_2}$$

As was the case with the control chart parameters for the subgroup averages, defining another set of factors will ease the computations, namely:

$D_3 = 1 - 3\ d_3 / d_2$ and $D_4 = 1 + 3\ d_3 / d_2$. These yield

$$UCL = \bar{R}D_4$$

$$\text{Center Line} = \bar{R}$$

$$LCL = \bar{R}D_3$$

The factors D_3 and D_4 depend only on *n*, and are tabled below.

Factors for Calculating Limits for $\bar{X}$ and R Charts

n	A_2	D_3	D_4
2	1.880	0	3.267
3	1.023	0	2.575
4	0.729	0	2.282
5	0.577	0	2.115
6	0.483	0	2.004
7	0.419	0.076	1.924
8	0.373	0.136	1.864
9	0.337	0.184	1.816
10	0.308	0.223	1.777

In general, the range approach is quite satisfactory for sample sizes up to around 10.

.

For larger sample sizes, using subgroup standard deviations is preferable. For small sample sizes, the relative efficiency of using the range approach as opposed to using standard deviations is shown in the following table.

Efficiency of Yon S versus R

n	Relative Efficiency
2	1.000
3	0.992
4	0.975
5	0.955
6	0.930
10	0.850

A typical sample size is 4 or 5, so not much is lost by using the range for such sample sizes.

Generally the construction of $\overline{X}$ and **R** chart

1. for construction of x bar charts: - the control charts for means is drawn on the graph paper by taking the sample number along the horizontal scale, (x-axis) and the statistics $\overline{X}$ along the vertical scale (y-axis). Sample points (sample means $\overline{X}_1$, $\overline{X}_2 \ldots \overline{X}_k$) are then plotted as plotted as points (dots) against the corresponding sample number. These points may or may not be joined. The central line is drawn as a bold (dark) horizontal line at $\mu = \dot{\mu}$ (if μ is known) or at $\overline{\overline{X}}$ (if μ is not known). The UCL $\overline{X}$ and LCL $\overline{X}$ are plotted as dotted horizontal lines at the computed values given.
2. Construction of R charts: - as incase of $\overline{X}$-charts, the samples or subgroups of numbers is taken along horizontal scale and the statistics (Range) is taken along vertical scale. The sample points R_1, R_2, $R_3 \ldots R_k$ are then plotted as points (dots) against the corresponding sample numbers. The central line is taken as the hold horizontal lines at R-bar and UCLR and LCLR are plotted as dotted horizontal lines at the computed values given.

Interpretations of $\overline{X}$ and R Charts:

1. The process is termed to be in statistical control if both the $\overline{X}$ and R charts exhibit control. i.e. all the sample points fall with in the control limits in both charts. Such situation implies that the process is operating under the influence of chance causes of variations in the product is natural variation. A process in control implies that there is no reason for worry as there are no apparent assignable or erratic causes of variation. If the process exhibits statistical control for quite some time and for sizable volume of out put then we believe that it will remain in control if the conditions of manufacture remain the same. It should be clearly understood that a process in control does not imply that its product is satisfactory. Product control is

achieved through sampling Inspection plan discussed later.

2. If one or more of the points in any or both the charts go out of the control limits we say that the process is out of control, i.e. it is not in state of statistical control. Such a situation indicates the presence of some assignable causes of erratic fluctuations which must be traced, identified and eliminated, so that the process may return to operation under stable statistical conditions. Some of the causes of assignable of variations are: defective or substandard raw materials, substandard or faulty tools or equipment. Inefficient or unskilled operators, negligence of the operator, break down of lubrication system, improper setting or wrong handling of the machines, change of working conditions, viz. , new materials or machines…etc.
3. $\overline{X}$-charts reveals undesirable variations between samples as far as their averages are concerned while, R-charts reveals any un-desirable variation with in the samples. The $\overline{X}$-charts primarily allow detection of shift in the process averages and assume that process variation remains essentially constant. It is quite possible that the same value for the mean of sample could conceal a shift in the variability of the process. it is because of this reason that both $\overline{X}$and R charts should be studied together to arrive at a decision if the process is in control or not . The choice between $\overline{X}$and R charts is a managerial problem. From practical point of view it is better to construct R chart first. If R charts indicates lack of control, i.e. the variability (dispersion) of the quality by the process is out of control, generally it is better not to construct $\overline{X}$-charts unless the quality dispersion is brought under control.

Example 1 : A machine is set to deliver packets of a given weight. 10 sample of size 5 each were recoded. Below a given relevant data:

Sample Number	1	2	3	4	5	6	7	8	9	10
Mean ($\overline{X}$)	15	17	15	18	17	14	18	15	17	16
Range (R)	7	7	4	9	8	7	12	4	11	5

Calculate the *central line and the control limit* for mean chart and the range chart also

draw the charts and then comment the state of control.

Solution: from the above table the conversion factor for n=5 are A_2=0.5, D_3=0,D_4=2.115.

From the above sample data we get:

$\overline{\overline{X}} = \frac{1}{10}\sum_{i=1}^{10} \overline{X}_i$ =162\10= 16.2

R (grand) = $\frac{1}{10}\sum_{i=1}^{10} R_i$ =74\10=7.4

As given above for n=5 we have:

A_2=0.5, D_3=0 and D_4=2.115.

Mean chart. 3-σ.control limits for $\overline{\overline{X}}$-charts are :

UCL $\overline{\overline{X}}$= $\overline{\overline{X}}$+A_2*R(Grand)

= 16.2+ 0.53*7.4 = 20.492

LCL $\overline{\overline{X}}$= $\overline{\overline{X}}$-A_2*R(Grand)

= 16.2-0.53*7.4 =11.908

CL $\overline{\overline{X}}$= $\overline{\overline{X}}$=16.2

.

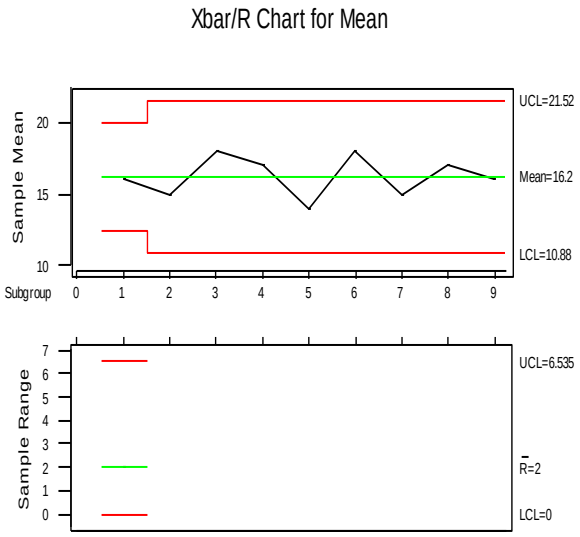


Figure 3: shows $\overline{\overline{X}}$ and R For above sample data.

Conclusion: Based on the above sample data of weights of a certain packets of products since all the sample points (means) fall with in the control limits. This chart shows that the process is in state of statistical process control.

Example 2: you are given the values of sample means ($\overline{\overline{X}}$) and the range(R) ten samples of five size of each. Draw the mean ($\overline{\overline{X}}$) and range(R) charts and comments on the state of control of the process.

Sample Number	1	2	3	4	5	6	7	8	9	10
$\overline{\overline{X}}$	43	49	37	44	45	37	51	46	43	47
R	5	6	5	7	7	4	8	6	4	6

(You may use the following control chart constants from the above table)

For n=5 are A_2=0.5, D_3=0 and D_4=2.115.

Solution: for mean chart

$$\overline{\overline{X}} = \frac{1}{10}\sum_{i=1}^{10}\overline{X}_i = 442\backslash 10 = 44.2$$

$$R(\text{Grand}) = \frac{1}{10}\sum_{i=1}^{10}R_i = 58\backslash 10 = 5.8$$

As given above for n=5 we have:

A_2=0.5, D_3=0 and D_4=2.115.

3-σ.control limits for $\bar{\bar{X}}$-charts are:

UCL $\bar{\bar{X}}$= $\bar{\bar{X}}$+A_2*R (Grand)

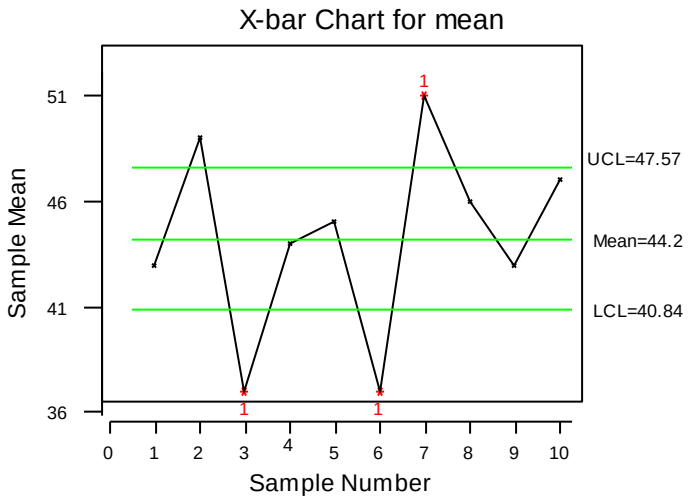
= 44.2+ 0.58*5.8 =47.562

LCL $\bar{\bar{X}}$= $\bar{\bar{X}}$-A_2*R (Grand)

= 44.2-0.58*5.8 =40.836

CL $\bar{\bar{X}}$= $\bar{\bar{X}}$=44.2

The control charts for mean is given below.



From the chart we can conclude that, we see that the sample points (sample mean) corresponding to the second, the third, the fourth and the seventh samples are lies out side the control limits. Hence the process is out of control. This implies that some assignable causes of variation are operating which should be detected and corrected.

Rage chart (R chart):

3-σ.control limits for R-charts are

UCLR= D_4*R (Grand)

= 2.115*5.8 =12.267

LCLR= D_3*R= 0*5.8 = 0

CLR= R (Grand) = 5.8

Like wise since the entire sample points (sample range) fall with in the control limits, R charts shows that the process is in statistical control. Although R-charts depict control, the process can not re-grade to be in statistical control since $\overline{X}$-charts shows lack of control.

Generally for the construction of S (standard deviation chart):

Since standard deviation is an ideal measure of dispersion, a combinations of control chart for mean ($\overline{X}$) and standard deviation (σ.) known as $\overline{X}$ and σ .charts theoretically more appropriate than the combination of $\overline{X}$ and R charts for controlling the process average and process variability.

$$c_4 = \sqrt{\frac{2}{n-1}}\ \frac{\left(\frac{n}{2}-1\right)!}{\left(\frac{n-1}{2}-1\right)!}$$

Where c_4 is a constant that depends on the sample size n. This constant is tabulated in most text books on statistical quality control and may be calculated using the above formula.

We make use of the factor c_4 described on the previous page.

The statistic $\bar{s}/c_4$ is an unbiased estimator of σ. Therefore, the parameters of the *S* chart would be

$$UCL = \bar{s} + 3\frac{\bar{s}}{c_4}\sqrt{1 - c_4^2}$$

$$\text{Center Line} = \bar{s}$$

$$LCL = \bar{s} - 3\frac{\bar{s}}{c_4}\sqrt{1 - c_4^2}$$

Similarly, the parameters of the $\overline{\text{X}}$chart would be

$$UCL = \bar{\bar{x}} + 3\frac{\bar{s}}{c_4\sqrt{n}}$$

$$\text{Center Line} = \bar{\bar{x}}$$

$$LCL = \bar{\bar{x}} - 3\frac{\bar{s}}{c_4\sqrt{n}}$$

$\bar{\bar{x}}$, the "grand" mean is the average of all the observations.

It is often convenient to plot the $\overline{\text{X}}$and *S* charts on one page (graph).

We can also the 3-$\boldsymbol{\sigma}$. Control limits for control charts for standard deviation is given by:

UCLs = $B_2\boldsymbol{\sigma}$.

LCLs = $B_1\boldsymbol{\sigma}$. Where $\boldsymbol{\sigma}$ is the process standard deviation. If $\boldsymbol{\sigma}$ is not known then its estimate is based on the average sample of standard deviation used. In this case ($\boldsymbol{\sigma}$ is unknown).

UCLs = B_4(S(grand))

LCLs =B_3(S(grand))

Where S (grand) = sum of sample standard deviation \number of samples.

The values of the constant B_1, B_2, B_3, and B_4 for different sample size are given in table 'X' in appendix.

The other steps in the construction of s-chart are same as the case of R-chart. The sample number is taken along horizontal scale and the statistic standard deviation is taken along the vertical scale. The sample standard deviation s_1, s_2, s_3….s_k are plotted as points

against the corresponding sample numbers. The central line is drawn at s(grand) and UCLs and LCLs are plotted as dotted line at the computed values given.

S-charts are interpreted the same as R-charts.

Remark: *s-charts vs. R-charts*: theoretically s-chart appears to be more appropriate than R-charts for controlling dispersion of the process. But the difficulty of computation makes the use of s-chart almost impractical in most industry work. In practice, R-chart is preferred to s-chart because of its computational case. Fortunately, for small samples (which is generally the case in control chat) the range, as a measure of dispersion is sufficiently stable to be useful, because for small samples the range R and the standard deviation s are likely to fluctuate together, i.e. is s is small(large) , R is also likely to be small(large).

Moreover**,** in most of the situations is quite in expensive to take samples from the process. Hence a reasonable amount of accuracy can be achieved by taking more samples rather than spending the valuable time in computing the sample the sample standard deviation. However, if sampling cost per unit is quite high then we can not afford to take more samples because of the cost of the time spent in computing the sample standard deviations will be relatively much less than the cost of talking samples. In such a situation the desired accuracy can be achieved more economically through the use of $\overline{X}$ and s charts.

Exercise: The following data gives reading 10 samples of size 6 in the production of a certain component.

Sample Number	1	2	3	4	5	6	7	8	9	10
Mean ($\overline{X}$)	383	508	505	532	557	337	514	614	707	753
S.D(s)	30.5	41.6	39.5	25.4	24.2	48.7	89	13.1	33.9	32.2
Range (R)	95	128	100	91	68	65	148	28	37	80

Draw the control charts for $\bar{X}$, R and σ. Comments the state of control in both three charts is there the same answer?

Time To Detection or Average Run Length (ARL)

Two important questions when dealing with control charts are:

1. How often will there be false alarms where we look for an assignable cause but nothing has changed?
2. How quickly will we detect certain kinds of systematic changes, such as mean shifts?

The ARL tells us, for a given situation, how long on the average we will plot successive control charts points before we detect a point beyond the control limits.

For an $\bar{X}$ chart, with no change in the process, we wait on the average $1/p$ points before a false alarm takes place, with p denoting the probability of an observation plotting outside the control limits. For a normal distribution, $p = .0027$ and the ARL is approximately 371.

A table comparing Shewhart $\bar{X}$ chart *ARL*'s to Cumulative Sum (CUSUM) *ARL*'s for various mean shifts is given later in this section.

2.4 Control Charts for Attributes

The Shewhart control chart plots quality characteristics that can be measured and expressed numerically. We measure weight, height, position, thickness, etc. If we cannot represent a particular quality characteristic numerically, or if it is impractical to do so, we then often resort to using a quality characteristic to sort or classify an item that is inspected into one of two "buckets". In cases where quality is measured as attributes (number of defects in a component or a product or a batch of components or products, number or proportion of defectives in a batch, etc.), attribute control charts

are used.

An example of a common quality characteristic classification would be designating units as "conforming units" or "nonconforming units". Another quality characteristic criteria would be sorting units into "non defective" and "defective" categories. Quality characteristics of that type are called attributes.

Note that there is a difference between "nonconforming to an engineering specification" and "defective" - a nonconforming unit may function just fine and be, in fact, not defective at all, while a part can be "in spec" and not function as desired (i.e., be defective).

Examples of quality characteristics that are attributes are the number of failures in a production run, the proportion of malfunctioning wafers in a lot, the number of people eating in the cafeteria on a given day, etc.

In spite of wide application of shewart's control charts for variables, viz. $\bar{X}$ and R charts or $\bar{X}$ and s charts are very powerful tools for detections of assignable causes of erratic fluctuations in productions from a repetitive process, they have a certain limitation as given below.

1. These charts can be used only when the quality characteristic can be measured quantitatively and expressed in a certain units of measurement. In other words, they can be used only when we can take numerical observations on the sample units.
2. Quite often, we make a large number of observations on each of the sampled unit. For example, we can take measurement of length, weight, diameter, thickness, tensile strength on each of the sample units, say, steel pipes. Each of these characteristics could be a possible candidate for $\bar{X}$ and R or $\bar{X}$ and s charts. Things become very difficult if the number of such candidates is very large because it is very inconvenient, impracticable and uneconomical to draw a very large number of control charts.

As an alternative of to $\overline{X}$ and R or $\overline{X}$ and s charts we have the control charts for attributes which are used:

(i) When we deal with quality characteristics which can not be measured quantitatively. In such cases the inspection of units is accompanied by classifying them as acceptable or non acceptable, defective or non defective.

(ii) When we deal with characteristics which are actually observed as attributes although they could be measured quantitatively.

Remark: we shall be using two words '*defect*' and '*defective*' in connection with control charts for attributes. Any instance of a characters tics or unit not conforming to specification (required standards) is termed as *defect.* A defective is a unit which contains more than allowable number (usually one) of defects.

Types of control charts for attributes

Control charts dealing with the number of *defects* or *nonconformities* are called *c* charts (for count).

Control charts dealing with the *proportion* or *fraction* of defective product are called *p* charts (for proportion).

There is another chart which handles *defects per unit*, called the *u* chart (for unit). This applies when we wish to work with the average number of nonconformities per unit of product.

Generally control charts for attributes are:

(a) control charts for number of defects per unit, i.e. c harts
(b) control charts for proportion or fraction of defectives p charts
(c) Control charts for number of defectives i.e. np chats or d charts.

2.4.1 Control Charts for Number of Defects per unit (c-chart)

As already pointed out defects as any instance of a characters tics or unit not conforming to one or more of the given specification or standards. Accordingly every defective unit contains one or more defects. For instance, a defective casting may further be examined for blow holes, could shuts, rough surface, weak structure etc…

The literature differentiates between *defects* and *defective*, which is the same as differentiating between *nonconformity* and *nonconforming units*. This may sound like splitting hairs, but in the interest of clarity let's try to unravel this man-made mystery.

Consider a wafer with a number of chips on it. The wafer is referred to as an "item of a product". The chip may be referred to as "a specific point". There exist certain specifications for the wafers. When a particular wafer (e.g., the item of the product) does not meet at least one of the specifications, it is classified as a nonconforming item. Furthermore, each chip, (e.g., the specific point) at which a specification is not met becomes a defect or nonconformity.

So, a nonconforming or defective item contains at least one defect or nonconformity. It should be pointed out that a wafer can contain several defects but still be classified as conforming. For example, the defects may be located at no critical positions on the wafer. If, on the other hand, the number of the so-called "unimportant" defects becomes alarmingly large, an investigation of the production of these wafers is warranted.

Control charts involving counts can be either for the *total number* of nonconformities (defects) for the sample of inspected units, or for the *average number* of defects per inspection unit.

The c chart is used when we count the number of defects per unit rather than classifying a unit as good or bad, i.e. defective or non-defective. In a number of manufacturing processes we come across situations where:

(i) The opportunities for the occurrence of a defect in any unit are very large.

.

(ii) The actual occurrences of a defect are rare, i.e. the probability of occurrences of a defect in any spot is very small.

Poisson approximation for numbers or counts of defects

Let us consider an assembled product such as a microcomputer. The opportunity for the occurrence of any given defect may be quite large. However, the probability of occurrence of a defect in any one arbitrarily chosen spot is likely to be very small. In such a case, the incidence of defects might be modeled by a *Poisson* distribution. Actually, the Poisson distribution is an approximation of the *binomial* distribution and applies well in this capacity according to the following rule of thumb:

The sample size n should be equal to or larger than 20 and the probability of a single success, p, should be smaller than or equal to .05. If $n \geq 100$, the approximation is excellent if np is also ≤ 10.

Illustrate Poisson approximation to binomial

To illustrate the use of the Poisson distribution as an approximation of a binomial distribution, consider the following comparison: Let p, the probability of a single success in $n = 200$ trials, be .025.

Find the probability of exactly 3 successes. If we assume that p remains constant then the solution follows the binomial distribution rules, that is:

$$p(x) = \binom{n}{x} p^x (1-p)^{n-x} = \binom{200}{3} .025^3 .975^{197} = 0.1399995$$

By the Poisson approximation we have

$$c = (200)(.025)$$

and

Before the control chart parameters are defined there is one more definition: *the inspection unit.*

practioners.

Example 1 may help to illustrate the construction of control limits for counts data. We are inspecting 25 successive wafers, each containing 100 chips. Here the wafer is the inspection unit. The observed number of defects are

Wafer Number	Number of Defects	Wafer Number	Number of Defects
1	16	14	16
2	14	15	15
3	28	16	13
4	16	17	14
5	12	18	16
6	20	19	11
7	10	20	20
8	12	21	11
9	10	22	19
10	17	23	16
11	19	24	31
12	17	25	13
13	14		

From this table we have

$$\bar{c} = \frac{\text{total number of defects}}{\text{total number of samples}} = \frac{400}{25} = 16$$

$$UCL = \bar{c} + 3\sqrt{\bar{c}} = 16 + 2\sqrt{16} = 28$$

Control Chart for Counts

Transforming Poisson Data

We have seen that the 3-sigma limits for a *c* chart, where *c* represents the number of nonconformities, are given by

$$\bar{c} \pm 3\sqrt{\bar{c}}$$

where it is assumed that the normal approximation to the Poisson distribution holds, hence the symmetry of the control limits. It is shown in the literature that the normal approximation to the Poisson is adequate when the mean of the Poisson is at least 5. When applied to the *c* chart this implies that the mean of the defects should be at least 5. This requirement will often be met in practice, but still, when the mean is smaller than 9 (solving the above

equation) there will be no lower control limit.

Let the mean be 10. Then the lower control limit = 0.513. However, $P(c = 0)$ = .000045, using the Poisson formula. This is only 1/30 of the assumed area of .00135. So one has to raise the lower limit so as to get as close as possible to .00135. From Poisson tables or computer software we find that $P(1)$ = .0005 and $P(2)$ = .0027, so the lower limit should actually be 2 or 3.

that are highly skewed (see Ryan and Schwertman (1997) for more about the possibly extreme consequences of doing this).

Uses of c charts: in spite of its restricted field of application as compared with $\bar{X}$ and R charts, a number of practical situations exist in many industries where c-chart is needed. Some of its field of application given below:

1. Number of defects observed in a bale of cloth, sheet of photographic film, roll of coated paper etc..
2. Number of defects in a galvanized sheet or a painted, plated or enamel surface of a given area.
3. Number of defects of all types in air crafts sub-assemblies or final assembly and so on.

Example 2: *during an examination of equal length of cloth, the following are the number of defects is observed.*

Sample number	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Number of defects	*2*	*3*	*4*	*0*	*5*	*6*	*7*	*4*	*3*	*2*

Draw a control chart for the number of defects and comment weather the process is under control or not?

Solution: let the number of defects per unit (equal length of cloth) be denoted by c. then the average number of defects in the 10 sample units is given by:

C (grand) = total number of defects \total number of samples

= 2+3+4+…+2\10

= 36\10 =3.6

The 3-σ control limits for c-charts are given by:

$$UCL = c + k\sqrt{c}$$

= 3.6 +3 * √3.6

= 3.6+3*1.8974

= 3.6+5.6922 = 9.2922

= 3.6 -3*√3.6

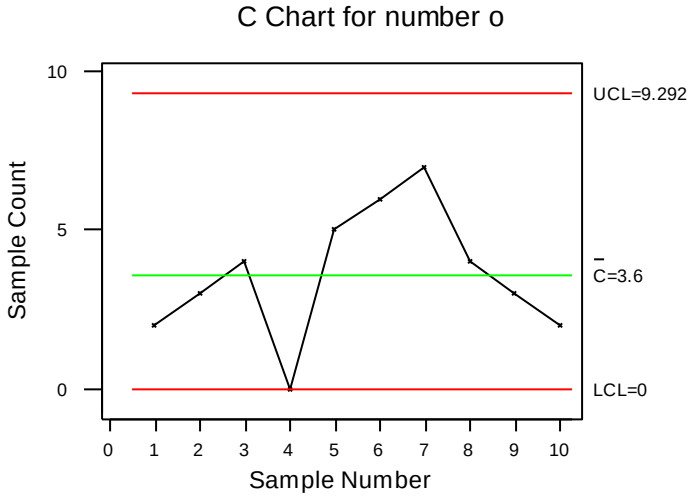
= 3.6- 5.6922

-2.0922 ~ = 0, since the number of defects can not be negative.

$$CenterLine = c$$

CLs= 3.6

The c- chart i.e. the control charts for the number of defects is obtained on plotting c values against the corresponding sample number, together with the control limits and is given below.



Since the entire sample points are with in the control limits, the process is in statistical control.

2.4.2 Control Charts for Proportion or Fraction of Defects (P charts)

P is the fraction defective in a lot or population

The proportion or fraction nonconforming (defective) in a population is defined as the ratio of the number of nonconforming items in the population to the total number of items in that population. The item under consideration may have one or more quality characteristics that are inspected simultaneously. If at least one of the characteristics does not conform to standard, the item is classified as nonconforming.

The fraction or proportion can be expressed as a decimal, or, when multiplied by 100, as a percent. The underlying statistical principles for a control chart for proportion nonconforming are based on the binomial distribution.

Let us suppose that the production process operates in a stable manner, such that the probability that a given unit will not conform to specifications is *p*. Furthermore, we assume that successive units produced are independent. Under these conditions, each unit that is produced is a realization of a Bernoulli random variable with parameter *p*. If a random sample of *n* units of product is selected and if *D* is the number of units that are nonconforming, the *D* follows a binomial distribution with parameters *n* and *p*

The binomial distribution model for number of defectives in a sample

$$p\{D = x\} = \binom{n}{x} p^x (1-p)^{n-x} \qquad x = 0,1,\cdots,n$$

The mean of ***D*** is ***np*** and the variance is ***np (1-p).*** The *sample* proportion nonconforming is the ratio of the number of nonconforming units in the sample, *D*, to the sample size *n*,

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

The mean and variance of this estimator are

$$\mu = p$$

And

This background is sufficient to develop the control chart for proportion or fraction nonconforming. The chart is called the *p*-chart.

P control charts for lot proportion defective

If the true fraction conforming *p* is known (or a standard value is given), then the center line and control limits of the fraction nonconforming control chart is

When the process fraction (proportion) *p* is not known, it must be estimated from the available data. This is accomplished by selecting *m* preliminary samples, each of size *n*. If there are D_i defectives in sample *i*, the fraction nonconforming in sample *i* is

And the average of these individuals sample fractions is

[redacted]

The $\bar{p}$ is used instead of *p* in the control chart setup.

Generally a control chart for fraction of defectives is used when sample unit as a whole classified as goods or bads, i.e. defective or non-defective.

Construction of p charts: as in the case of $\bar{X}$ and R charts we take the sample number along the horizontal scale and the statistic 'p' along the vertical scale. The sample fraction defective p_1, p_2, $p_3 \ldots p_k$ are plotted against the corresponding sample numbers as

points (dots). The central line is drawn as a dark horizontal line at $\bar{p}$(grand) and UCLp and LCLp are plotted as dotted horizontal lines at the computed value given.

Remarks: 1. since p can not be negative, if LCLp computed from the above formula comes out to be negative then it is taken as zero.

2 Interpretation of p charts: p charts interpreted exactly similarly as an $\bar{X}$ or R charts. If all the sample points fall with in the control limits, the process is termed to be in statistical control. If one or more of the points go out sides the control limits, it is said to be out of control. If the point goes above the UCLp, it reflects the lack of the statistical control since it has changed for the worse. Such a point, known as *high spot*, indicates deterioration in the lot quality. Reasons for this should be traced and eliminated. If the point goes below the LCLp, it again exhibits lack of control as the process has changed for the better. Such a point, known as *low spot*, indicates improvement in the lot quality. Reasoning for this should be looked for and incorporated in to the process if possible.

Example of a p-chart

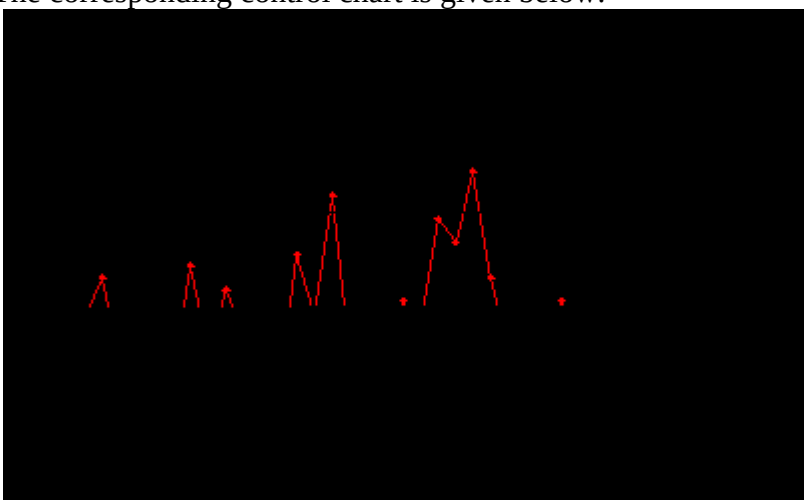
A numerical example will now be given to illustrate the above mentioned principles.
The location of chips on a wafer is measured on 30 wafers. On each wafer 50 chips are measured and a defective is defined whenever a miss registration, in terms of horizontal and/or vertical distances from the center, is recorded. The results are

Sample Number	Fraction Defectives	Sample Number	Fraction Defectives	Sample Number	Fraction Defectives
1	.24	11	.10	21	.40
2	.30	12	.12	22	.36
3	.16	13	.34	23	.48
4	.20	14	.24	24	.30
5	.08	15	.44	25	.18
6	.14	16	.16	26	.24
7	.32	17	.20	27	.14
8	.18	18	.10	28	.26
9	.28	19	.26	29	.18

10	.20	20	.22	30	.12

Sample proportions control chart

The corresponding control chart is given below:



Since from above charts two sample points are lies out of the control limit, indicates the process is out of control.

2.4.3 Control Charts for Number of Defectives (np or d charts)

If the sample size is constant for all samples, say n, then the sampling distribution of the statistic,

d= number of defective in the sample = np is given by,

E (d) =np

S.E (d) = $\sqrt{npQ}$

Hence, the 3-σ control limits for np charts are given by

E (d) ± S.E (d) = np ± $\sqrt{npQ}$

If $\hat{p}$ is the known or specified values of p then

UCL=n + $\sqrt{npQ}$

LCL = n - $\sqrt{npQ}$

CL = n $\hat{p}$

If p is not known, then its unbiased estimate provided by $\bar{p}$(grand) given in previous charts i.e. p –charts

UCLd== n $\bar{p}$(grand) +3

LCLd= n $\bar{p}$(grand) -3 $\sqrt{\bar{P}(grand)(1-\bar{P}(grand))}$

CLd= n $\bar{p}$(grand)

Remarks: np charts vs. p charts; if the sample size is varies from sample to sample, then np chart would be quite uncomfortable to use because the central lines as well as the control limits would vary from sample to sample. In such a case p chart would be better to use. However, in case of constant sample size for all samples any one of d or p charts may be used but, inpractice, p chart commonly used.

Example: *during an examination of equal length of cloth, the following are the number of defects is observed.*

Sample number	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Number of defectives	*2*	*3*	*4*	*0*	*5*	*6*	*7*	*4*	*3*	*2*

Draw a control chart for the number of defectives (np charts) and comment weather the process is under control or not?

Solution:- d =2+3+…+2=36

An estimate of the process of defective is given by p(grand)

N=1+2+3+…+10=55

P (grand) =36\55=0.65455

1- P (grand) =1-0.65455=0.345454

So, the 3-σ control limits for np charts are given by:

UCLd== n $\bar{p}$(grand) +3 $\sqrt{\bar{P}(grand)(1-\bar{P}(grand))}$

$= 55*0.6545+3\sqrt{55*0.6545*0.345454}$ **= *8.154***

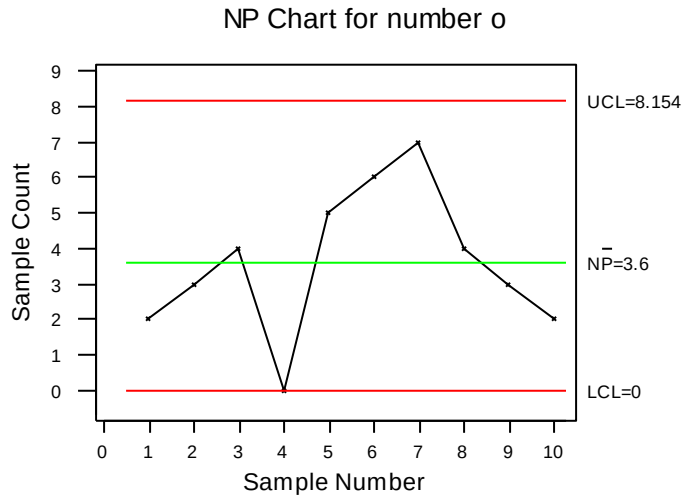
LCLd= n $\bar{p}$(grand) -3 $\sqrt{\bar{P}(grand)(1-\bar{P}(grand))}$

= *55*0.6545-3*$\sqrt{55*0.6545*0.345454}$ = ***-0.243*** since there is no negative number of defectives LCLd=0

CLd= n $\bar{p}$(grand)

= *55*0.654545*=***3.6*** since there are 10 samples per each defectives.

So, the control charts for the number of defective units is obtained in plotting the number of defectives against the corresponding sample number is given below.



Since one sample point corresponding to the 4th sample lies out side the control limits, the process is not in a state of statistical control.

2.5 Process and Measurement System Capability Analysis

Process capability compares the output of an *in-control* process to the specification limits by using *capability indices*. The comparison is made by forming the ratio of the spread between the process specifications (the specification "width") to the spread of the process values, as measured by 6 process standard deviation units (the process "width").

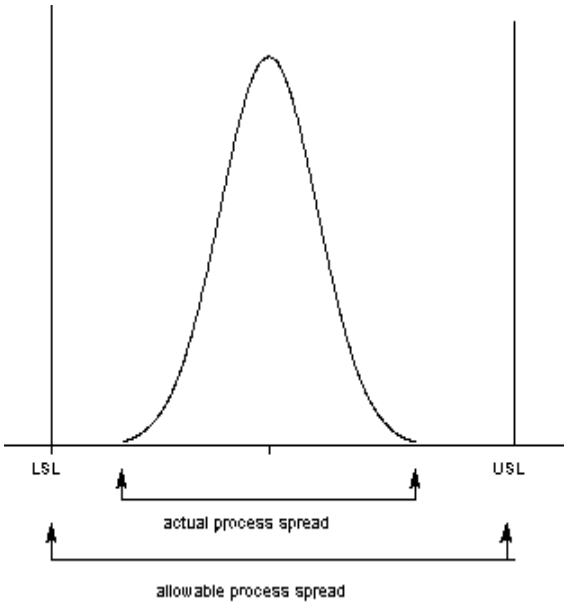
A process capability index uses both the process variability and the process specifications to determine whether the process is "capable"

Process Capability Indices

We are often required to compare the output of a stable process with the process specifications and make a statement about how well the process meets specification.
To do this we compare the natural variability of a stable process with the process specification limits.

A capable process is one where almost all the measurements fall inside the specification limits.

This can be represented pictorially by the plot below:



There are several statistics that can be used to measure the capability of a process: C_p, C_{pk}, C_{pm}.

Most capability indices estimates are valid only if the sample size used is 'large enough'.

Large enough is generally thought to be about 50 independent data values.

The C_p, C_{pk}, *and* C_{pm} statistics assume that the population of data values is normally distributed.

Assuming a two-sided specification, if and are the mean and standard deviation, respectively, of the normal data and USL, LSL, and T are the upper and lower specification limits and the target value, respectively, then the population capability indices are defined as follows:

Definitions of various process capability indices

Sample estimates of capability indices

Sample estimators for these indices are given below. (Estimators are indicated with a "hat" over them).

The estimator for C_{pk} can also be expressed as $C_{pk} = C_p(1-k)$, where k is a scaled distance between the midpoint of the specification range, m, and the process mean, .

Denote the midpoint of the specification range by m = process mean,

(The absolute sign takes care of the case when). To determine the estimated value, , we estimate by . Note that .

The estimator for the C_p index, adjusted by the k factor, is

Since , it follows that .

Plot showing C_p for
To get an idea of the value of the C_p statistic for varying process widths, consider the following plot

This can be expressed numerically by the table below:

Translating capability into "rejects"	*USL - LSL*	6	8	10	12

Where ppm = parts per million and ppb = parts per billion. Note that the reject figures are based on the assumption that the distribution is centered at μ.

We have discussed the situation with two spec. limits, the USL and LSL. This is known as the *bilateral* or two-sided case. There are many cases where only the lower or upper specifications are used. Using one spec limit is called *unilateral* or one-sided. The corresponding capability indices are

One-sided specifications and the corresponding capability indices

$$C_{pu} = \frac{\text{allowable upper spread}}{\text{actual upper spread}} = \frac{USL - \mu}{3\sigma}$$

and

$$C_{pl} = \frac{\text{allowable lower spread}}{\text{actual lower spread}} = \frac{\mu - LSL}{3\sigma}$$

Where μ and σ are the process mean and standard deviation, respectively.

Estimators of C_{pu} and C_{pl} are obtained by replacing μ and σ by and s,

respectively.

The following relationship holds

$C_p = (C_{pu} + C_{pl})$ /2.
This can be represented pictorially by

Note that we also can write:

$C_{pk} = \min \{C_{pl}, C_{pu}\}$.

Confidence Limits for Capability Indices

Confidence intervals for indices

Assuming normally distributed process data, the distribution of the sample follows from a Chi-square distribution and and have distributions related to the non-central *t* distribution. Fortunately, approximate confidence limits related to the normal distribution have been derived. Various approximations to the distribution of have been proposed, including those given by Bissell (1990), and we will use a normal approximation.

The resulting formulas for confidence limits are given below:

100(1-α)% Confidence Limits for C_p

Where

.

= degrees of freedom

Confidence Intervals for C_{pu} and C_{pl}

Approximate 100(1-α)% confidence limits for C_{pu} with sample size n are:

With z denoting the percent point function of the standard normal distribution. If is not known, set it to.

Limits for C_{pl} are obtained by replacing by .

Confidence Interval for C_{pk}

Zhang et al. (1990) derived the exact variance for the estimator of C_{pk} as well as an approximation for large n. The reference paper is Zhang, Stenback and Wardrop (1990), "Interval Estimation of the process capability index", *Communications in Statistics: Theory and Methods*, 19(21), 4455-4470.

The variance is obtained as follows:

Let

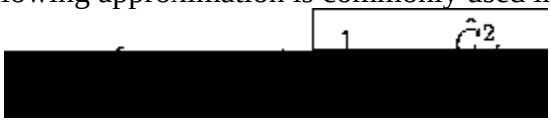
Then

$$= (d^2/36)(n-1)(n-3)$$

Their approximation is given by:

where

The following approximation is commonly used in practice



It is important to note that the sample size should be at least 25 before these approximations are valid. In general, however, we need $n \geq 100$ for capability studies. Another point to observe is that variations are not negligible due to the randomness of capability indices.

Capability Index Example

For a certain process the USL = 20 and the LSL = 8. The observed process average, $\overline{X}$= 16, and the standard deviation, *s* = 2. From this we obtain

$$\hat{C}_p = \frac{USL - LSL}{6s} = \frac{20-8}{6(2)} = 1.0$$

This means that the process is capable as long as it is located at the midpoint, m = (USL + LSL)/2 = 14.

But it doesn't, since = 16. The factor is found by

and

We would like to have at least 1.0, so this is not a good process. If possible, reduce the variability or/and center the process. We can compute the and

$$\hat{C}_{pu} = \frac{USL - \bar{x}}{3s} = \frac{20 - 16}{3(2)} = 0.6667$$

$$\hat{C}_{pl} = \frac{\bar{x} - LSL}{3s} = \frac{16 - 8}{3(2)} = 1.3333$$

From this we see that the , which is the smallest of the above indices, is 0.6667. Note that the formula is the algebraic equivalent of the *min* { , } definition.

What happens if the process is not approximately normally distributed?

What you can do with non-normal data

The indices that we considered thus far are based on normality of the process distribution. This poses a problem when the process distribution is not normal. Without going into the specifics, we can list some remedies.

1. Transform the data so that they become approximately normal. A popular transformation is the Box-Cox transformation
2. Use or develop another set of indices, that apply to non normal distributions. One statistic is called C_{npk} (for non-parametric C_{pk}). Its estimator is calculated by

$$\hat{C}_{npk} = \min\left[\frac{USL - median}{p(.995) - median}, \frac{median - LSL}{median - p(.005)}\right]$$

Where p (0.995) is the 99.5th percentile of the data and p(.005) is the 0.5th percentile of the data.

There is, of course, much more that can be said about the case of non normal data. However, if a Box-Cox transformation can be successfully performed, one is encouraged to use it.

Exercise

1. Given below are the values of samples mean ($\bar{X}$) and the range (R) for ten samples of size 5 each.

Sample number	1	2	3	4	5	6	7	8	9	10
$\bar{X}$	43	49	37	44	45	37	51	46	43	47
R	5	6	7	7	7	4	8	6	4	6

Given ***(A_3 = 0.58 for n=5)***

Draw appropriate mean chart and range chart and comment on the state of control of the process.

2. Construct the control chart for mean and range for the following data on the bases of fuses, samples of 5 being taken every nour (each set of 5 has been arranged in ascending order of magnitude). Comment on weather the production seems to be under control, assuming that these are the first data.

42	42	19	36	42	51	60	18	15	69	64	61
65	45	24	54	51	74	60	20	30	109	90	78
75	68	80	69	57	75	72	27	39	113	93	94
78	72	81	77	59	78	95	42	62	118	109	109
87	90	81	84	78	132	138	60	84	153	112	136

3. The following data shows the values of sample mean $\bar{X}$ and the range **R** ten samples of 5 each. Calculate the value of central line and control limits for mean chart, and range chart and determine weather the process is in control.

Sample number	1	2	3	4	5	6	7	8	9	10
Mean	11.2	11.8	10.8	11.6	11.0	9.6	10.4	9.6	10.6	10.0
Range	7	4	8	5	7	4	8	4	7	9

(Conversion factor n=5 are $\mathbf{A_2}$=0.577, $\mathbf{D_3}$=0 and $\mathbf{D_4}$=2.115)

4. The following data gives readings for 10 samples of size 6 in the production of a certain component.

Sample	1	2	3	4	5	6	7	8	9	10
Mean	383	505	532	508	557	337	514	614	707	753
S .D	30.5	41.6	39.5	32.2	25.4	24.2	48.7	89	13.1	33.9
Range	95	128	100	91	68	65	148	28	37	80

Draw the control charts for $\overline{X}$, σ and **R.** Calculate the limits of $\overline{X}$, in two ways. Can within

8. The average number of defectives in 22 sampled lots of 2000 rubber belts each was found to be 10%. Draw an appropriate control charts.

9. A TV voltage stabilizer manufacturer cheeks the quality of 50 units of his product daily for 15 days and finds the fraction of non conforming units and number of defectives as follows.

Days	Fraction defectives	No. of defectives
1	0.10	5
2	0.20	10
3	0.06	3
4	0.04	2
5	0.16	8
6	0.02	1
7	0.08	4
8	0.06	3
9	0.02	1
10	0.16	8
11	0.12	6
12	0.14	7
13	0.08	4
14	0.10	5
15	0.06	3

(i) Construct 3-sigma trial control limit for fraction defectives.

(ii) Construct 3-sigma trial control limits for np chart.

10. A manufacture of a soft drink uses control charts to check the CO_2 contents of the drink. In a prerun of m= 26 samples of size n= 4 the following data of CO_2 content

Chapter 3

3. Other Statistical Process-Monitoring and Control Techniques

3.1 Cumulative Sum Control Charts (CUSUM control charts)

CUSUM is an efficient alternative to Shewhart procedures

CUSUM charts, while not as intuitive and simple to operate as Shewhart charts, have been shown to be more efficient in detecting small shifts in the mean of a process. In particular, analyzing ARL's for CUSUM control charts shows that they are better than Shewhart control charts when it is desired to detect shifts in the mean that are 2 sigma or less.

CUSUM works as follows: Let us collect *k* samples, each of size *n*, and compute the mean of each sample. Then the cumulative sum (CUSUM) control chart is formed by plotting one of the following quantities

Definition of cumulative sum

$$S_m = \sum_{i=1}^{m} (\bar{x}_i - \hat{\mu}_0) \quad \text{or} \quad S'_m = \frac{1}{\sigma_{\bar{x}}} \sum_{i=1}^{m} (\bar{x}_i - \hat{\mu}_0)$$

Against the sample number *m*, where $\hat{\mu}_0$ is the estimate of the in-control mean and $\sigma_{\bar{x}}$ is the known (or estimated) standard deviation of the sample means. The choice of which of these two quantities is plotted is usually determined by the statistical software package. In either case, as long as the process remains in control centered at $\hat{\mu}_0$, the cusum plot will show variation in a random pattern centered about zero. If the process mean shifts upward, the charted cusum points will eventually drift upwards, and vice versa if the process mean decreases.

V-Mask used to determine if process is out of control

A visual procedure proposed by Barnard in 1959, known as *the V-Mask*, is sometimes used to determine whether a process is out of control. More often, the tabular form of the *V*-Mask is preferred. The tabular form is illustrated later in this section.

\

A V-Mask is an overlay

is more common in the literature, θ= 1/2 the vertex angle) as the design parameters, and we would end up with the same *V*-Mask.

In practice, designing and manually constructing a *V*-Mask is a complicated procedure. A cusum spreadsheet style procedure shown below is more practical, unless you have statistical software that automates the *V*-Mask methodology. Before describing the spreadsheet approach, we will look briefly at an example of a software *V*-Mask.

JMP example of V-Mask

JMP example of V-Mask

An example will be used to illustrate how to construct and apply a *V*-Mask procedure using JMP. The 20 data points

324.925, 324.675, 324.725, 324.350, 325.350, 325.225,
324.125, 324.525, 325.225, 324.600, 324.625, 325.150,
328.325, 327.250, 327.825, 328.500, 326.675, 327.775,
326.875, 328.350

are each the average of samples of size 4 taken from a process that has an estimated mean of 325. Based on process data, the process standard deviation is 1.27 and therefore the sample means used in the cusum procedure have a standard deviation of $1.27/4^{1/2} = 0.635$.

After inputting the 20 sample means and selecting "control charts" from the pull down "Graph" menu, JMP displays a "Control Charts" screen and a "CUSUM Charts" screen. Since each sample mean is a separate "data point", we choose a constant sample size of 1. We also choose the option for a two sided Cusum plot shown in terms of the original data.

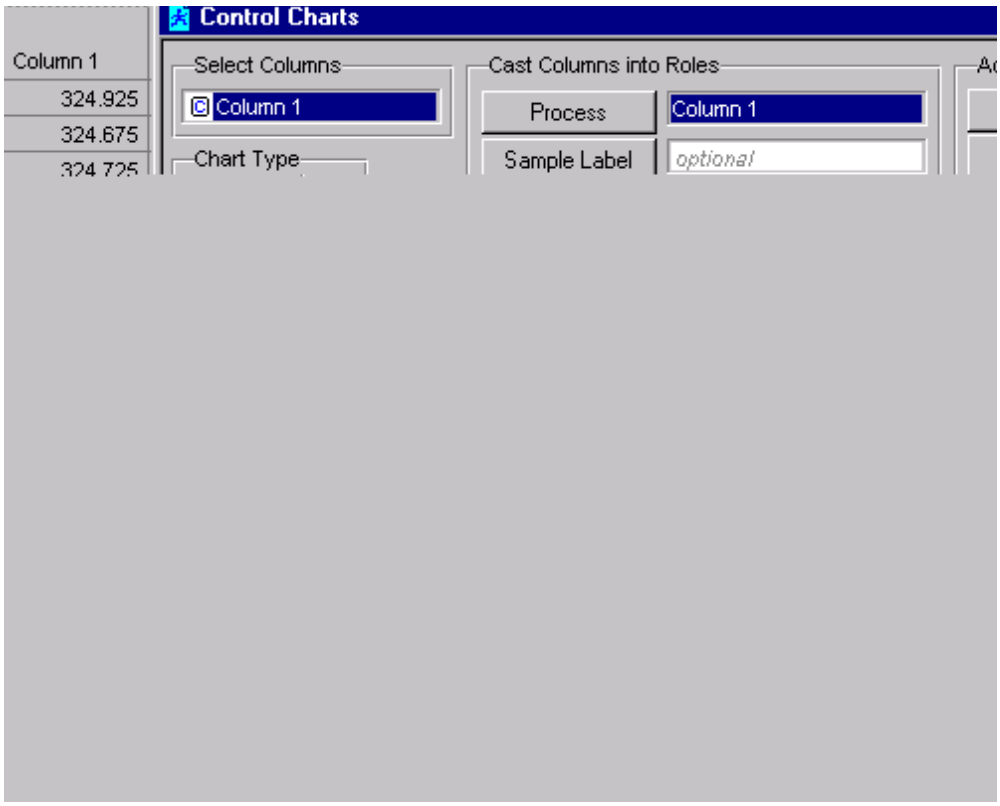
JMP allows us a choice of either designing via the method using *h* and *k* or using an *alpha* and *beta* design approach. For the latter approach we must specify

- α, the probability of a false alarm, i.e., concluding that a shift in the process has occurred, while in fact it did not
- β, the probability of not detecting that a shift in the process mean has, in fact, occurred
- δ(Delta), the amount of shift in the process means that we wish to detect, expressed as a multiple of the standard deviation of the data points (which are the sample means).

Note: Technically, alpha and beta are calculated in terms of one sequential trial where we monitor S_m until we have either an out-of-control signal or S_m returns to the starting point (and the monitoring begins, in effect, all over again).

JMP menus for inputting options to the cusum procedure

In our example we choose an α of 0.0027 (equivalent to the plus or minus 3 sigma criteria used in a standard Shewhart chart), and a β of 0.01. Finally, we decide we want to quickly detect a shift as large as 1 sigma, which sets delta = 1. The screen below shows all the inputs.

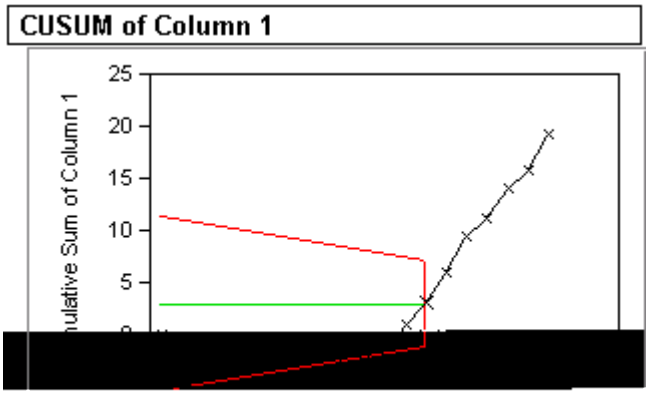


JMP output from CUSUM procedure

When we click on chart we see the *V*-Mask placed over the last data point. The mask clearly indicates an out of control situation.

We next "grab" the *V*-Mask and move it back to the first point that indicated the process was out of control. This is point number 14, as shown below.

JMP CUSUM chart after moving V-Mask to first out of control point



Rule of thumb for choosing h and k

Note: A general rule of thumb (Montgomery) if one chooses to design with the *h* and *k* approach, instead of the alpha and beta method illustrated above, is to choose *k* to be half the delta shift (.5 in our example) and *h* to be around 4 or 5.

Tabular or Spreadsheet Form of the *V*-Mask

A spreadsheet approach to cusum monitoring

Most users of cusum procedures prefer tabular charts over the *V*-Mask. The *V*-Mask is actually a carry-over of the pre-computer era. The tabular method can be quickly implemented by standard spreadsheet software.

To generate the tabular form we use the *h* and *k* parameters expressed in the original data units. It is also possible to use sigma units.

The following quantities are calculated:

$$S_{hi}(i) = \max(0, S_{hi}(i\text{-}1) + x_i - \hat{\mu}_0 - k)$$

$$S_{lo}(i) = \max(0, S_{lo}(i\text{-}1) + \hat{\mu}_0 - k - x_i))$$

where $S_{hi}(0)$ and $S_{lo}(0)$ are 0. When either S_{hi}(i) or S_{lo}(i) exceeds h, the process is out of control.

Example of spreadsheet calculation

Construct a cusum tabular chart for the example described above. For this example, the JMP parameter table gave *h* = 4.1959 and *k* = .3175. Using these design values, the tabular form of the example is

$\hat{\mu}_0$	*h*	*k*
325	4.1959	0.3175

We will construct a cusum tabular chart for the example described above. For this example, the JMP parameter table gave *h* = 4.1959 and *k* = .3175. Using these design values, the tabular form of the example is

$\hat{\mu}_o$	h	k
325	4.1959	0.3175

Group	x	x-325	Increase in mean x-325-k	S_{hi}	Decrease in mean 325-k-x	S_{lo}	Cusum
1	324.93	0.07	-0.39	0.00	-0.24	0.00	-0.007
2	324.68	0.32	-0.64	0.00	0.01	0.01	-0.40
3	324.73	0.27	-0.59	0.00	-0.04	0.00	-0.67
4	324.35	0.65	-0.97	0.00	0.33	0.33	-1.32
5	325.35	0.35	0.03	0.03	-0.67	0.00	-0.97
6	325.23	0.23	-0.09	0.00	-0.54	0.00	-0.75
7	324.13	0.88	-1.19	0.00	0.56	0.56	-1.62
8	324.53	0.48	-0.79	0.00	0.16	0.72	-2.10
9	325.23	0.23	-0.09	0.00	0.54	0.17	-1.87
10	324.60	0.40	-0.72	0.00	0.08	0.25	-2.27
11	324.63	0.38	-0.69	0.00	0.06	0.31	-2.65
12	325.15	0.15	-0.17	0.00	0.47	0.00	-2.50
13	328.33	3.32	3.01	3.01	-3.64	0.00	0.83
14	327.25	2.25	1.93	4.94*	-0.57	0.00	3.08
15	327.83	2.82	2.51	7.45*	-3.14	0.00	5.90
16	328.50	3.50	3.18	10.63*	-3.82	0.00	9.40
17	326.68	1.68	1.36	11.99*	-1.99	0.00	11.08
18	327.78	2.77	2.46	14.44*	-3.09	0.00	13.85
19	326.88	1.88	1.56	16.00*	-2.19	0.00	15.73
20	328.35	3.35	3.03	19.04*	-3.67	0.00	19.08

* = out of control signal

3.2.1 Cumulative Sum (CUSUM) Average Run Length

Time to Detection or Average Run Length (ARL)

Waiting time to signal "out of control"

Two important questions when dealing with control charts are:

3. How often will there be false alarms where we look for an assignable cause but nothing has changed?
4. How quickly will we detect certain kinds of systematic changes, such as mean shifts?

The ARL tells us, for a given situation, how long on the average we will plot successive control charts points before we detect a point beyond the control limits.

For an $\bar{X}$chart, with no change in the process, we wait on the average 1/*p* points before a false alarm takes place, with *p* denoting the probability of an observation plotting outside the control limits. For a normal distribution, *p* = .0027 and the ARL is approximately 371.

A table comparing Shewhart $\bar{X}$chart *ARL*'s to Cumulative Sum (CUSUM) *ARL*'s for various mean shifts is given later in this section.

The Average Run Length of Cumulative Sum Control Charts

The ARL of CUSUM

The operation of obtaining samples to use with a cumulative sum (CUSUM) control chart consists of taking samples of size *n* and plotting the cumulative sums

$$S_r = \sum_{i=1}^{r} (\bar{x}_i - k) \quad or \quad S_r = \sum_{i=1}^{r} (\bar{x}_i - k)/\sigma_{\bar{x}} \quad (\text{standardized})$$

Versus the sample number *r*, where $\bar{x}_i$the sample is mean and *k* is a reference value.

In practice, k might be set equal to (+ $_1$)/2, where is the estimated in-control mean, which is sometimes known as the *acceptable quality level*, and $_1$ is referred to as the

in control), and a low ARL, L_1, when the process mean shifts to an unsatisfactory level.

In order to determine the parameters of a CUSUM chart, the acceptable and reject able quality levels along with the desired respective ARL ' s are usually specified. The design parameters can then be obtained by a number of ways. Unfortunately, the calculations of the ARL for CUSUM charts are quite involved.

There are several nomographs available from different sources that can be utilized to find the ARL's when the standardized *h* and *k* are given. Some of the nomographs solve the unpleasant integral equations that form the basis of the exact solutions, using an approximation of Systems of Linear Algebraic Equations (SLAE). This Handbook used a computer program that furnished the required ARL's given the standardized *h* and *k*. An example is given below:

Example of finding ARL's given the standardized h and k

mean shift (k = .5)	$h\sqrt{n}/\sigma$ 4	5	Shewart $\bar{X}$
0	336	930	371.00
.25	74.2	140	281.14
.5	26.6	30.0	155.22
.75	13.3	17.0	81.22
1.00	8.38	10.4	44.0
1.50	4.75	5.75	14.97
2.00	3.34	4.01	6.30
2.50	2.62	3.11	3.24
3.00	2.19	2.57	2.00
4.00	1.71	2.01	1.19

Using the table

If k = .5, then the shift of the mean (in multiples of the standard deviation of the mean) is obtained by adding .5 to the first column. For example to detect a mean shift of 1 sigma at $h = 4$, the ARL = 8.38. (at first column entry of .5).

The last column of the table contains the ARL's for a Shewhart control chart at selected mean shifts. The ARL for Shewhart = $1/p$, where p is the probability for a point to fall outside established control limits. Thus, for 3-sigma control limits and assuming normality, the probability to exceed the upper control limit = .00135 and to fall below the lower control limit is also .00135 and their sum = .0027. (These numbers come from standard normal distribution tables or computer programs, setting $z = 3$). Then the ARL = 1/.0027 = 370.37. This says that when a process is in control one expects an out-of-control signal (false alarm) each 371 runs.

ARL if a 1 sigma shift has occurred

When the means shifts up by 1 sigma, then the distance between the upper control limit and the shifted mean is 2 sigma (instead of 3 σ). Entering normal distribution tables with $z = 2$ yields a probability of p = .02275 to exceed this value. The distance between the shifted mean and the lower limit is now 4 sigma and the probability of $\bar{X} < -4$ is only .000032 and can be ignored. The ARL is 1 / .02275 = 43.96.

Shewhart is better for detecting large shifts, CUSUM is faster for small shifts

The conclusion can be drawn that the Shewhart chart is superior for detecting large shifts and the CUSUM scheme is faster for small shifts. The break-even point is a function of h, as the table shows.

3.3 Exponentially Weighted Moving Average Control Charts (EWMA)

EWMA statistic

The Exponentially Weighted Moving Average (EWMA) is a statistic for monitoring the process that averages the data in a way that gives less and less weight to data as they are further removed in time.

Comparison of Shewhart control chart and EWMA control chart techniques

For the Shewhart chart control technique, the decision regarding the state of control of the process at any time, *t,* depends solely on the most recent measurement from the process and, of course, the degree of 'trueness' of the estimates of the control limits from historical data. For the EWMA control technique, the decision depends on the EWMA statistic, which is an exponentially weighted average of all prior data, including the most recent measurement.

By the choice of weighting factor, , the EWMA control procedure can be made sensitive to a small or gradual drift in the process, whereas the Shewhart control procedure can only react when the last data point is outside a control limit.

Definition of EWMA

The statistic that is calculated is:

$$EWMA_t = Y_t + (1-) EWMA_{t-1} \quad \text{for } t = 1, 2, ..., n.$$

Where

- $EWMA_0$ is the mean of historical data (target)
- Y_t is the observation at time *t*
- *n* is the number of observations to be monitored including $EWMA_0$
- $0 < \leq 1$ is a constant that determines the depth of memory of the EWMA.

The equation is due to Roberts (1959).

Choice of weighting factor

The parameter determines the rate at which 'older' data enter into the calculation of the EWMA statistic. A value of = 1 implies that only the most recent measurement influences the EWMA (degrades to Shewhart chart). Thus, a large value of = 1 gives more weight to recent data and less weight to older data; a small value of gives more weight to older data. The value of is usually set between 0.2 and 0.3 (Hunter) although this choice is somewhat arbitrary. Lucas and Saccucci (1990) give tables that help the user select .

Variance of EWMA statistic

The estimated variance of the EWMA statistic is approximately

$$s^2_{ewma} = (\ \ / (2- \ \ )) \ s^2$$

when *t* is not small, where *s* is the standard deviation calculated from the historical data.

Definition of control limits for EWMA

The center line for the control chart is the target value or $EWMA_0$. The control limits are:

$$UCL = EWMA_0 + ks_{ewma}$$

$$LCL = EWMA_0 - ks_{ewma}$$

where the factor *k* is either set equal 3 or chosen using the Lucas and Saccucci (1990) tables. The data are assumed to be independent and these tables also assume a normal population.

As with all control procedures, the EWMA procedure depends on a database of measurements that are truly representative of the process. Once the mean value and standard deviation have been calculated from this database, the process can enter the monitoring stage, provided the process was in control when the data were collected. If not, then the usual Phase 1 work would have to be completed first.

Example of calculation of parameters for an EWMA control chart

To illustrate the construction of an EWMA control chart, consider a process with the following parameters calculated from historical data:

$EWMA_0 = 50$

$s = 2.0539$

with λ chosen to be 0.3 so that $\lambda/(2-\lambda) = .3 / 1.7 = 0.1765$ and the square root = 0.4201.
The control limits are given by

$UCL = 50 + 3\,(0.4201)(2.0539) = 52.5884$

$LCL = 50 - 3\,(0.4201)\,(2.0539) = 47.4115$

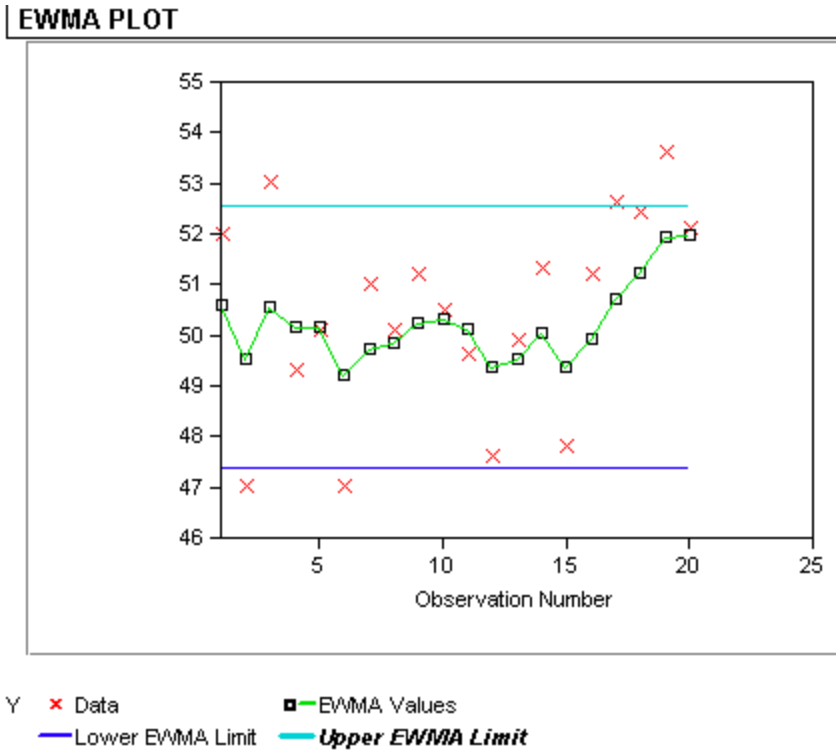
Sample data Consider the following data consisting of 20 points where 1-10 are on the top row from left to right and 11-20 are on the bottom row from left to right:

52.0 47.0 53.0 49.3 50.1 47.0
51.0 50.1 51.2 50.5 49.6 47.6
49.9 51.3 47.8 51.2 52.6 52.4
53.6 52.1

EWMA statistics for sample data

These data represent control measurements from the process which is to be monitored using the EWMA control chart technique. The corresponding EWMA statistics that are computed from this data set are:

50.00 50.60 49.52 50.56 50.18
50.16 49.12 49.75 49.85 50.26
50.33 50.11 49.36 49.52 50.05
49.34 49.92 50.73 51.23 51.94



Interpretation of EWMA control chart

The red dots are the raw data; the jagged line is the EWMA statistic over time. The chart tells us that the process is in control because all $EWMA_t$ lie between the control limits. However, there seems to be a trend upwards for the last 5 periods.

3.4 Other Univariate Statistical Process Monitoring And Control Charts

Individual control charts

Samples are Individual Measurements

Moving range used to derive upper and lower limits

Control charts for individual measurements, e.g., the sample size = 1, use the *moving range* of two successive observations to measure the process variability.

The moving range is defined as

$$MR_i = |x_i - x_{i-1}|$$

which is the absolute value of the first difference (e.g., the difference between two consecutive data points) of the data. Analogous to the Shewhart control chart, one can plot both the data (which are the individuals) and the moving range.

Individuals control limits for an observation

For the control chart for individual measurements, the lines plotted are:

$$UCL = \bar{x} + 3\frac{\overline{MR}}{1.128}$$

$$\text{Center Line} = \bar{x}$$

$$LCL = \bar{x} - 3\frac{\overline{MR}}{1.128}$$

where $\overline{X}$ is the average of all the individuals and $\overline{MR}$ is the average of all the moving ranges of two observations. Keep in mind that either or both averages may be replaced by a standard or target, if available. (Note that 1.128 is the value of d_2 for $n = 2$).

Example: The following example illustrates the control chart for individual observations. A new process was studied in order to monitor flow rate. The first 10 batches resulted in the following table.

Batch Number	Flowrate x	Moving Range MR
1	49.6	
2	47.6	2.0
3	49.9	2.3
4	51.3	14
5	47.8	3.5
6	51.2	3.4
7	52.6	1.4
8	52.4	0.2
9	53.6	1.2
10	52.1	1.5
	$\overline{X}$= 50.81	$\overline{MR}$= 1.8778

Limits for the moving range chart

This yields the parameters below.

$$UCL = \bar{x} + 3\frac{\overline{MR}}{1.128} = 50.81 + 3\frac{1.8778}{1.128} = 55.8041$$

$$\text{Center Line} = \bar{x} = 50.81$$

Example of individuals chart

The control chart is given below

The process is in control, since none of the plotted points fall outside either the *UCL* or *LCL*.

Alternative for constructing individuals control chart

Note: Another way to construct the individuals chart is by using the standard deviation. Then we can obtain the chart from

$$\bar{x} \pm 3s/c_4$$

It is preferable to have the limits computed this way for the start of Phase 2.

3.5 Multivariate Control Charts

Multivariate control charts and Hotelling's T^2 statistic

It is a fact of life that most data are naturally multivariate. Hotelling in 1947 introduced a statistic which uniquely lends itself to plotting multivariate observations. This statistic, appropriately named Hotelling's T^2, is a scalar that combines information from the dispersion and mean of several variables. Due to the fact that computations are laborious and fairly complex and require some knowledge of matrix algebra, acceptance of multivariate control charts by industry was slow and hesitant.

Multivariate control charts now more accessible

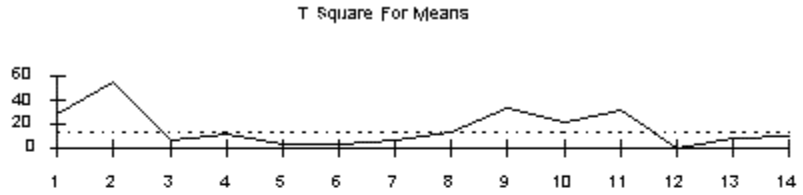
Nowadays, modern computers in general and the PC in particular have made complex calculations accessible and during the last decade, multivariate control charts were given more attention. In fact, the multivariate charts which display the Hotelling T^2 statistic became so popular that they sometimes are called Shewhart charts as well (e.g., Crosier, 1988), although Shewhart had nothing to do with them.

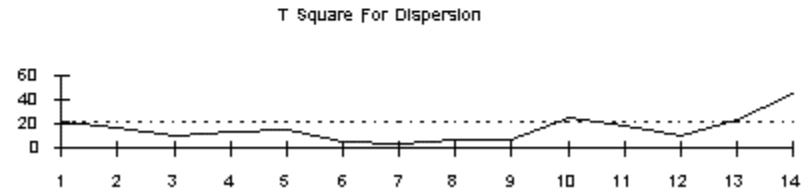
Hotelling charts for both means and dispersion

As in the univariate case, when data are grouped, the T^2 chart can be paired with a chart that displays a measure of variability within the subgroups for all the analyzed characteristics. The combined T^2 and T_d^2(dispersion) charts are thus a multivariate counterpart of the univariate $\overline{X}$and *S* (or $\overline{X}$and *R*) charts.

Hotelling mean and dispersion control

An example of a Hotelling T^2 and T_d^2 pair of charts is given below charts.





Interpretation of sample Hotelling control charts*:* Each chart represents 14 consecutive measurements on the means of four variables. The T^2 chart for means indicates an out-of-control state for groups 1, 2 and 9-11. The T^2_d chart for dispersions indicate that groups 10, 13 and 14 are also out of control. The interpretation is that the multivariate system is suspect. To find an assignable cause, one has to resort to the individual univariate control charts or some other univariate procedure that should accompany this multivariate chart.

3.5.1 Elements of Multivariate Analysis

Multivariate analysis

Multivariate analysis is a branch of statistics concerned with the analysis of multiple measurements, made on one or several samples of individuals. For example, we may wish to measure length, width and weight of a product.

Multiple measurement, or observation, as row or column vector

A multiple measurement or observation may be expressed as

x = **[4 2 0.6]**

referring to the physical properties of length, width and weight, respectively. It is customary to denote multivariate quantities with bold letters. T6ly. It is

3.5.2 Hotelling Control Charts

Definition of Hotelling's T^2 "distance" statistic

The Hotelling T^2 distances is a measure that accounts for the covariance structure of a multivariate normal distribution. It was proposed by Harold Hotelling in 1947 and is called Hotelling T^2. It may be thought of as the multivariate counterpart of the Student's-*t* statistic.

The T^2 distance is a constant multiplied by a quadratic form. This quadratic form is obtained by multiplying the following three quantities:

1. The vector of deviations between the observations and the mean **m,** which is expressed by **(X-m)'**,
2. The inverse of the covariance matrix, $\mathbf{S}^{-1}$,
3. The vector of deviations, **(X-m**).

It should be mentioned that for independent variables, the covariance matrix is a diagonal matrix and T^2 becomes proportional to the sum of squared standardized variables.

In general, the higher the T^2 value, the more distant is the observation from the mean. The formula for computing the T^2 is:

$$T^2 = c\left(\mathbf{X}-\mathbf{m}'\right)\mathbf{S}^{-1}\left(\mathbf{X}-\mathbf{m}\right)$$

The constant *c* is the sample size from which the covariance matrix was estimated.

T^2 readily graph able: The T^2 distances lend themselves readily to graphical displays and as a result the T^2-chart is the most popular among the multivariate control charts.

Estimation of the Mean and Covariance Matrix

Mean and Covariance matrices

Let $\mathbf{X}_1,...\mathbf{X}_n$ be n p-dimensional vectors of observations that are sampled independently from $N_p(\mathbf{m}, \Sigma)$ with $p < n$-1, with Σ the covariance matrix of $\mathbf{X}$. The observed mean vector $\bar{\mathbf{X}}$ and the sample dispersion matrix

$$\mathbf{S} = \frac{1}{n-1}\sum_{i=1}^{n}\left(\mathbf{X_i} - \bar{\mathbf{X}}\right)\left(\mathbf{X_i} - \bar{\mathbf{X}}\right)'$$

Are the unbiased estimators of $\mathbf{m}$ and Σ, respectively.

3.5.3 Principal component control charts

Problems with T^2 charts: Although the T^2 chart is the most popular, easiest to use and interpret method for handling multivariate process data, and is beginning to be widely accepted by quality engineers and operators, it is not a panacea. First, unlike the univariate case, the scale of the values displayed on the chart is not related to the scales of any of the monitored variables. Secondly, when the T^2 statistic exceeds the upper control limit (UCL), the user does not know which particular variable(s) caused the out-of-control signal.

Run univariate charts along with the multivariate ones

With respect to scaling, we strongly advise to run individual univariate charts in tandem with the multivariate chart. This will also help in honing in on the culprit(s) that might have caused the signal. However, individual univariate charts cannot explain situations that are a result of some problems in the covariance or correlation between the variables. This is why a dispersion chart must also be used.

Another way to monitor multivariate data: Principal Components control charts

Another way to analyze the data is to use *principal components*. For each multivariate measurement (or observation), the principal components are linear combinations of the standardized *p* variables (to standardize subtract their respective targets and divide by their standard deviations). The principal components have two important advantages:

1. The new variables are uncorrelated (or almost)
2. Very often, a few (sometimes 1 or 2) principal components may capture most of the variability in the data so that we do not have to use all of the *p* principal components for control.

Eigenvalues: Unfortunately, there is one big disadvantage: The identity of the original variables is lost! However, in some cases the specific linear combinations corresponding to the principal components with the largest *eigenvalues* may yield meaningful measurement units. What is being used in control charts are the *principal factors.*

A principal factor is the principal component divided by the square root of its eigenvalue.

3.5.4 Multivariate Exponentially Weighted Moving Average (EWMA) Charts

Multivariate EWMA Control Chart

Univariate EWMA model: - The model for a univariate EWMA chart is given by:

$$Z_i = \lambda X_i + (1 - \lambda) Z_{i-1} \quad i = 1, 2, ..., n$$

where Z_i is the ith EWMA, X_i is the ith observation, Z_0 is the average from the historical data, and $0 < \leq 1$.

Multivariate EWMA model:- In the multivariate case, one can extend this formula to

$$Z_i = \Lambda X_i + (1 - \Lambda) Z_{i-1}$$

where Z_i is the *i*th EWMA vector, X_i is the the *i*th observation vector $i = 1, 2, ..., n$, Z_0 is the vector of variable values from the historical data, Λ is the diag ($_1$, $_2$, ... , $_p$) which is a diagonal matrix with $_1$, $_2$, ... , $_p$ on the main diagonal, and *p* is the number of variables; that is the number of elements in each vector.

Illustration of multivariate EWMA:- The following illustration may clarify this. There are *p* variables and each variable contains *n* observations. The input data matrix looks like:

The quantity to be plotted on the control chart is

Simplification:- It has been shown (Lowry et al., 1992) that the (*k*, *l*)th element of the covariance matrix of the *i*th EWMA, , is

where is the (*k*,*l*)th element of

Table for selected values of and i

The following table gives the values of (1-) ^{2i} for selected values of and *i*.

	2*i*								
1 -	**4**	**6**	**8**	**10**	**12**	**20**	**30**	**40**	**50**
.9	.656	.531	.430	.349	.282	.122	.042	.015	.005
.8	.410	.262	.168	.107	.069	.012	.001	.000	.000
.7	.240	.118	.058	.028	.014	.001	.000	.000	.000
.6	.130	.047	.017	.006	.002	.000	.000	.000	.000
.5	.063	.016	.004	.001	.000	.000	.000	.000	.000
.4	.026	.004	.001	.000	.000	.000	.000	.000	.000
.3	.008	.001	.000	.000	.000	.000	.000	.000	.000
.2	.002	.000	.000	.000	.000	.000	.000	.000	.000
.1	.000	.000	.000	.000	.000	.000	.000	.000	.000

Specified formula not required:-

It should be pointed out that a well-meaning computer program does not have to adhere to the simplified formula, and potential inaccuracies for low values for and *i* can thus be avoided.

MEWMA computer out put for Lowry data:-

Here is an example of the application of an MEWMA control chart. To facilitate comparison with existing literature, we used data from Lowry et al. The data were simulated from a bivariate normal distribution with unit variances and a correlation coefficient of 0.5. The value for = .10 and the values for T_i^2 were obtained by the equation given above. The covariance of the MEWMA vectors was obtained by using the non-simplified equation. That means that for each MEWMA control statistic, the computer computed a covariance matrix, where *i* = 1, 2, ...10. The results of the computer routine are:

********** **********
* Multi - Variate EWMA Control Chart *
****

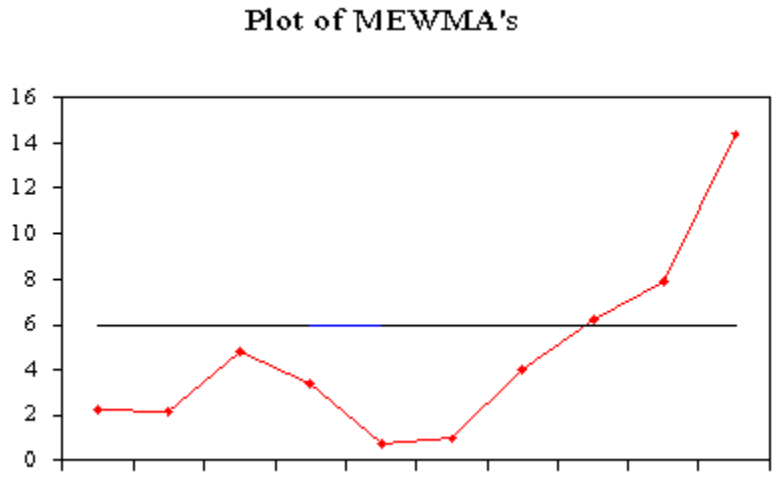
DATA SERIES 1	2	MEWMA Vector 1	2	MEWMA STATISTIC	
- 1.190	0.590	- 0.119	0.059	2.1886	
0.120	0.900	- 0.095	0.143	2.0697	
- 1.690	0.400	- 0.255	0.169	4.8365	
0.300	0.460	- 0.199	0.198	3.4158	
0.890	- 0.750	- 0.090	0.103	0.7089	
0.82 0	0.980	0.001	0.191	0.9268	
- 0.300	2.280	- 0.029	0.400	4.0018	
0.630	1.750	0.037	0.535	6.1657	
1.560	1.580	0.189	0.639	7.8554	
1.460	3.050	0.316	0.880	14.4158	

VEC	XBAR	MSE	Lamda
1	.260	1.200	0.100
2	1.124	1.774	0.100

The UCL = 5.938 for α= .05. Smaller choices of α are also used.

Sample MEWMA plots

The following is the plot of the above MEWMA



Chapter 4

4 Acceptance Sampling

4.1 Introduction

So far we confined our attention to '*process control* 'which is achieved through the powerful tool of shewhart's control charts. In this case producer had the complete control of the process which made the product, but when it comes to the marketing of the product, the problem is different. Here we have to take in to consideration the requirements of the customers and the firms or companies who receive the end products from the process. For this we need that it termed as '*product control'* in which the producer wants to ensure himself that the manufactured goods are according to the specifications of the customers or the receiving firms or companies and do not contain a large number of defectives. For this, it is imperative that he should have his product examined at strategic points, which is designed as receiving inspection. The inspection may be on sample basis or census basis. The basic tool used for the examination of the finished product is called *acceptance sampling* in which the decision to accept or reject a lot is made on the basis of random samples draw from it. Acceptance sampling plan invented by Dodge, H.F.and Roming, H.G. are powerful tools for product control. They basically designed these plans for lot quality protection so that a lot as poor as a given quality has a small chance of being accepted. They aimed at:

(i) Providing a definite assurance against accepting any unsatisfactory lot.

(ii) Minimizing the amount of inspection (and hence the inspection costs) subject to the degree of protection provide in the above (i).

Like control charts for variables and attributes we have sampling inspection plans for attributes and variables. The acceptance sampling plans for attributes are relatively easy to carry on and are described here briefly.

4.2 Concepts of Acceptance Sampling

It looks to the point to inspect each and every items produced by a manufacturing unit and make sure about the quality specification before releasing it for sale. But cent per cent inspection has its own weakness.

Firstly, due to fatigue of checking a large number of items, the efficency of inspection goes down and hence one can not expect that no defective or non-conforming item will not be left out after inspection.

Secondly, cent per cent inspection is impossibility in cases where the produce or items are perished under inspection such as inspection for life of electric bulbs, life of battery cells, combustibility of coke, etc.

Thirdly, the item and cost are other two important economic factors which discourage a manufacture from 100 percent inspections. So the acceptance or rejection of a lot is usually based on the inspection of sample drawn from the lots at regular intervals during the manufacturing process. Usually people call it *acceptance sampling plan*. Thus under acceptance sampling plan one takes the decision whether a lot is to be accepted or rejected. Broadly, three purposes are served by sampling inspection plans.

(i) It enables to know whether the process is producing the product which meets the quality specification or not.

(ii) It reveals whether the furnished product is good for marketing or not.

(iii) It minimizes the risk of the consumer and protects the producer from future losses.

In spite of many factors against cent per cent inspections, it is still preferred under special situations such as:

(i) A defective item may cause danger to life.

(ii) A defect may stop the whole function of the system.

(iii) The lot size is small

(iv) The income quality is very poor.

Before discussing the inspection plans we shall first explain the various terms and concepts.

Definitions of basic Acceptance Sampling terms

Deriving a plan, within one of the categories listed above, is discussed in the pages that follow. All derivations depend on the properties you want the plan to have. These are described using the following terms:

- ***Acceptable Quality Level (AQL)***: The AQL is a percent defective that is the base line requirement for the quality of the producer's product. The producer would like to design a sampling plan such that there is a *high probability of accepting* a lot that has a defect level less than or equal to the AQL.
- ***Lot Tolerance Percent Defective (LTPD)***: The LTPD is a designated high defect level that would be unacceptable to the consumer. The consumer would like the sampling plan to have a *low probability of accepting* a lot with a defect level as high as the LTPD.
- ***Type I Error (Producer's Risk)***: This is the probability, for a given (*n,c*) sampling plan, of rejecting a lot that has a defect level equal to the AQL. The producer suffers when this occurs, because a lot with acceptable quality was rejected. The symbol α is commonly used for the Type I error and typical values for α range from 0.2 to 0.01.
- ***Type II Error (Consumer's Risk):*** This is the probability, for a given (*n,c*) sampling plan, of accepting a lot with a defect level equal to the LTPD. The consumer suffers when this occurs, because a lot with unacceptable quality was accepted. The symbol β is commonly used for the Type II error and typical values range from 0.2 to 0.01.
- ***Operating Characteristic (OC) Curve:*** This curve plots the probability of accepting the lot (Y-axis) versus the lot fraction or percent defectives (X-axis). *The OC curve is the primary tool for displaying and investigating the properties of a LASP*.
- ***Average Outgoing Quality (AOQ):*** A common procedure when sampling and testing is non-destructive, is to 100% inspect rejected lots and replace all defectives with good units. In this case, all rejected lots are made perfect and the only defects left are those in lots that were accepted. *AOQ*'s refer to the long term

defect level for this combined LASP and 100% inspection of rejected lots process. If all lots come in with a defect level of exactly *p*, and the OC curve for the chosen (*n,c*) LASP indicates a probability p_a of accepting such a lot, over the long run the *AOQ* can easily be shown to be:

$$AOQ = \frac{p_a p(N-n)}{N}$$

Where *N* is the lot size.

- ***Average Outgoing Quality Level (AOQL):*** A plot of the *AOQ* (Y-axis) versus the incoming lot *p* (X-axis) will start at 0 for $p = 0$, and return to 0 for $p = 1$ (where every lot is 100% inspected and rectified). In between, it will rise to a maximum. This maximum, which is the worst possible long term *AOQ*, is called the *AOQL*.
- ***Average Total Inspection (ATI):*** When rejected lots are 100% inspected, it is easy to calculate the *ATI* if lots come consistently with a defect level of *p*. For a LASP (*n,c*) with a probability p_a of accepting a lot with defect level *p*, we have

  $$ATI = n + (1 - p_a)(N - n)$$

  Where *N* is the lot size.

- ***Average Sample Number (ASN):*** For a single sampling LASP (*n,c*) we know each and every lot has a sample of size *n* taken and inspected or tested. For double, multiple and sequential LASP's, the amount of sampling varies depending on the number of defects observed. For any given double, multiple or sequential plan, a long term *ASN* can be calculated assuming all lots come in with a defect level of *p*. A plot of the *ASN*, versus the incoming defect level *p*, describes the sampling efficiency of a given LASP scheme.

4.3 Lot By Lot Acceptance Sampling

The acceptance or rejection of a lot is based on inspections of a sample drawn from a submitted lot. Usually the submission based on the number or proportion of

defectives present in the sample according to attributes under consideration. But the item classified as defective and non- defective on the basis of measurable quality characteristics as well. For instance, an item heavier than a fixed weight or shorter than a fixed weight or shorter than a fixed length can be classified as defective.

LASP is a sampling scheme and a set of rules

A lot acceptance sampling plan (LASP) is a sampling scheme and a set of rules for making decisions. The decision, based on counting the number of defectives in a sample, can be to accept the lot, reject the lot, or even, for multiple or sequential sampling schemes, to take another sample and then repeat the decision process.

Types of acceptance plans to choose from

Lot acceptance sampling plans (LASPs) fall in one of the following:-

- **Single sampling plans:** One sample of items is selected at random from a lot and the disposition of the lot is determined from the resulting information. These plans are usually denoted as (*n*, *c*) plans for a sample size *n*, where the lot is rejected if there are more than *c* defectives. *These are the most common (and easiest) plans to use although not the most efficient in terms of average number of samples needed.*
- **Double sampling plans:** After the first sample is tested, there are three possibilities:
    1. Accept the lot
    2. Reject the lot
    3. No decision

    If the outcome is (3), and a second sample is taken, the procedure is to combine the results of both samples and make a final decision based on that information.

- **Multiple sampling plans:** This is an extension of the double sampling plans where more than two samples are needed to reach a conclusion. The advantage of multiple sampling is smaller sample sizes.

.

- Sequential sampling plans: This is the ultimate extension of multiple sampling where items are selected from a lot one at a time and after inspection of each item a decision is made to accept or reject the lot or select another unit.
    - **Skip lot sampling plans:** Skip lot sampling means that only a fraction of the submitted lots are inspected.

The final choice is a tradeoff decision: - Making a final choice between single or multiple sampling plans that have acceptable properties is a matter of deciding whether the average sampling savings gained by the various multiple sampling plans justifies the additional complexity of these plans and the uncertainty of not knowing how much sampling and inspection will be done on a day-by-day basis.

4.3.1 Single Sampling Plan

A sampling plan in which a decision about the acceptance or rejection of a lot is based on one of the sample that has been inspected. Suppose a lot consists of N items having a proportion of defectives as P and in all D defectives, i.e. D=NP. A sample of size n drawn from a submitted lot which contains d defectives. Let c be the maximum allowable number of defectives in the sample, c is known as *acceptance number*. The quantities n and c can either be determined by lot quality protection approach or by an average quality protection approach.

The procedure to arrive at a decision about the acceptance or rejection of a lot is as follows:-

(i) Select a random sample of size n from the submitted lot.

(ii) Inspect each and every unit included in the sample and classify them as defective or non-defective. Suppose the number of defectives in the sample is d. the criteria for the acceptance or rejection of lot is,

(a) If $d \leq c$, the lot is accepted. In this case all the defective items in the sample are replaced by non-defectives and the lot is release for marketing.

(b) If d>c, the lot is rejected. In this situation, inspect the whole lot and replace all defectives found in the lot by non defectives and release the lot for sale.

In single sampling plan, the variable d follows the hyper geometric distribution. Various probabilities under this plan can be worked out by the following formula.

1. The probability of obtaining d defectives in a sample of size n from the lot is,

$$P_d = \frac{\binom{NP}{d}\binom{N-NP}{n-d}}{\binom{N}{n}}$$

2. The probability of accepting the lot having not more than c defectives in the sample of n items is,

$$P_a = \frac{\sum_{d=0}^{c}\binom{NP}{d}\binom{N-NP}{n-d}}{\binom{N}{n}}$$

3. The consumer risk when the lot having the proportion defectives p_t accepted.

$$P_c = \frac{\sum_{d=0}^{c}\binom{NPt}{d}\binom{N-NPt}{n-d}}{\binom{N}{n}}$$

4. The producer's risk when the lot containing a proportion of defectives is rejected.

$P_p =$ _____

= 1- _____

From the above expression we draw the following conclusions

(a)

Number of defectives is approximately binomial

It is instructive to show how the points on this curve are obtained, once we have a sampling plan (*n*,*c*) - later we will demonstrate how a sampling plan (*n*,*c*) is obtained.

We assume that the lot size *N* is very large, as compared to the sample size *n*, so that removing the sample doesn't significantly change the remainder of the lot, no matter how many defects are in the sample. Then the distribution of the number of defectives, *d*, in a random sample of *n* items is approximately binomial with parameters *n* and *p*, where *p* is the fraction of defectives per lot.

The probability of observing exactly *d* defectives is given by

$$P(d) = f(d) = \frac{n!}{d!(n-d)!} p^d (1-p)^{n-d}$$

The probability of acceptance is the probability that *d*, the number of defectives, is less than or equal to *c*, the accept number. This means that

$$P_a = P\{d \le c\} = \sum_{d=0}^{c} \frac{n!}{d!(n-d)!} p^d (1-p)^{n-d}$$

Sample table for Pa, Pd using the binomial distribution

Using this formula with *n* = 52 and *c*=3 and *p* = .01, .02, ...,.12 we find

P_a	P_d
.998	.01
.980	.02
.930	.03
.845	.04
.739	.05
.620	.06
.502	.07
.394	.08
.300	.09
.223	.10

.162 .11
.115 .12

Equations for calculating a sampling plan with a given OC curve

In order to design a sampling plan with a specified OC curve one needs two designated points. Let us design a sampling plan such that the probability of acceptance is 1-α for lots with fraction defective p_1 and the probability of acceptance is β for lots with fraction defective p_2. Typical choices for these points are: p_1 is the *AQL*, p_2 is the *LTPD* and α, β are the Producer's Risk (Type I error) and Consumer's Risk (Type II error), respectively.

If we are willing to assume that binomial sampling is valid, then the sample size n, and the acceptance number c are the solution to

$$1-\alpha=\sum_{d=0}^{c}\frac{n!}{d!(n-d)!}p_1^d(1-p_1)^{n-d}$$

$$\beta=\sum_{d=0}^{c}\frac{n!}{d!(n-d)!}p_2^d(1-p_2)^{n-d}$$

These two simultaneous equations are nonlinear so there is no simple, direct solution. There are however a number of iterative techniques available that give approximate so that composition of a computer program posses few problems.
We can also calculate the *AOQ* for a (n,c) sampling plan, provided rejected lots are 100% inspected and defectives are replaced with good parts100% inspected and defectives are replaced with good parts.

Assume all lots come in with exactly a p_0 proportion of defectives. After screening a rejected lot, the final fraction defectives will be zero for that lot. However, accepted lots have fraction defective p_0. Therefore, the outgoing lots from the inspection stations are a mixture of lots with fractions defective p_0 and 0. Assuming the lot size is N, we have.

$$AOQ=\frac{p_a p(N-n)}{N}$$

For example, let N = 10000, n = 52, c = 3, and p, the quality of incoming lots, = 0.03. Now at p = 0.03, we glean from the OC curve table that p_a = 0.930 and

AOQ = (.930)(.03)(10000-52) / 10000 = 0.02775.

Sample table of AOQ versus p

Setting p = .01, .02, ..., .12, we can generate the following table

AOQ	p
.0010	.01
.0196	.02
.0278	

Interpretation of AOQ plot:- From examining this curve we observe that when the incoming quality is very good (very small fraction of defectives coming in), then the outgoing quality is also very good (very small fraction of defectives going out). When the incoming lot quality is very bad, most of the lots are rejected and then inspected. The "duds" are eliminated or replaced by good ones, so that the quality of the outgoing lots, the *AOQ*, becomes very good. In between these extremes, the *AOQ* rises, reaches a maximum, and then drops.

The maximum ordinate on the *AOQ* curve represents the worst possible quality that results from the rectifying inspection program. It is called the **average outgoing quality limit,** (***AOQL***).

From the table we see that the *AOQL* = 0.0372 at $p = .06$ for the above example.

One final remark: if $N >> n$, then the $AOQ \sim p_a\, p$.

Calculating the Average Total Inspection

What is the total amount of inspection when rejected lots are screened?

If all lots contain zero defectives, no lot will be rejected.

If all items are defective, all lots will be inspected, and the amount to be inspected is *N*.

Finally, if the lot quality is $0 < p < 1$, the average amount of inspection per lot will vary between the sample size *n*, and the lot size *N*.

Let the quality of the lot be *p* and the probability of lot acceptance be p_a, then the *ATI* per lot is

$$ATI = n + (1 - p_a)(N - n)$$

For example, let $N = 10000$, $n = 52$, $c = 3$, and $p = .03$ we know from the OC table that $p_a = 0.930$. Then $ATI = 52 + (1-.930)(10000 - 52) = 753$. (Note that while 0.930 was rounded to three decimal places, 753 were obtained using more decimal places.)

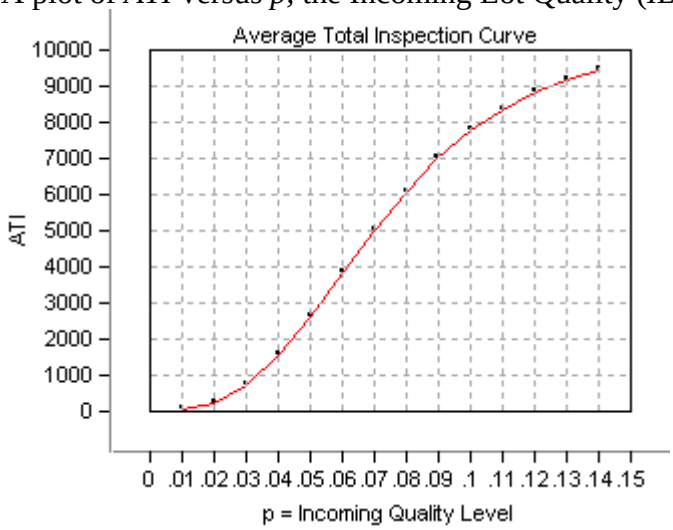
Sample table of ATI versus p

Setting *p*= .01, .02, ...14 generates the following table

ATI	P
70	.01
253	.02
753	.03
1584	.04
2655	.05
3836	.06
5007	.07
6083	.08
7012	.09
7779	.10
8388	.11
8854	.12
9201	.13
9453	.14

Plot of ATI versus p

A plot of *ATI* versus *p,* the Incoming Lot Quality (*ILQ*) is given below.



Let is proportion of defective items in the batch

X = number of defectives in the sample, we accept the batch if X where c is constant.

otherwise reject the batch if X c

Case 1 where N is infinite

X~ Bin (n,) X = 0,1,2…..n

Example 2:-We might assume c=0 , the conditions " be 99% of a certain rejecting a batch with 2% of defectives

Solution :- , OC(

$OC(0.02) \leq 0.01$

but for single stage sampling $OC(\theta) = (1-\theta)^n$

$= (1\text{-}0.02)^n$

$(1\text{-}0.02)^n \leq 0.01$

By taking natural logarithm both sides we get,

nlog(0.98)

n =227.9

= n 228 since n must be an integer ,so that we can "take a sample of 228 observation accept the batch if every object is good and reject if there is defectives"

Example 3 (instead of fixing c , fix n) :- let take n =100 what will give us a 95% chance of rejecting a batch that has 5% defectives.

Solution: - we can calculate OC (0.05) , the value of c is 0, 1, 2, 3…n =100.

c= 0,$OC(0.05) = (1\text{-}0.05)^{100}$

$= (0.95)^{100}$

= 0.00592

C=1,OC(0.05) = p(x

= p(x=0) + p(x=1) by using the binomial distribution we can get

$= (0.95)^{100} + (0.05)^1 (0.95)^{99}$

= 0.037

C=2 OC(0.05) =p(x = p(x=0) + p(x=1) +p(x=2)

by using the binomial distribution we can get

$= (0.95)^{100} + (0.05)^1 (0.95)^{99} + (0.05)^2 (0.95)^{98}$

= 0.11

OC(0.05) =

Example 3:- (fix c\n) a buyers test a sample of n fuses from infinite batch (population) of fuses. She will accept the batch if the sample contains defectives. she is considering by taking a sample of n = 25,50,75,100 etc…fuses .how large should "n" be if she wants to be 90% of certain of rejecting a batch with 10% of defectives? i.e.c\n =0.04 find the value of c and n.

Solution :- let us check possible solutions C=1,n=25

C=2,n=50

C=3, n=75

C=4, n=100

OC(0.1) C=1,n=25

OC(0.1) = p(x = p(x=0) +p(x =1)

$= (0.9)^{25} + (0.1)^{1} (0.9)^{24} = 0.27$

OC(0.1) = 0.27 >10%

This is not solution set.

C=2,n=50, OC(0.1) =P(x =p(x =0)+ p(x =1) +p(x =2)

$= (0.9)^{50} + \binom{50}{1} (0.1)^{1} (0.9)^{49} + (0.1)^{2} (0.9)^{48}$

= 0.11>10%

This is also not solution set.

C=3,n=75, OC(0.1) =P(x = P(x=0) +…P(x=3)

We finally get as, the same procedure above, = 0.059<10%

OC(0.1) = 0.059

This is a solution set.

we need n and c=3

Case 2 N is finite (N< :-

Here we have a batch of N items, M of this are defectives ; is proportion of defectives =M\N

Note:- can only take the values

0, 1\N,2\N, 3\N, ….N-1\N,1

X= number of defective in the sample has hyper- geometric distribution

P(X=x) =

where x , n-x

P(X=x) = The probability of getting ‘x ‘defective in the sample.

Example 4 :- Suppose N=100,c=0,choose n to make sure OC(0.05) i.e. we have a batch of size 1000 we use the rule “accept if the sample contains no defectives” choose n so that we are 95% of rejecting a batch is 5% defectives.

Choose the sample observation n when OC(0.05)

since , M=N =50

OC(x =0\M=50)

= P(every sample observation have no defective\M=50)

= 950\100

by inspection OC(0.05) when n

Example 5:- Suppose the batch size is N= 40 and we take c=0 how can we choose n so that OC (0.05) for this value of n calculate the OC functions?

Solution N=40, ,

M=N =0.05

OC(0.05) = P(accept the batch\M=2)

= P(1^{st} item is good) P(2^{nd} item is good\$1^{st}$ item is good) P(3^{rd} item is good\$1^{st}$ and 2^{nd} item is good) …… P(n^{th} item is good \$1^{st}$,2^{nd} ,3^{rd} ….$(n-1)^{th}$ item is good)

OC(0.05) = P(X =

=

38\40)\(40-(n+1)) =(40-n)(39-n)\(40

(40-n)(39-n)\(40 also by inspection n=31 =0.046

We need to choose n=31 , n is the minimum sample observation we must use n

How double sampling plans work:-

Double and multiple sampling plans were invented to give a questionable lot another chance. For example, if in double sampling the results of the first sample are not conclusive with regard to accepting or rejecting, a second sample is taken. Application of double sampling requires that a First sample of size n_1 is taken at random from the (large) lot. The number of defectives is then counted and compared to the first sample's acceptance number a_1 and rejection number r_1.

Denote the number of defectives in sample 1 by d_1 and in sample 2 by d_2, then:

If d_1 a_1, the lot is accepted.

If d_1 r_1, the lot is rejected.

If $a_1 < d_1 < r_1$, a second sample is taken.

If a second sample of size n_2 is taken, the number of defectives, d_2, is counted.

The total number of defectives is $D_2 = d_1 + d_2$. Now this is compared to the acceptance number a_2 and the rejection number r_2 of sample 2. In double sampling, $r_2 = a_2 + 1$ to ensure a decision on the sample.

If D_2

Design of a double sampling plan

The parameters required to construct the OC curve are similar to the single sample case. The two points of interest are (p_1, 1-α) and (p_2, β, where p_1 is the lot fraction defective for plan 1 and p_2 is the lot fraction defective for plan 2. As far as the respective sample sizes are concerned, the second sample size must be equal to, or an even multiple of, the first sample size.

There exist a variety of tables that assist the user in constructing double and multiple sampling plans. The index to these tables is the p_2/p_1 ratio, where $p_2 > p_1$.

One set of tables, taken from the Army Chemical Corps Engineering Agency for α= .05 and β= .10, is given below:

Tables for $n_1 = n_2$

$R = p_2/p_1$	accept numbers c_1	c_2	approximation of pn_1 P = .95	values for P = .10
11.90	0	1	0.21	2.50
7.54	1	2	0.52	3.92
6.79	0	2	0.43	2.96
5.39	1	3	0.76	4.11
4.65	2	4	1.16	5.39
4.25	1	4	1.04	4.42
3.88	2	5	1.43	5.55
3.63	3	6	1.87	6.78
3.38	2	6	1.72	5.82
3.21	3	7	2.15	6.91
3.09	4	8	2.62	8.10
2.85	4	9	2.90	8.26
2.60	5	11	3.68	9.56
2.44	5	12	4.00	9.77
2.32	5	13	4.35	10.08
2.22	5	14	4.70	10.45
2.12	5	16	5.39	11.41

Tables for $n_2 = 2n_1$

R = p_2/p_1	accept numbers c_1	c_2	approximation of pn_1 P = .95	values for P = .10
14.50	0	1	0.16	2.32
8.07	0	2	0.30	2.42
6.48	1	3	0.60	3.89
5.39	0	3	0.49	2.64
5.09	0	4	0.77	3.92
4.31	1	4	0.68	2.93
4.19	0	5	0.96	4.02
3.60	1	6	1.16	4.17
3.26	1	8	1.68	5.47
2.96	2	10	2.27	6.72
2.77	3	11	2.46	6.82
2.62	4	13	3.07	8.05
2.46	4	14	3.29	8.11
2.21	3	15	3.41	7.55
1.97	4	20	4.75	9.35
1.74	6	30	7.45	12.96

***Example of a double sampling plan*:-**

We wish to construct a double sampling plan according to

$p_1 = 0.01$ $\alpha = 0.05$ $p_2 = 0.05$ $\beta = 0.10$ and $n_1 = n_2$

The plans in the corresponding table are indexed on the ratio

$R = p_2/p_1 = 5$

We find the row whose R is closet to 5. This is the 5th row ($R = 4.65$). This gives $c_1 = 2$ and $c_2 = 4$.

The value of n_1 is determined from either of the two columns labeled pn_1.

The left holds α constant at 0.05 ($P = 0.95 = 1 - \alpha$) and the right holds β constant at 0.10. ($P = 0.10$).

Then holding α constant we find $pn_1 = 1.16$ so $n_1 = 1.16/p_1 = 116$.

And, holding β constant we find $pn_1 = 5.39$, so $n_1 = 5.39/p_2 = 108$. Thus the desired sampling plan is

$n_1 = 108 \quad c_1 = 2 \quad n_2 = 108 \quad c_2 = 4$

If we opt for $n_2 = 2n_1$, and follow the same procedure using the appropriate table, the plan is:

$n_1 = 77 \quad c_1 = 1 \quad n_2 = 154 \quad c_2 = 4$

The first plan needs less samples if the number of defectives in sample 1 is greater than 2, while the second plan needs less samples if the number of defectives in sample 1 is less than 2.

ASN Curve for a Double Sampling Plan

Since when using a double sampling plan the sample size depends on whether or not a second sample is required, an important consideration for this kind of sampling is the Average Sample Number (*ASN*) curve.

This curve plots the *ASN* versus *p'*, the true fraction defective in an incoming lot.

We will illustrate how to calculate the *ASN* curve with an example.

Consider a double-sampling plan $n_1 = 50$, $c_1 = 2$, $n_2 = 100$, $c_2 = 6$,

where n_1 is the sample size for plan 1, with accept number c_1, and n_2, c_2, are the sample size and accept number, respectively, for plan 2.

Let $p' = .06$. Then the probability of acceptance on the first sample, which is the chance of getting two or less defectives, is .416 (using binomial tables). The probability of rejection on the second sample, which is the chance of getting more than six defectives, is (1-.971) = .029.

The probability of making a decision on the first sample is .445, equal to the sum of .416 and .029.

With complete inspection of the second sample, the *average* size sample is equal to the size of the first sample times the probability that there will be only one sample

plus the size of the combined samples times the probability that a second sample will be necessary.

For the sampling plan under consideration,

the *ASN* with complete inspection of the second sample for a p' of .06 is

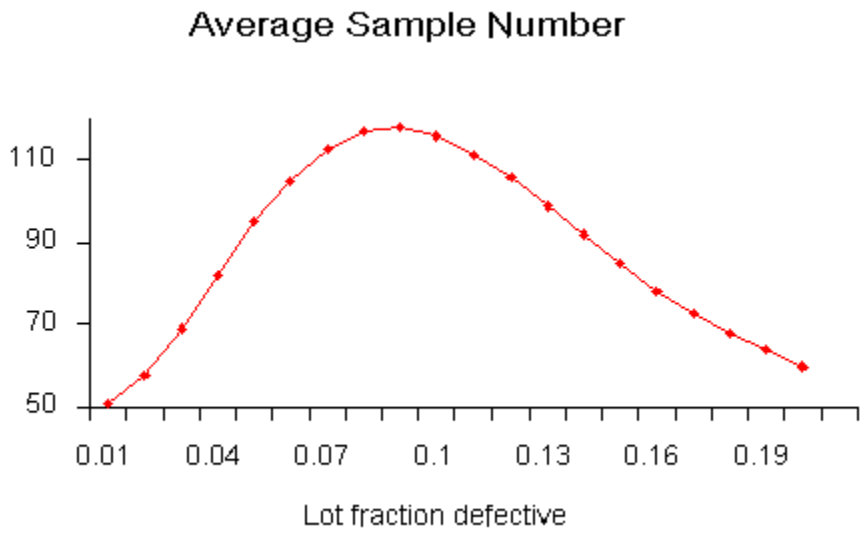
$$50(.445) + 150(.555) = 106$$

The general formula for an average sample number curve of a double-sampling plan with complete inspection of the second sample is

$$ASN = n_1P_1 + (n_1 + n_2)(1 - P_1) = n_1 + n_2(1 - P_1)$$

where P_1 is the probability of a decision on the first sample.

The graph below shows a plot of the *ASN* versus p'.



4..2.5 Multiple Stage Sampling

Multiple sampling is an extension of double sampling.

It involves inspection of 1 to *k* successive samples as required to reach an ultimate decision.

Mil-Std 105D suggests $k = 7$ is a good number.

Multiple sampling plans are usually presented in tabular form:

Procedure for multiple sampling

The procedure commences with taking a random sample of size n_1 from a large lot of size *N* and counting the number of defectives, d_1.

if $d_1 \leq a_1$ the lot is accepted.

if $d_1 \geq r_1$ the lot is rejected.

if $a_1 < d_1 < r_1$, another sample is taken.

If subsequent samples are required, the first sample procedure is repeated sample by sample.

For each sample, the total number of defectives found at any stage, say stage *i*, is

$$D_i = \sum_{j=1}^{i} d_j$$

This is compared with the acceptance number a_i and the rejection number r_i for that stage until a decision is made.

Sometimes acceptance is not allowed at the early stages of multiple sampling;

However, rejection can occur at any stage.

Efficiency measured by the ASN:-

Efficiency for a multiple sampling scheme is measured by the *average sample number (ASN)*

Required for a given Type I and Type II set of errors.

The number of samples needed when following a multiple sampling scheme may vary from trial to trial, and the *ASN* represents the average of what might happen over many trials with a fixed incoming defect level

4.2.6 Sequential Sampling Plan

Description of sequential sampling graph :-

The cumulative observed number of defectives is plotted on the graph.

For each point, the x-axis is the total number of items thus far selected, and the y-axis is the total number of observed defectives. If the plotted point falls within the parallel lines the process continues by sample. As soon as a point falls on or above the upper line, the lot is rejected. And when a point falls on or line, the lot is accepted. The process can theoretically last until the lot is 100% inspected. However, as a rule of thumb, sequential-sampling plans are truncated after the number inspected reaches three times would have been inspected using a corresponding single sampling plan.

Equations for the limit lines

The equations for the two limit lines are functions of the parameters p_1, α, p_2, and β.

$$x_a = -h_1 + sn \quad \text{(acceptance line)}$$

where

Instead of using the graph to determine the fate of the lot, one can resort to generating tables (with the help of a computer program)

Example of a sequential sampling plan

As an example, let $p_1 = .01$, $p_2 = .10$, $\alpha = .05$, $\beta = .10$. The resulting equations are

Both acceptance numbers and rejection numbers must be integers. The acceptance number is the next integer less than or equal to x_a and the rejection number is the next integer greater than or equal to x_r. Thus for $n = 1$, the acceptance number = -1, which is impossible, and the rejection number = 2, which is also impossible. For $n = 24$, the acceptance number is 0 and the rejection number = 3.

.

The results for n =1, 2, 3... 26 are tabulated below.

n inspect	n accept	n reject	n inspect	n accept	n reject
1	x	x	14	x	2
2	x	2	15	x	2
3	x	2	16	x	3
4	x	2	17	x	3
5	x	2	18	x	3
6	x	2	19	x	3
7	x	2	20	x	3
8	x	2	21	x	3
9	x	2	22	x	3
10	x	2	23	x	3
11	x	2	24	0	3
12	x	2	25	0	3
13	x	2	26	0	3

So, for n = 24 the acceptance number is 0 and the rejection number is 3. The "x" means that acceptance or rejection is not possible.

Other sequential plans are given below.

n inspect	n accept	n reject
49	1	3
58	1	4
74	2	4
83	2	5
100	3	5
109	3	6

The corresponding single sampling plan is (52,2) and double sampling plan is (21,0), (21,1).

Efficiency measured by ASN

Efficiency for a sequential sampling scheme is measured by the *average sample number (ASN)* required for a given Type I and Type II set of errors. The number of samples needed when following a sequential sampling scheme may vary from trial to trial, and the *ASN* represents the average of what might happen over many trials with a fixed incoming defect level. Good software for designing sequential sampling schemes will calculate the *ASN* curve as a function of the incoming defect level.

4.2.7 Skip Lot Sampling Plan

Skip Lot sampling means that only a fraction of the submitted lots are inspected.
This mode of sampling is of the cost-saving variety in terms of time and effort.
However skip-lot sampling should only be used when it has been demonstrated that the quality of the submitted product is very good.
Implementation of skip-lot sampling plan:-

A skip-lot sampling plan is implemented as follows:

1. Design a single sampling plan by specifying the alpha and beta risks and the consumer/producer's ris called "the reference sampling plan".
2. Start with normal lot-by-lot inspection, using the reference plan.
3. When a pre-specified number, *i*, of consecutive lots are accepted, switch to inspecting only a fraction *f* selection of the members of that fraction is done at random.
4. When a lot is rejected return to normal inspection.

The f and i parameters

The parameters *f* and *i* are essential to calculating the probability of Acceptance for a skip-lot sampling plan. In this scheme, *i*,

t

called the *clearance number,* is a positive integer and the sampling fraction f is such that $0 < f < 1$. Hence, when $f = 1$ there is no longer skip-lot Sampling.

The calculation of the acceptance probability for the skip-lot sampling plan performed via the following formula

where P is the probability of accepting a lot with a given proportion of

incoming defectives p, from the OC curve of the single sampling plan.

The following relationships hold:

for a given i, the smaller is f, the greater is P_a

for a given f, the smaller is i, the greater is P_a

An illustration of a a skip-lot sampling plan is given below

ASN of skip-lot sampling plan:-

An important property of skip-lot sampling plans is the average sample number (*ASN*).

The *ASN* of a skip-lot sampling plan is

$$ASN_{skip\text{-}lot} = (F)(ASN_{reference})$$

where *F* is defined by

$$F = \frac{f}{(1-f)P^i + f}$$

Therefore, since $0 < F < 1$, it follows that the *ASN* of skip-lot sampling is smaller than the

ASN of the reference sampling plan.

In summary, skip-lot sampling is preferred when the quality of the submitted lots is excellent and the supplier can demonstrate a proven track record.

4.3 Curves For Sampling Plan

The following curves are quite essential in understanding the behavior and operation of sampling inspection plan.

1.

Applying OC-Curve to Evaluate Sampling Plans

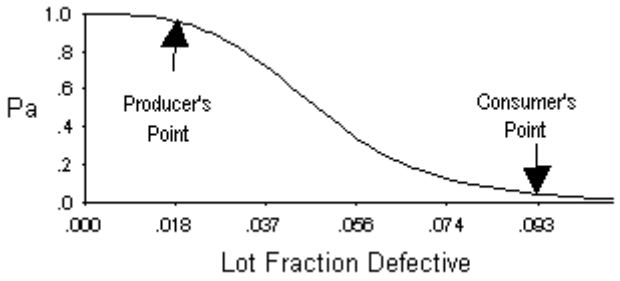


Figure 1. OC-Curve. (Operating Characteristic)

The Operating Characteristic (OC) curve shows the probability of acceptance, Pa, for any level of lot quality. See *Figure 1.* On the horizontal axis is the quality characteristic.

This OC-curve enables you to evaluate the probability of acceptance for any true lot quality level-on a what-if basis. This way, you can design sampling plans that **perform** the way you want.

Interpret the curve according to this example:

1. If the lot quality is 0.093 fractions defective, then the probability of acceptance, Pa, is 0.05.
2. If the lot quality is 0.018 fractions defective, then the probability of acceptance, Pa, is 0.95.

APPLYING OC-CURVES to COMPARE SAMPLING PLANS

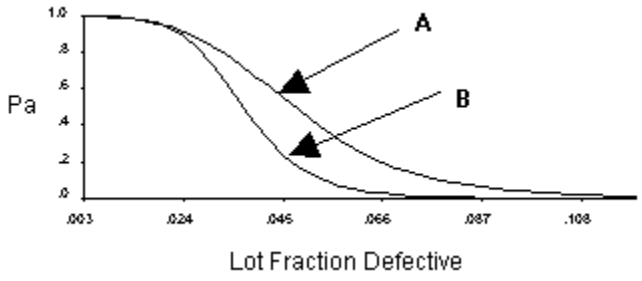


Figure 2. Comparing Alternative Plans, A and B

You can use OC-curves to compare alternative plans. See *Figure 2.* Choose between the plans by their relative ability to detect reject able lots. You should expect that the steeper the curve, the larger the sample size.

Complete this picture by comparing the costs of the sampling to the resulting performance.

APPLYING OC-CURVES to DESIGN -- THE TWO-POINT METHOD

The Two-Point method for developing acceptance sampling plans requires that you specify two points of the operating characteristic curve (OC-curve).

Producer's Point

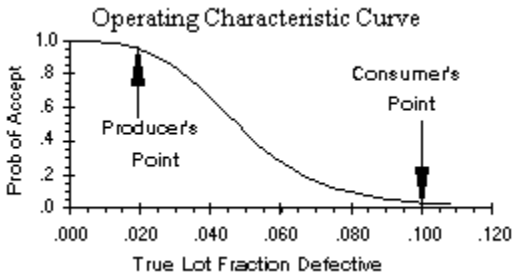


Figure 3. The Two-Point Method

The **producer's point** controls the acceptance of lots that are at an acceptable quality level. (See *figure 3*) The goal: prevent good lots from being rejected.

Consumer's Point

The **consumer's point** controls the rejection of lots that are at a reject able quality level. (See *figure 3*) The goal: prevent bad lots from being accepted.

Decision Table Defines the Two Points

Matching the OC Curves of Different Types of Plan

An important ability is to match sampling plans by their OC-curves. Two matched plans have the same operating characteristic curve, but different decision rules. You can safely choose between matched plans for economy, knowing they offer equal protection. The following table shows useful matches.

Plan A	Plan B
Attribute	Variables
Fixed-n	Double, Multiple, Sequential
Variables - Known standard deviation	Variables - Unknown standard deviation

The following series of sampling plan examples shows how various types of plan are matched to the same OC-curve. The OC-curve of this example is characterized by the two points: (AQL=0.01, =0.05) and (RQL=0.10, =0.05). For convenience we will refer to the curve as "**OC-Curve X**".

Attribute Sequential Sampling Plan (Matched to OC-Curve X)

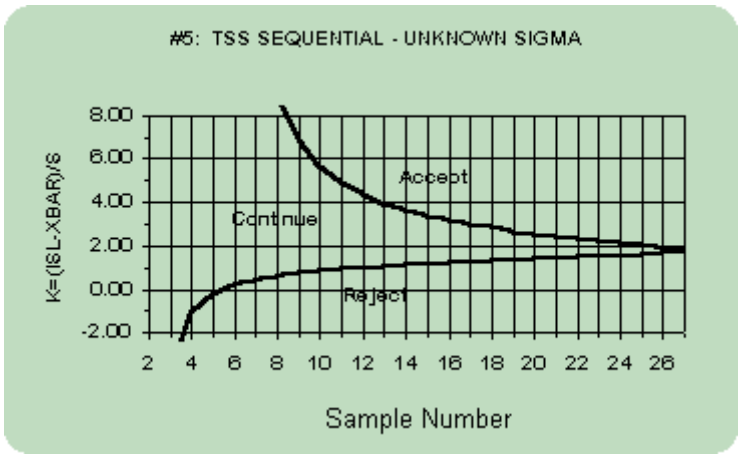
Sequential Probability Ratio method (SPR)

(AQL=0.01, =0.05) and (RQL=0.10, =0.05)

SPR Sequential		
(n)	(Ac)	(Re)
1	###	12.55
2	1.13	8.74
3	2.40	7.48
4	3.04	6.84
5	3.42	6.46
6	3.67	6.21
7	3.85	6.03
8	3.99	5.89
9	4.09	5.78
10	4.18	5.70
11	4.25	5.63
12	4.31	5.57
13	4.36	5.52

Variables Sequential Sampling Plan, unknown Sigma
(Matched to OC-Curve X)

TSS = **Truncatable Single Sample**
(AQL=0.01, =0.05) and (RQL=0.10, =0.05)



Variables
TSS Sequential
Unknown Sigma

(n)	K (Re)	K (Ac)
2		
3	-3.83	
4	-1.03	
5	-0.20	
6	0.21	
7	0.46	
8	0.63	8.94
9	0.76	6.86
10	0.87	5.66
11	0.96	4.88
12	1.03	4.33
13	1.10	3.92
14	1.16	3.61
15	1.21	3.36
16	1.26	3.15
17	1.30	2.98
18	1.35	2.83
19	1.39	2.69
20	1.43	2.58
21	1.47	2.47
22	1.51	2.37
23	1.55	2.27
24	1.59	2.18
25	1.64	2.08
26	1.70	1.96
27	1.80	1.80

Relationship of Sampling to Control Charts

Shewhart Control Charts are not Acceptance Plans

Shewhart Xbar and R charts analyze processes that involve a series of lots produced over time. They concern the relationship of the subgroups to each other, and not to any externally imposed specification. Use Xbar and R charts to discover the factors that contribute to process variability.

Shewhart charts cannot ensure against accepting poor or recallable lots. Even an in-control characteristic can have a substantial fraction of non-conformities. Xbar and R charts do not control the consumer's risk (ß) of accepting RQL or recallable lots.

Acceptance Control Charts.

Acceptance control charts are acceptance sampling plans that you convert into chart form for implementation. They control the producers point **and** the consumer's point of the oc-curve. Acceptance charts provide a valid visible means for making acceptance sampling decisions.

Acceptance Sampling for Acceptance Decisions

The best way to make the accept/reject decision - whether a process is out-of-control or in-control-is to use both the producer's risk and the consumer's risk. In other words, honor **both** the process capability and the product specifications

2. ***Average Outgoing Quality (AOQ) Curve:-***

AOQ curve shows the relation ship between incoming lot quality p and average out going quality AOQ. For a single sampling plan AOQ curves given by

$$AOQ = \frac{P\,(N-n)\,.Pa\,(p)}{N}$$

Where symbols have their usual meaning. If sample size n is small compared with lot size N, so that sampling fraction n\N can being noted. This can be approximated by

$$AOQ = P\left[N - n \diagup N\right] Pa\,(p)$$

.

$= P\left[1 - n/N\right] Pa\ (p)$

$\Rightarrow AOQ \approx P.Pa\ (p)$

2 Average Total Inspection (A.T.I) Curve:-

It depicts graphically the relation ship between the incoming lot quality 'p' and the average total number of units inspected per lot, including sampling and sorting 100%. For instance in case of single sampling plan A.T.I curve is given by:

ATI = n+ (N-n) $\left[1 - Pa(p)\right]$, since n items to be inspected in each case and the remaining N-n items will be inspected only if lot is rejected (i.e. if d>c); Pa (p) being the probability of acceptance for a lot of incoming quality p.

Exercise

1. Write down OC function for a binomial sampling scheme with c = 0 (your answer is based on a function of n). Differentiate this function and show its derivative at 0 equals' n.

2. Suppose we have a binomial sampling scheme and take c = 0. how large should we take n if we want to make sure :

    a. We are at least 99% certain rejection of a batch which has 8% defectives?

    b. We are at least 90% certain of accept a batch which has 1% of defectives?

3. Suppose we intend to create a binomial sampling scheme and we want make sure:

    a. We are at least 90% of certain rejecting a batch which has 10% of defectives?

b. We are at least 90% of a certain acceptance a batch which has 3% of defectives?

c. The scheme uses c = 0

It is possible to create this scheme? If it is, what are the possible values of n we can take?

4. Suppose we have a binomial sampling scheme with n = 40. what value should we take for c we want to make sure :

a. We are at least 90% of certain rejecting a batch which has 10% of defectives?

b. We are at least 90% certain of accept a batch which has 1% of defectives?

It is possible to create the scheme so that we are 90% of certain rejecting a batch which has 10% of defectives and 90% of certain of accepting a batch which has1% of defectives?

5. write down OC function

a. for binomial sampling scheme with c = 1(your answer will be a function of n)

b. Differentiate this function and show that its derivative at 0 equals' n. compares this with your answers to question 1.interpriate your answer.

c. Calculate OC (0101), OC (0.03), OC (0.05), OC (0.15) and OC (0.10) for binomial sampling scheme with c = 1 and n = 80.sktech the function.

.

6. Suppose we have a binomial sampling scheme with c = 1. Show that if we want to make sure we are at least 95% certain rejecting a batch that has 10% of defectives we need to take $n \geq 46$.

7. An engineer receives a large batch of items and need to decide whether or not she will accept the batch. She consider using the following sampling schemes

   (i). take a sample of 65 items and accept the batch if there are no defective items in the sample.

   (ii). Take a sample of 100 items and accept the batch if there is at most one defective items in the sample.

   Use your answers to question 1 and 5 to write down the OC functions corresponding to each of these samples. Show that OC1 (0.05) is approximately equal to OC2 (0.05). Sketch the two OC functions on the same graph. What does this tell you about the two schemes?

8. Consider the binomial sampling with c = 0.002n. Derive the following table (note this is quick to do in excel using the function BINOMDIST, but can be quite long to do by hand. Just check one or two entries if not familiar with excel or don't have access to it.)

n	θ		
	0.01	0.05	0.10
50	0.911	0.279	0.034
100	0.921	0.118	0.002
150	0.935	0.005	0.000
200		0.026	0.000

9. Suppose we have a batch size N. What is the OC function of the hyper geometric sampling scheme with c =1.

10. Consider two acceptance sampling scheme defined by the following rules:

(i). the single stage binomial sampling schemes with n = 59 and c = 0. “Take a sample size of 59 and accept the batch if it contains no defectives”

(ii). in double – stage binomial sampling scheme with n = 70, c1 = 0, d1 =2 and c2 = 1:”take a sample of size 70. Then

a. If it contains no defectives accept the batch.
b. If it contains two or more defectives reject the batch.
c. If it contains exactly one defective take a further sample of size m and accept the batch if this sample has no defectives.

Calculate the OC function for each of these schemes. We saw from the module OC1 (0.05) = 0.05. Calculate also OC1 (0.01) and OC1 (0.10); find the smallest values of m so that OC2 (0.05) <0.05. For this value of m calculate OC2 (0.01), OC2 (0.05) and OC2 (0.10). Interpret your results.

11. consider the double stage binomial sampling scheme with n = 50, m = 80, c1 = 0, d1 = 2 and c2 =1: this means we take a sample of size 50 then:

a. if it contains no defectives accept the batch
b. if it contain two or more defectives reject the batch
c. If it contains exactly one defective take a further sample of size 80 and accept the batch if this sample has no defective, calculate OC functions of these schemes. Calculate OC(0.01), OC(0.05), and OC(0.10)
d. The probability that the batch is accepted after the second stage.
e. The probability that the batch is rejected after the second stage, suppose θ = 0.1. Calculate the expected number of items sampled.

12. Write down the OC function of the one stage acceptance sampling scheme with c= 0 which values of n define an admissible scheme in each of the following cases:

a. RQL 5%, OC (RQL) < 5%

b. AQL = 1% OC(AQL)>90%

c. AOQL<1%

13. |Suppose we take k =2 what values of n make sure

(i). we are at least 90% certain of rejecting a batch which has 7% of defectives?

(ii) We are at least 99% certain rejection of which has 1% of defectives?

14. Design an acceptance sampling scheme that satisfy OC (RQL) < β and OC (AQL) >1- α for the values given below. (Assume N= ∞). State the value of n and k you would use.

(i). RQL =10%, AQL = 1%. $\beta = 5\%, \alpha = 10\%$

(ii). RQL= 20%, AQL=10%, $\beta = 10\%, \alpha = 10\%$

(iii)RQL= 5%, AQL=1%, $\alpha = 1\%$, $\beta = 1\%$

Chapter Five

5. Reliability and Life Testing

5.1 Introduction

5HOLDELOLW\ LV VRPHWKLQJ WKDW HYHU\RQH EHOLHYHV WKH\ XQGHUVWDQG EXW QHHGV D ZDUQLQJ ZKHQ LW FRPHV WR WKH VDPH EXW QHHGV D GHILQLWLRQ LV WKH VDPH EXW QHHGV D GHILQLWLRQ

$Q RWKHU ZD\ RI ORRNLQJ DW WKLV LV WR FRQVLGHU UHSHDWDELOLW\ RU FRQVLVWHQF\ LQ

in an experiment. For maintaining reliability internally, a researcher will use as many repeat sample groups as possible, to reduce the chance of an abnormal sample group skewing the results.

Reliability and Statistics

Physical scientists expect to obtain exactly the same results every single time, due to the relative predictability of the physical realms. If you are a nuclear physicist or an inorganic chemist, repeat experiments should give exactly the same results, time after time.

Ecologists and social scientists, on the other hand, understand fully that achieving exactly the same results is an exercise in futility. Research in these disciplines incorporates random factors and natural fluctuations and, whilst any experimental design must attempt to eliminate confounding variables and natural variations, there will always be some disparities.

The key to performing a good experiment is to make sure that your results are as reliable as is possible; if anybody repeats the experiment, powerful statistical tests will be able to compare the results and the scientist can make a solid estimate of statistical reliability.

The definition of reliability vs. validity

Reliability and validity are often confused, but the terms actually describe two completely different concepts, although they are often closely inter-related. This distinct difference is best summed up with an example:

A researcher devises a new test that measures IQ more quickly than the standard IQ test:

- If the new test delivers scores for a candidate of 87, 65, 143 and 102, then the test is not reliable or valid, and it is fatally flawed.
- If the test consistently delivers a score of 100 when checked, but the candidates real IQ is 120, then the test is reliable, but not valid.
- If the researcher's test delivers a consistent score of 118, then that is pretty close, and the test can be considered both valid and reliable.

Reliability is an essential component of validity but, on its own, is not a sufficient measure of validity. A test can be reliable but not valid, whereas a test cannot be valid yet unreliable.

Reliability, in simple terms, describes the repeatability and consistency of a test. Validity defines the strength of the final results and whether they can be regarded as accurately describing the real world.

The definition of reliability –an example

Imagine that a researcher discovers a new drug that she believes helps people to become more intelligent, a process measured by a series of mental exercises. After analyzing the results, she finds that the group given the drug performed the mental tests much better than the control group.

For her results to be reliable, another researcher must be able to perform exactly the same experiment on another group of people and generate results with the same statistical significance. If repeat experiments fail, then there may be something wrong with the original research.

The real difference between reliability and validity is mostly a matter of definition. Reliability estimates the consistency of your measurement, or more simply the degree to which an instrument measures the same way each time it is used in under the same conditions with the same subjects. Validity, on the other hand, involves the degree to which you are measuring what you are supposed to, more simply, the accuracy of your measurement. It is my belief that validity is more important than reliability because if an instrument does not accurately measure what it is supposed to, there is no reason to use it even if it measures consistently (reliably).

5.2 Definition of Reliability

Definition: Reliability is the consistency of your measurement, or the degree to which an instrument measures the same way each time it is used under the same condition with the same subjects. In short, it is the repeatability of your measurement. A measure is considered reliable if a person's score on the same test given twice is similar. It is important to remember that reliability is not measured, it is estimated.

There are two ways that reliability is usually estimated: test/retest and internal consistency.

Test/Retest

Test/retest is the more conservative method to estimate reliability. Simply put, the idea behind test/retest is that you should get the same score on test 1 as you do on test 2. The three main components to this method are as follows:

1). Implement your measurement instrument at two separate times for each subject; 2).Compute the correlation between the two separate measurements; and 3) Assume there is no change in the underlying condition (or trait you are trying to measure) between test 1 and test 2.

Internal-Consistency

Internal consistency estimates reliability by grouping questions in a questionnaire that measure the same concept. For example, you could write two sets of three questions that measure the same concept (say class participation) and after collecting the responses, run a correlation between those two groups of three questions to determine if your instrument is reliably measuring that concept.

One common way of computing correlation values among the questions on your instruments is by using Cronbach's Alpha. In short, Cronbach's alpha splits all the questions on your instrument every possible way and computes correlation values for them all (we use a computer program for this part). In the end, your computer output generates one number for Cronbach's alpha - and just like a correlation coefficient, the closer it is to one, the higher the reliability estimate of your instrument. Cronbach's alpha is a less conservative estimate of reliability than test/retest.

The primary difference between test/retest and internal consistency estimates of reliability is that test/retest involves two administrations of the measurement instrument, whereas the internal consistency method involves only one administration of that instrument.

5.2.1 Relation and Difference between Validity and Reliability

Definition: Validity is the strength of our conclusions, inferences or propositions. More formally, Cook and Campbell (1979) define it as the "best available approximation to the truth or falsity of a given inference, proposition or conclusion." In short, were we right? Let's look at a simple example. Say we are studying the effect of strict attendance policies on class participation. In our case, we saw that class participation did increase after the

policy was established. Each type of validity would highlight a different aspect of the relationship between our treatment (strict attendance policy) and our observed outcome (increased class participation).

Types of Validity:

There are four types of validity commonly examined in social research.

1. Conclusion validity asks is there a relationship between the program and the observed outcome? Or, in our example, is there a connection between the attendance policy and the increased participation we saw?

2. Internal Validity asks if there is a relationship between the program and the outcome we saw, is it a causal relationship? For example, did the attendance policy cause class participation to increase?

3. Construct validity is the hardest to understand in my opinion. It asks if there is there a relationship between how I operational zed my concepts in this study to the actual causal relationship I'm trying to study/? Or in our example, did our treatment (attendance policy) reflect the construct of attendance, and did our measured outcome - increased class participation - reflect the construct of participation? Overall, we are trying to generalize our conceptualized treatment and outcomes to broader constructs of the same concepts.

4. External validity refers to our ability to generalize the results of our study to other settings. In our example, could we generalize our results to other classrooms?

Threats to Internal Validity

There are three main types of threats to internal validity - single group, multiple group and social interaction threats.

Single Group Threats apply when you are studying a single group receiving a program or treatment. Thus, all of these threats can be greatly reduced by adding a control group that is comparable to your program group to your study.

A ***History Threat*** occurs when an historical event affects your program group such that it causes the outcome you observe (rather than your treatment being the cause). In our earlier example, this would mean that the stricter attendance policy did not cause an increase in class participation, but rather, the expulsion of several students due to low participation from school impacted your program group such that they increased their participation as a result.

A ***Maturation Threat*** to internal validity occurs when standard events over the course of time cause your outcome. For example, if by chance, the students who participated in your study on class participation all "grew up" naturally and realized that class participation increased their learning (how likely is that?) - That could be the cause of your increased participation, not the stricter attendance policy.

A ***Testing Threat*** to internal validity is simply when the act of taking a pre-test affects how that group does on the post-test. For example, if in your study of class participation, you measured class participation prior to implementing your new attendance policy, and students became forewarned that there was about to be an emphasis on participation, they may increase it simply as a result of involvement in the pretest measure - and thus, your outcome could be a result of a testing threat - not your treatment.

An ***Instrumentation Threat*** to internal validity could occur if the effect of increased participation could be due to the way in which that pretest was implemented.

A ***Mortality Threat*** to internal validity occurs when subjects drop out of your study, and this leads to an inflated measure of your effect. For example, if as a result of a stricter attendance policy, most students drop out of a class, leaving only those more serious students in the class (those who would participate at a high level naturally) - this could mean your effect is overestimated and suffering from a mortality threat.

The last single group threat to internal validity is a ***Regression Threat***. This is the most intimating of them all (just its name alone makes one panic). Don't panic. Simply put, a regression threat means that there is a tendency for the sample (those students you study for example) to score close to the average (or mean) of a larger population from the pretest to the posttest. This is a common occurrence, and will happen between almost any two variables that you take two measures of. Because it is common, it is easily remedied through either the inclusion of a control group or through a carefully designed research plan (this is discussed later). For a great discussion of regression threats, go to Bill Trochim's Center for Social Research Methods.

In sum, these single group threats must be addressed in your research for it to remain credible. One primary way to accomplish this is to include a control group comparable to your program group.

Multiple Group Threats to internal validity involve the comparability of the two groups in your study, and whether or not any other factor other than your treatment causes the outcome. They also (conveniently) mirror the single group threats to internal validity.

A ***Selection-History*** threat occurs when an event occurring between the pre and post test affects the two groups differently.

A ***Selection-Maturation*** threat occurs when there are different rates of growth between the two groups between the pre and post test.

Selection-Testing threat is the result of the different effect from taking tests between the two groups.

A ***Selection-Instrumentation*** threat occurs when the test implementation affects the groups differently between the pre and post test.

A ***Selection-Mortality*** Threat occurs when there are different rates of dropout between the groups which leads to you detecting an effect that may not actually occur.

Finally, a ***Selection-Regression*** threat occurs when the two groups regress towards the mean at different rates.

Diffusion or "Imitation of Treatment occurs when the comparison group learns about the program group and imitates them, which will lead to an equalization of outcomes between the groups.

Compensatory Rivalry means that the comparison group develops a competitive attitude towards the program group, and this also makes it harder to detect an effect due to your treatment rather than the comparison groups' reaction to the program group.

Resentful Demoralization is a threat to internal validity that exaggerates the posttest differences between the two groups. This is because the comparison group (upon learning of the program group) gets discouraged and no longer tries to achieve on their own.

Compensatory Equalization of Treatment is the only threat that is a result of the actions of the research staff - it occurs when the staff begins to compensate the comparison group to be "fair" in their opinion, and this leads to equalization between the groups and makes it harder to detect an effect due to your program.

Generally, the real difference between reliability and validity is mostly a matter of definition. Reliability estimates the consistency of your measurement, or more simply the degree to which an instrument measures the same way each time it is used in under the same conditions with the same subjects. Validity, on the other hand, involves the degree to which you are measuring what you are supposed to, more simply, the accuracy of your measurement. It is my belief that validity is more important than reliability because if an instrument does not accurately measure what it is supposed to, there is no reason to use it even if it measures consistently (reliably).

Reliability does not imply validity. That is, a reliable measure is measuring something consistently, but not necessarily what it is supposed to be measuring. For example, while there are many reliable tests of specific abilities, not all of them would be valid for predicting, say, job performance. In terms of accuracy and precision, reliability is

precision, while validity is accuracy. An example often used to illustrate the difference between reliability and validity in the experimental sciences involves a common bathroom scale. If someone who is 200 pounds steps on a scale 10 times and gets readings of 15, 250, 95, 140, etc., the scale is not reliable. If the scale consistently reads

Inter-Rater or Inter-Observer Reliability

When multiple people are giving assessments of some kind or are the subjects of some test, then similar people should lead to the same resulting scores. It can be used to calibrate people, for example those being used as observers in an experiment.

Inter-rater reliability thus evaluates reliability *across different people*.

Two major ways in which inter-rater reliability is used are (a) testing how similarly people *categorize* items, and (b) how similarly people *score* items.

This is the best way of assessing reliability when you are using observation, as observer bias very easily creeps in. It does, however, assume you have multiple observers, which is not always the case.

Inter-rater reliability is also known as *inter-observer reliability* or *inter-coder reliability*.

Examples

Two people may be asked to categorize pictures of animals as being dogs or cats. A perfectly reliable result would be that they both classify the same pictures in the same way.

Observers being used in assessing prisoner stress are asked to assess several 'dummy' people who are briefed to respond in a programmed and consistent way. The variation in results from a standard gives a measure of their reliability.

In a test scenario, an IQ test applied to several people with a true score of 120 should result in a score of 120 for everyone. In practice, there will be usually be some variation between people.

Whenever you use humans as a part of your measurement procedure, you have to worry about whether the results you get are reliable or consistent. People are notorious for their inconsistency. We are easily distractible. We get tired of doing repetitive tasks. We daydream. We misinterpret.

So how do we determine whether two observers are being consistent in their observations? You probably should establish inter-rater reliability outside of the context of the measurement in your study. After all, if you use data from your study to establish reliability, and you find that reliability is low, you're kind of stuck. Probably it's best to do this as a side study or pilot study. And, if your study goes on for a long time, you may want to reestablish inter-rater reliability from time to time to assure that your raters aren't changing.

There are two major ways to actually estimate inter-rater reliability. If your measurement consists of categories -- the raters are checking off which category each observation falls in you can calculate the percent of agreement between the raters. For instance, let's say you had 100 observations that were being rated by two raters. For each observation, the rater could check one of three categories. The other major way to estimate inter-rater reliability is appropriate when the measure is a continuous one. There, all you need to do is calculate the correlation between the ratings of the two observers. For instance, they might be rating the overall level of activity in a classroom on a 1-to-7 scale.

Test-Retest Reliability

An assessment or test of a person should give the same results whenever you apply the test.

Test-retest reliability evaluates reliability across *time*.

Reliability can vary with the many factors that affect how a person responds to the test, including their mood, interruptions, time of day, etc. A good test will largely cope with such factors and give relatively little variation. An unreliable test is highly sensitive to such factors and will give widely varying results, even if the person re-takes the same test half an hour later.

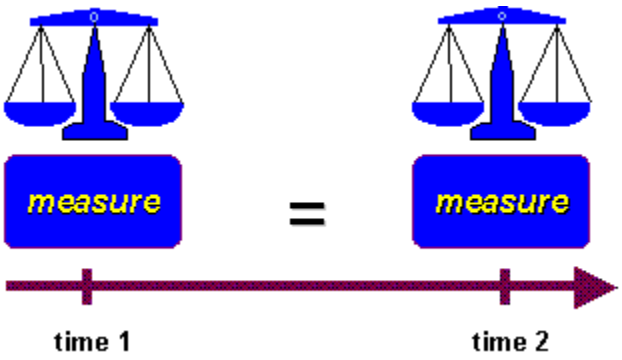
Generally speaking, the longer the delay between tests, the greater the likely variation. Better tests will give less retest variation with longer delays.

Examples

Various questions for a personality test are tried out with a class of students over several years. This helps the researcher determine those questions and combinations that have better reliability.

In the development of national school tests, classes of children are given several tests that are intended to assess the same abilities. A week and a month later, they are given the same tests. With allowances for learning, the variation in the test and retest results are used to assess which tests have better test-retest reliability.

We estimate test-retest reliability when we administer the same test to the same sample on two different occasions. This approach assumes that there is no substantial change in the construct being measured between the two occasions. The amount of time allowed between measures is critical. We know that if we measure the same thing twice that the correlation between the two observations will depend in part by how much time elapses between the two measurement occasions. The shorter the time gap, the higher the correlation; the longer the time gap, the lower the correlation. This is because the two observations are related over time -- the closer in time we get the more similar the factors that contribute to error. Since this correlation is the test-retest estimate of reliability, you can obtain considerably different estimates depending on the interval.



Parallel-Forms Reliability

One problem with questions or assessments is knowing what questions are the best ones to ask. A way of discovering this is do two tests in parallel, using different questions.

Parallel-forms reliability evaluates different questions and question sets that seek to assess the same construct.

Parallel-Forms evaluation may be done in combination with other methods, such as *Split-half,* which divides items that measure the same construct into two tests and applies them to the same group of people.

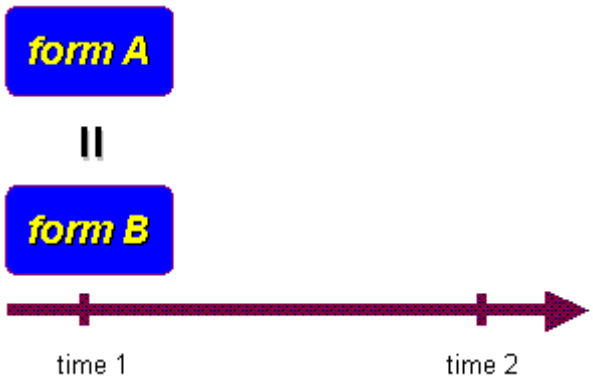
Examples

An experimenter develops a large set of questions. They split these into two and administer them each to a randomly-selected half of a target sample.

In development of national tests, two different tests are simultaneously used in trials. The test that gives the most consistent results is used, whilst the other (provided it is sufficiently consistent) is used as a backup.

In parallel forms reliability you first have to create two parallel forms. One way to accomplish this is to create a large set of questions that address the same construct and then randomly divide the questions into two sets. You administer both instruments to the same sample of people. The correlation between the two parallel forms is the estimate of reliability. One major problem with this approach is that you have to be able to generate lots of items that reflect the same construct. This is often no easy feat. Furthermore, this approach makes the assumption that the randomly divided halves are parallel or equivalent. Even by chance this will sometimes not be the case. The parallel forms approach is very similar to the split-half reliability described below. The major difference is that parallel forms are constructed so that the two forms can be used independent of each other and considered equivalent measures. For instance, we might be concerned about a testing threat to internal validity. If we use Form A for the pretest and Form B for the posttest, we minimize that problem. it would even be better if we randomly assign individuals to receive Form A or B on the pretest and then switch them on the posttest. With split-half reliability we have an instrument that we wish to use as a single measurement instrument and only develop randomly split halves for purposes of estimating reliability.

.



Internal Consistency Reliability

When asking questions in research, the purpose is to assess the response against a given construct or idea. Different questions that test the same construct should give consistent results.

Internal consistency reliability evaluates individual questions in comparison with one another for their ability to give consistently appropriate results.

Average inter-item correlation compares correlations between all pairs of questions that test the same construct by calculating the mean of all paired correlations.

Average item total correlation takes the average inter-item correlations and calculates a total score for each item, then averages these.

Split-half correlation divides items that measure the same construct into two tests, which are applied to the same group of people, then calculates the correlation between the two total scores.

Cronbach's alpha calculates an equivalent to the average of all possible split-half correlations and is calculated thus:

a = (N . r -bar) / (1 + (N -1) . r -bar)

Where N is the number of components,
and r -bar is the average of all Pearson correlation coefficients

In internal consistency reliability estimation we use our single measurement instrument administered to a group of people on one occasion to estimate reliability. In effect we

judge the reliability of the instrument by estimating how well the items that reflect the same construct yield similar results. We are looking at how consistent the results are for different items for the same construct within the measure. There are a wide variety of internal consistency measures that can be used.

Average Inter-item Correlation

The average inter-item correlation uses all of the items on our instrument that are designed to measure the same construct. We first compute the correlation between each pair of items, as illustrated in the figure. For example, if we have six items we will have 15 different item pairings (i.e., 15 correlations). The average inter item correlation is simply the average or mean of all these correlations. In the example, we find an average inter-item correlation of .90 with the individual correlations ranging from .84 to .95.

Average Item total Correlation

This approach also uses the inter-item correlations. In addition, we compute a total score for the six items and use that as a seventh variable in the analysis. The figure shows the six item-to-total correlations at the bottom of the correlation matrix. They range from .82 to .88 in this sample analysis, with the average of these at .85.

Split-Half Reliability

In split-half reliability we randomly divide all items that purport to measure the same construct into two sets. We administer the entire instrument to a sample of people and calculate the total score for each randomly divided half. the split-half reliability estimate, as shown in the figure, is simply the correlation between these two total scores. In the example it is .87.

Cronbach's Alpha

that's not how we compute it. Notice that when I say we compute all possible split-half estimates, I don't mean that each time we go a measure a new sample! That would take forever. Instead, we calculate all split-half estimates from the same sample. Because we measured our entire sample on each of the six items, all we have to do is have the computer analysis do the random subsets of items and compute the resulting correlations. The figure shows several of the split-half estimates for our six item example and lists them as SH with a subscript. Just keep in mind that although Cronbach's Alpha is equivalent to the average of all possible split half correlations we would never actually calculate it that way. Some clever mathematician (Cronbach, I presume!) figured out a way to get the mathematical equivalent a lot more quickly.

Comparison of Reliability Estimators

Each of the reliability estimators has certain advantages and disadvantages. Inter-rater reliability is one of the best ways to estimate reliability when your measure is an observation. However, it requires multiple raters or observers. As an alternative, you could look at the correlation of ratings of the same single observer repeated on two different occasions. For example, let's say you collected videotapes of child-mother interactions and had a rater code the videos for how often the mother smiled at the child. To establish inter-rater reliability you could take a sample of videos and have two raters code them independently. To estimate test-retest reliability you could have a single rater ple of videos and

they yielded consistent results. If you get a suitably high inter-rater reliability you could then justify allowing them to work independently on coding different videos. You might use the test-retest approach when you only have a single rater and don't want to train any others. On the other hand, in some studies it is reasonable to do both to help establish the reliability of the raters or observers.

The parallel forms estimator is typically only used in situations where you intend to use the two forms as alternate measures of the same thing. Both the parallel forms and all of the internal consistency estimators have one major constraint -- you have to have multiple items designed to measure the same construct. This is relatively easy to achieve in certain contexts like achievement testing (it's easy, for instance, to construct lots of similar addition problems for a math test), but for more complex or subjective constructs this can be a real challenge. If you do have lots of items, Cronbach's Alpha tends to be the most frequently used estimate of internal consistency.

The test-retest estimator is especially feasible in most experimental and quasi-experimental designs that use a no-treatment control group. In these designs you always have a control group that is measured on two occasions (pretest and posttest). the main problem with this approach is that you don't have any information about reliability until you collect the posttest and, if the reliability estimate is low, you're pretty much sunk.

Each of the reliability estimators will give a different value for reliability. In general, the test-retest and inter-rater reliability estimates will be lower in value than the parallel forms and internal consistency ones because they involve measuring at different times or with different raters. Since reliability estimates are often used in statistical analyses of quasi-experimental designs (e.g., the analysis of the nonequivalent group design), the fact that different estimates can differ considerably makes the analysis even more complex.

Reliability Testing

Reliability testing is the cornerstone of a reliability engineering program. It provides the most detailed form of reliability data because the conditions under which the data are collected can be carefully controlled and monitored. Furthermore, reliability tests can be

designed to uncover particular suspected failure modes and other problems. The type of reliability testing a product undergoes will change along different points of its life cycle, but the overriding goal is to insure that data from all or most of the tests were generated under similar enough conditions so that an "apples to apples" comparison can be made of the product's reliability characteristics at different points in the product's life. It is for this reason that consistent and thorough reliability specifications and a standard definition of failure are up-front requirements to implementing reliability testing.

A properly designed series of tests, particularly during the product's earlier design stages, can generate data that would be useful in the implementation of a reliability growth tracking program. This will provide information that will be helpful in making management decisions regarding scheduling, development cost projections and so forth. This information will also be useful in planning the development cycle of future products.

There are several different kinds of tests, including: *Customer Usage Profiling, Development Testing and Manufacturing Testing*

Customer Usage Profiling

An important requirement for designing useful reliability tests is to have a good idea of how the product is actually going to be used in the field. The tests should be based on a realistic expectation of the customer usage, rather than estimates or "gut feelings" about the way the customer will use the product. Tests based on mere speculation may result in a product that has not been rigorously tested and consequently may run into operational difficulties due to use stress levels being higher than anticipated. On the other hand, tests that are designed with a strong basis of information on how the product will be used will be more realistic and result in an optimized design that will exhibit fewer failures in field.

Customer usage profiles can be designed to actively gather information on how the customers are actually using an organization's product. This design can range from a simple questionnaire to sophisticated instrumentation within the product that feeds back detailed information about its operation. An incentive is often useful to get customers to sign on for a usage measurement program, particularly if it is an intrusive process that

.

involves the installation of data collection equipment. Additionally, customers are often eager to participate in these programs in the knowledge that the information that they provide will ultimately result in a more reliable and user-friendly product.

Developmental Testing

Developmental testing occurs during the early phases of the product's life cycle, usually from project inception to product design release. It is vital to be able to characterize the reliability of the product as it progresses through its initial design stages so that the reliability specifications will be met by the time the product is ready for release. With a multitude of design stages and changes that could affect the product's reliability, it is necessary to closely monitor how the product's reliability grows and changes as the product design matures. There are a number of different test types that can be run during this phase of a product's life cycle to provide useful reliability information:

- **Component-level Testing**: Although component-level testing can continue throughout the development phase of a product, it is most likely to occur very early in the process. This may be due to the unavailability of parts in the early stages of the development program. There may also be special interest in the performance of a specific component if it has been radically redesigned or if there is a separate or individual reliability specification for that component. In many cases, component-level testing is undertaken to begin characterizing a product's reliability even though full system-level test units are unavailable or prohibitively expensive. However, system-level reliability characterization can be achieved through component-level testing. This is possible if sufficient understanding exists to characterize the interaction of the components. If this is the case, the system-level reliability can be modeled based on the configuration of components and the result of component reliability testing.
- **System-level Testing**: Although the results of component-level tests can be used to characterize the reliability of the entire system, the ideal approach is to test the entire system, particularly if that is how the reliability is specified. That is, if the technical specifications call out a reliability goal for a specific system or

configuration of components, that entire system or configuration should be tested to compare the actual performance with the stated goal. Although early system-level test units may be difficult to obtain, it is advisable to perform reliability tests at the system level as early in the development process as possible. At the very least, comprehensive system-level testing should be performed immediately prior to the product's release for manufacturing in order to verify design reliability. During such system-level reliability testing, the units under test should be from a homogeneous population and should be devoted solely to the specific reliability test. The results of the reliability test could be skewed or confounded by "piggybacking" other tests along with it and this practice should be avoided. A properly conducted system-level reliability test will be able to provide valuable engineering information above and beyond the raw reliability data.

- **Environmental and Accelerated Testing**: It may be necessary in some cases to institute a series of tests in which the system is tested at extreme environmental conditions or with other stress factors accelerated above the normal levels of use. It may be that the product would not normally fail within the time constraints of the test and, in order to get meaningful data within a reasonable time, the stress factors must be accelerated. In other cases, it may be necessary to simulate different operating environments based on where the product will be sold or operated. Regardless of the cause, tests like these should be designed, implemented and analyzed with care. Depending on the nature of the accelerating stress factors, it is easy to draw incorrect conclusions from the results of these tests. A good understanding of the proper accelerating stresses and the design limits of the product are necessary to be able to implement a meaningful accelerated reliability test. For example, one would not want to design an accelerated test that would overstress the product and introduce failure modes that would not normally be encountered in the field. Given that there have been a lot of incredible claims about the capability of accelerated testing and the improbably high acceleration factors that can supposedly be produced, care needs to be taken when setting up this type of reliability testing program.

Shipping Tests: Although shipping tests do not necessarily qualify as reliability tests per se, shipping tests or simulations designed to test the impact on the product of shipping and handling should be a part of the reliability testing program. This is because the effects of shipping will often have an impact on the reliability of the product as experienced by the customer. As such, it may be useful to incorporate shipping tests alongside the normal reliability testing. For example, it may be a good idea to put the units of a final design release reliability test through a non-destructive shipping test prior to the actual reliability testing in order to better simulate actual use conditions.

Manufacturing Testing

The testing that takes place after a product design has been released for production generally tends to measure the manufacturing process rather than the product, under the assumption that the released product design is final and good. However, this is not necessarily the case, as post-release design changes or feature additions are not uncommon. It is still possible to obtain useful reliability information from manufacturing testing without diluting any of the process-oriented information that these tests are designed to produce.

- **Functionality Testing and Burn-In:** This type of testing usually falls under the category of operation verification. In these tests, a large proportion, if not all, of the products coming off of the assembly line are put on a very short test in order to verify that they are functioning. In some situations, they may be run for a predetermined "burn-in" time in order to weed out those units that would have early infantile failures in the field. Although it may not be possible to collect detailed reliability information from this type of testing, what is lost in quality is made up for in quantity. With the proper structuring, these tests can provide a fairly good picture of early-life reliability behavior of the product.
- **Extended Post-Production Testing**: This type of testing usually gets implemented near the end or shortly after the product design is released to production. It is useful to structure these types of tests to be identical to the final reliability verification tests conducted at the end of the design phase. The purpose of

these tests is to assess the effects of the production process on the reliability of the product. In many cases, the test units that undergo reliability testing prior to the onset of actual production are hand-built or carefully adjusted prior to the beginning of the reliability tests. By replicating these tests with actual production units, potential problems in the manufacturing process can be identified before many units are shipped.

Design/Process Change Verification: This type of testing is similar to the extended post-production testing in that it should closely emulate the reliability verification testing that takes place at the end of the design phase. This type of testing should occur at regular intervals during production or immediately following a post-release design change or a change in the manufacturing process. These changes can have a potentially large effect on the reliability of the product and these tests should be adequate, in terms of duration and sample size, to detect such changes.

.