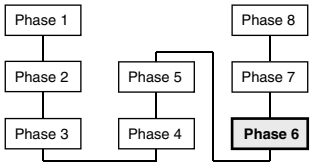
8

Product Performance and Production



8.1 Introduction

This chapter deals with reliability performance during phase 6 (stage III, level III) of the product life cycle. Phase 6 deals with production – and reliability performance is here the reliability of the items produced. As such, this phase is relevant mainly for standard products (for large volume production of items, such as cellular phones, domestic appliances, cars) and in some cases for custom-built products (for small volume production of items, such as ships, airplanes). Obviously, this is not relevant for a single custom-built product.

Production is the process of transforming inputs (raw materials, components) into finished products. The process is complex and can involve several sub-phases, with one or more operations carried out in each sub-phase. The reliability of the products produced (AP-III) will, in general, differ from that predicted by the design (PP-III of phase 3) or the prototype produced under strict laboratory conditions (PP-III of phase 4) and, in fact, is almost always lower. This difference is due to variability in the operations and in the inputs. Through proper quality control of the inputs and operations, we try to ensure that AP-III is kept close to DP-III. This chapter deals with variability issues that can cause AP-III to deviate from DP-III and quality control techniques to ensure that this does not occur.

The outline of the chapter is as follows. Section 8.2 deals with phase 6 for standard products. Section 8.3 looks at production process and occurrence of non-conforming items, Section 8.4 discusses the effect of quality variations on product reliability, and Section 8.5 deals with testing during production. Section 8.6 looks at alternative approaches to quality control and Section 8.7 deals with optimal quality control. The chapter concludes with a case study on cellular phones in Section 8.8.

8.2 Phase 6 for Standard Products

Figure 8.1 shows the key elements that influence the reliability (AP-III) of the produced items. Before production can commence, the manufacturer needs to design the production process. This involves providing the equipment and the resources needed to carry out the operations (e.g., casting, soldering, heating) at several sub-phases. The number of sub-phases needed depends on the complexity of the product. The production starts at sub-phase J (component level) and proceeds through several sub-phases leading to the sub-phase where the final operations (involving subsystems) take place to assemble the final product. The process state characterizes the condition of the equipment (e.g., cutting tool) and/or the settings (e.g., the temperature of the soldering element). The design process involves defining the acceptable limits for the states for the different equipment and the settings. When the process state is within these limits, it is said to be *in control*. With age and operations, the state will tend to degrade (e.g., cutting tool becoming blunt due to wear, and temperature setting deviating) and can go outside the specified limits. When this happens, the process is said to be *out of control*.

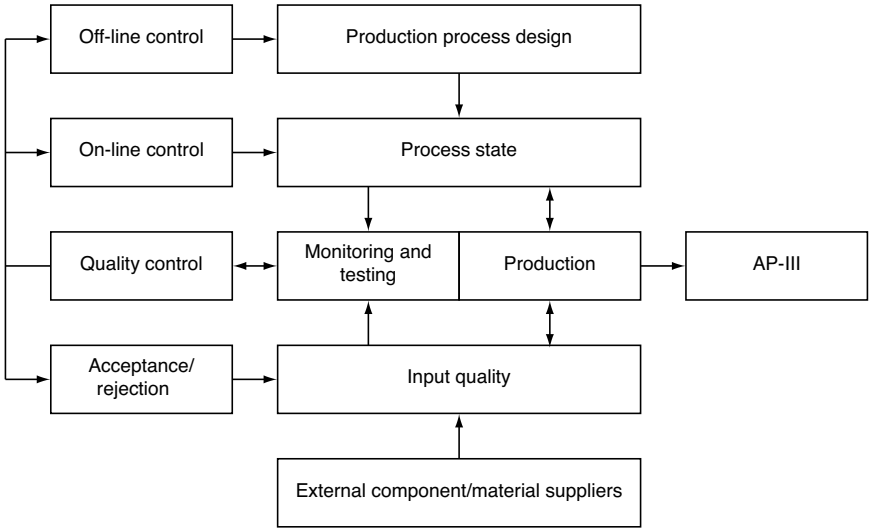


Figure 8.1. Key elements of phase 6

Items that meet the design reliability (and/or other) performance are called *conforming* items and those that do not, are called *non-conforming* items. Ideally, there should be no non-conforming item produced when the process is in control. However, this is an ideal case and is seldom true. In general, a very small fraction of the items produced can be non-conforming. When the process goes out of control, the fraction of non-conforming items produced increases significantly and this has serious implications for the manufacturer. Non-conforming items are less reliable than

conforming items and, as a result, lead to high customer dissatisfaction and warranty costs. Similarly, when the inputs (components or material obtained from external suppliers) do not meet the specification, the items produced are non-conforming even with the process being in control.

The variation in the output (conforming versus non-conforming) can be categorized into two groups:

1. Variation due to uncontrollable causes. By and large, nothing can be done about this source of variation, except to modify the process.
2. Variation due to assignable causes. This is due to a process going out of control, inputs not conforming to specifications and/or errors by human operators. This can be controlled through effective quality control schemes and process modifications (e.g., machine adjustment, replacement of worn out parts, and additional training of operators).

Quality control deals with controlling the variation resulting from assignable causes. In the case of process control, the approaches used can be broadly grouped into two categories (i) *on-line* and (ii) *off-line*. In the case of input quality control, a variety of techniques for accepting/rejecting have been proposed where the aim is to reject bad batches (that contain a high fraction of non-conforming components or the material properties do not meet the required specifications) and accept good batches (that contain a very small fraction of non-conforming components or material properties meet the required specifications).

To ensure that the actual performance AP-III matches the desired performance DP-III requires a proper quality control plan. Figure 8.2 shows the key elements of such a plan. It involves periodic sampling and testing of items (at component, final product, and at one or more intermediate sub-phases), monitoring of the process (inspection of equipment) and operations, to draw inference about input quality, process state and the actual performance AP-III. Inference is drawn on whether AP-III is in agreement with DP-III or not. If they are in agreement, the process continues with no change. If not, a root cause analysis is used to identify the cause and the corrective action needed. These different actions are indicated in Figure 8.2 and can involve on-line or off-line quality control, renegotiating with the supplier or even changing supplier.

8.3 Production Process and Occurrence of Non-conforming Items

The type of manufacturing process used depends on the demand for the product and is determined by economic considerations. If the demand is high, then it is economical to use a continuous production process. If the demand is medium, then it is more economical to use a batch production process, where items are produced in lots (or batches). Finally, if the demand is low, then flexible manufacturing is used.

In all cases, the state of the manufacturing process has a significant impact on the occurrence of non-conforming items. As discussed earlier, the process state can

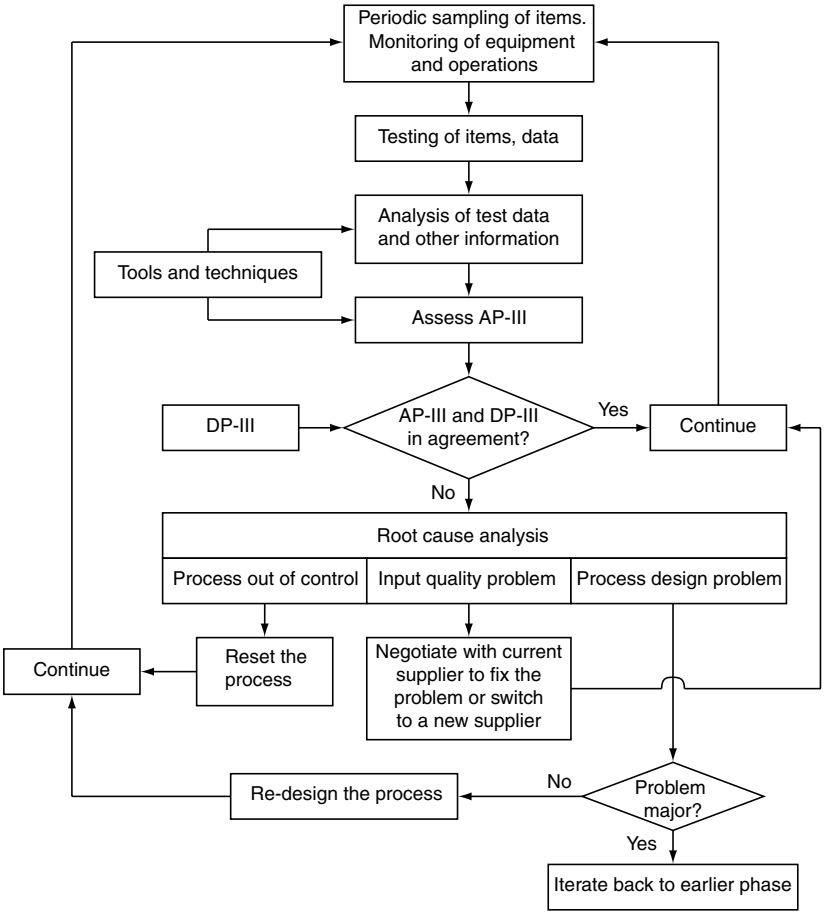


Figure 8.2. Quality control plan

be modelled as being in one of two possible states: (i) in control and, (ii) out of control. When the process state is in control, all the assignable causes are under control and the probability that an item produced is non-conforming is very small. Usually, this probability can be as small as 10^{-3} to 10^{-6} for properly designed production processes. The process state changes from in control to out of control due to one or more of the process parameters being no longer within the required interval. This increases the probability that an item is non-conforming.

8.3.1 Modelling Occurrence of Non-conforming Items

The modelling of the occurrence of non-conforming items depends on the type of production process. We consider both continuous and batch production. Let ϕ_0 and

ϕ_1 denote the probability that an item produced is conforming when the process is in control and out of control, respectively. In general, $\phi_0 \gg \phi_1$.[1]

Continuous Production

The production process starts in control and after a random length of time it changes to out of control. Once the process state changes from in control to out of control, it remains in that state until it is brought back to in control through some corrective action. The fraction of conforming items produced depends on the relative fractions of time the process is in control and out of control, respectively. This depends on the control plan to detect the change in process state. The change from in control to out of control can be either gradual or sudden as indicated in Figure 8.3.

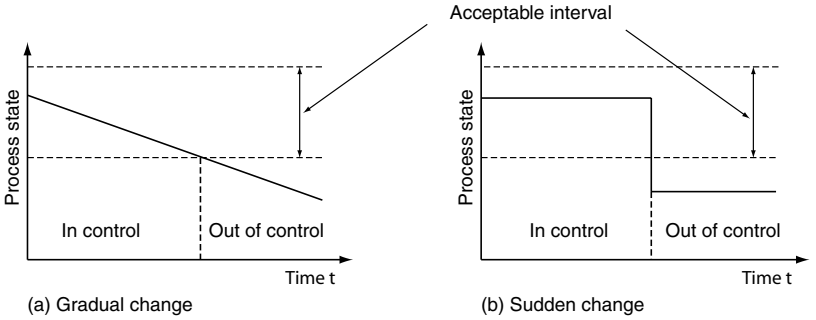


Figure 8.3. Change in process state

Batch Production

Let Q denote the lot size. At the start of each lot production, the process state is checked to ensure that it is in control. If the process state is in control at the start of the production of an item, it can change to out of control with probability $1 - \nu$, or stay in control with probability ν. Once the state changes to out of control, it remains there until completion of the lot.

8.4 Effect of Quality Variations on Reliability Performance

We look at the case where the reliability performance of the product is modelled by a ROCOF function. Let the desired ROCOF (obtained from phase 3) be as indicated in Figure 8.4. This reliability is achieved through specifying the desired reliabilities

[1] Porteus (1986) considers the extreme case, $\phi_0 = 1$, implying that all items produced are conforming when the state is in control, and $\phi_1 = 0$, implying that all items produced are non-conforming when the process is out of control. Djamaludin et al. (1994) consider the general case where $0 < \phi_0 \leq 1$, $0 \leq \phi_1 < 1$ and $\phi_0 > \phi_1$.

for the various components of the product. Let $R_c(t)$ denote the desired reliability function for some conforming component (an element of DP-III). The associated failure distribution function is given by $F_c(t) = 1 - R_c(t)$.

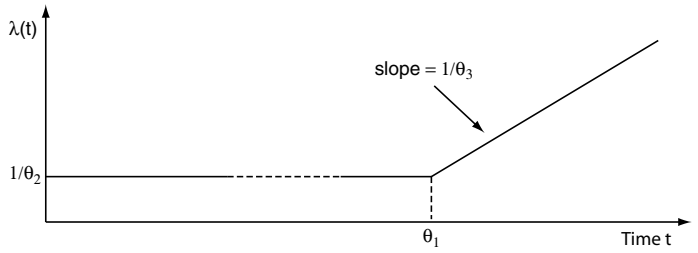


Figure 8.4. Desired ROCOF (from phase 3)

8.4.1 Variation in Component Quality

Let $F_n(t)$ denote the failure distribution of a non-conforming component and let p denote the probability that the component is non-conforming. Note that $F_n(t) > F_c(t)$, implying that the reliability performance of a non-conforming item is inferior to that of a conforming item. Then, the actual failure distribution function of the component (produced or bought from an external source) can be modelled by

$$F_a(t) = (1 - p)F_c(t) + pF_n(t) \tag{8.1}$$

The actual component reliability, $R_a(t) = 1 - F_a(t)$, is an element of AP-III.

The effect of component non-conformance on the ROCOF is as shown in Figure 8.5. When $p = 0$, the hump disappears as the ROCOF is unaffected. As p increases, the hump becomes more pronounced.

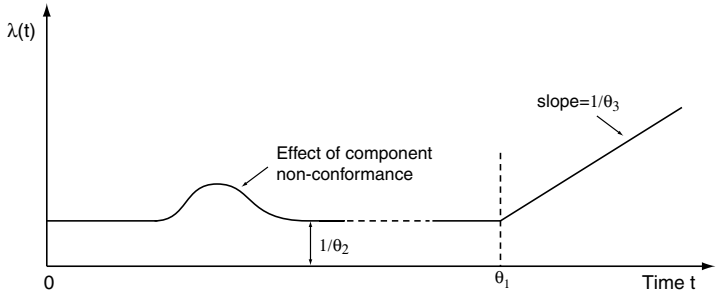


Figure 8.5. Effect of component non-conformance on ROCOF

8.4.2 Variations in Assembly Operations

Even with all the components conforming to design specification, an item can fail early due to chance errors in the assembly operations (e.g., misalignment, dry solder joint). This failure can be viewed as a new failure mode with a decreasing failure rate. The effect of this failure mode on the ROCOF is as shown in Figure 8.6. As can be seen, this actual ROCOF has a higher value over the initial life of items indicating a higher likelihood of an early failure due to assembly errors which decreases with time. As a result, the ROCOF has a "bathtub" shape. These early failures are also

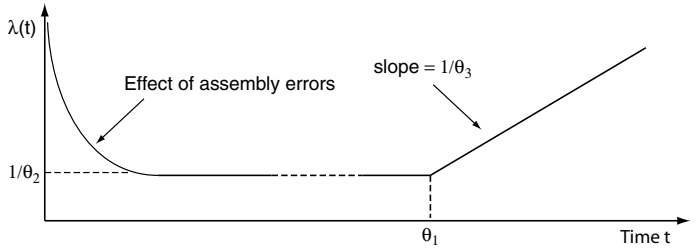


Figure 8.6. Effect of assembly error on ROCOF

termed "teething problems" or "infant mortality" and are detected within a relative short time period after an item is put into operation. In this case, the actual ROCOF, $\lambda_a(t)$ (an element of AP-III), is given by

$$\lambda_a(t) = \lambda_d(t) + \lambda_e(t) \tag{8.2}$$

where $\lambda_d(t)$ is the desired ROCOF (an element of DP-III) and $\lambda_e(t)$ is the failure rate associated with the failure distribution for the new failure mode with $\lambda_e(0) = q$. A higher value of q corresponds to a higher likelihood of failure from assembly errors.

8.4.3 Combined Effects of Component Non-conformance and Assembly Errors

Since most products are complex involving several components and many different assembly operations, the net effect of quality variation on the actual ROCOF is as indicated in Figure 8.7. The shape of the ROCOF is often referred to as the "roller coaster" shape (Wong, 1989). Note that there can be several humps.

8.5 Testing During Production

The purpose of testing during manufacturing is to eliminate assembly errors, defects, and early component failures. The type of testing to be done depends on the product (electrical, mechanical or electronic). For very expensive products (e.g., commercial satellites) requiring a very high level of reliability, 100% testing would be employed.

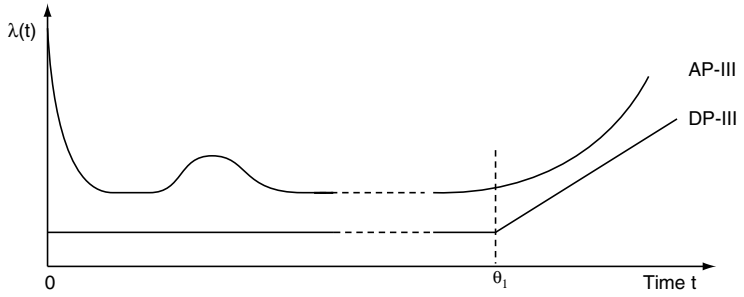


Figure 8.7. Total effect of quality variations on ROCOF

For most other products (particularly consumer durables), only a very small fraction is tested.

The level of testing and the type of tests can change over the period of production. For a new product, in the early stages of production, considerable testing is required to establish the process characteristics and the effect of process parameters on the reliability of the product. As the product matures, the testing requirements are reduced. Testing is also done under various environmental conditions and often the tests are carried in an accelerated mode to reduce the time needed for testing. Two types of testing used are

1. Environmental stress screening
2. Burn-in

Environmental stress screening involves subjecting an item (component or assembly) to various environmental extremes to identify and eliminate manufacturing defects prior to customer use. Typical methods used are temperature cycling, random vibrations, electrical stress, thermal stress, and so on.

Burn-in is a process used to improve outgoing quality and is discussed in a later section.

Note. In some cases the testing takes very little time (e.g., measuring some physical characteristic such as tolerances or solder joint). In other cases, such as life testing, the test duration needs to be defined. In this case the test data is a combination of failure and censored data. For very highly reliable items we tend to use accelerated degradation tests. In this case, the test data is the condition (e.g., wear) of the item under test and from this we infer the useful life of the item.[2]

8.6 Quality Control

The main aim of quality control is to ensure that the effect of quality variations is small so that the actual performance AP-III is fairly close to the desired performance

[2] For more details, see, for example Rausand and Høyland (2004); Nelson (1990).

DP-III. This implies ensuring that p in (8.1) and $q(=\lambda_e(0))$ in (8.2) are below some specified upper limits. For process variations, the control can be achieved through off-line control and/or on-line control. For input quality variations, the control is achieved through proper rules for accepting/rejecting.

8.6.1 Off-line Control of Production Process

As mentioned earlier, a production process is affected by several factors – some controllable and others not. Taguchi (1986) proposed a method for determining optimal settings for the controllable factors, taking into account the influence of the uncontrollable factors. The method uses well-known concepts from design of experiments combined with the concept of "signal-to-noise" ratio from electrical communication engineering. Since the pioneering work of Taguchi, there has been considerable development in the design of optimal and robust manufacturing processes.[3]

8.6.2 On-line Control of Production Process

The aim of on-line control is to prevent the occurrence of non-conforming items through actions that attempt to ensure that the process is in control during the production run. The approach used depends on the type of production process.

Continuous Production

In continuous production, the process begins in control and may change to out of control with the passage of time. The aim is to detect the change and bring the process back in control as fast as possible. Control charts are used for this purpose.

The underlying principle of a control chart is simple. Samples of items are taken periodically and the sample statistics (e.g., sample mean, sample standard deviation, number or fraction non-conforming) are plotted. When the process is in control, the sample statistics should assume values within some specified interval with high probability. When the process goes out of control, it is more likely to assume values outside the specified interval. As such, plotting of the sample statistics provides a means for detecting the change in process state using some rules. Many different rules have been proposed for different charts.

Control charts can be grouped into two broad categories:

Variable charts: These are based on continuous-valued measurements (e.g., physical dimension, hardness).

Attribute charts: These are based on integer-valued measurements (e.g., counts of flaws, such as the number of dry solder joints).

There are many different variable and attribute charts. The more commonly used variable charts are (i) $\bar{X}$ chart, (ii) R chart, and (iii) CUSUM chart, and for the

[3] Details can be found in many books; see for example, Dehnad (1989), Moen et al. (1991), and Peace (1993).

attribute charts, the most commonly used charts are (i) p chart and (ii) np chart. We briefly discuss the $\bar{X}$ chart.[4] The control chart does not indicate the cause for the change in the state of the process and we need to use tools, such as root cause analysis, to determine the cause.

$\bar{X}$ Chart

Here the variable being observed is assumed to be normally distributed with mean μ and variance σ^2 when the process is in control. When the process goes out of control, either the mean changes and/or the variance increases. A sample of size n is taken at regular intervals. Let x_{ji} denote the observed value for the ith item in sample j. Note that $j = 1, 2, \ldots$ and $i = 1, 2, \ldots, n$ for each j. The sample mean for sample j is given by

$$\bar{x}_j = \frac{1}{n} \sum_{i=1}^{n} x_{ji} \tag{8.3}$$

This statistic is plotted on the control chart. The chart has a centre line and two control lines. The centre line is a horizontal line corresponding to the nominal mean μ or the overall sample mean (the average of the $\bar{x}_j$'s). The two control lines (or control limits) are parallel to the centre line and at a distance $3\sigma/\sqrt{n}$ on either side of the centre line. The two warning lines (or warning limits) are similarly drawn, but at a distance $2\sigma/\sqrt{n}$ on either side of the centre line, as indicated in Figure 8.8.

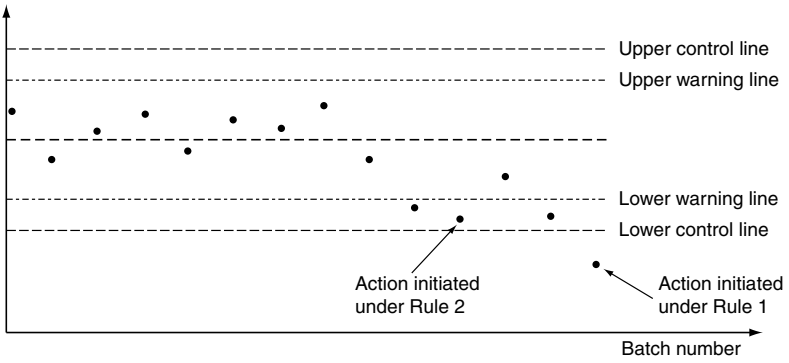


Figure 8.8. Typical control chart

As mentioned earlier, if the sample plots lie within the control limits, we can conclude that the process is in control with high probability and no action is needed. When one or more sample plots fall outside the limits, this is taken as an indicator

[4] Most books on quality control discuss the different charts; see for example, Grant and Leavensworth (1988), Montgomery (1985), Sinha and Willborn (1985), Ryan (1989), and Evans and Lindsay (1996).

of a change in the process state from in control to out of control. In fact, many rules have been developed to determine when a process is out of control and action is to be taken. Two such rules are:

Rule 1: A single point falling outside the control limits.
Rule 2: Two out of three points in a row falling above or below the warning line.

The instant at which action is initiated depends on the rule as indicated in Figure 8.8.

When it is inferred that the process is out of control, the process is stopped to check if a change has indeed occurred. If so, corrective actions to restore it to its in control state are initiated. If not, no corrective action is needed and production resumes. For any given stopping rule, there are two types of errors (or wrong decisions):

Type 1 error: A false alarm that leads to a stoppage of the production when the process is in control.
Type 2 error: Corrective action not being initiated for a certain length of time subsequent to the process changing from in control to out of control.

Ideally we would like to have the probabilities of both of these wrong decisions to be zero. Unfortunately, this is not possible and the probabilities depend on the rules used for initiating corrective actions.[5]

Batch Production

As indicated earlier, in batch production, the process starts in control and can go to out of control during the production of an item with probability $1-\nu$. This affects the number of non-conforming items in the lot. Let N_c denote the number of conforming items in a lot. This is a random variable which can take on integer values in the interval $[0, Q]$. The expected fraction of conforming items in a lot of size Q is given by[6]

$$\phi(Q) = \frac{\nu(\phi_0 - \phi_1)(1 - \nu^Q)}{(1-\nu)Q} + \phi_1 \tag{8.4}$$

where ϕ_0 and ϕ_1 denote the probability that an item produced is conforming when the process is in control and out of control, respectively.

It is easily seen that $\phi(Q)$ is a decreasing sequence in Q implying that the expected fraction of conforming items in a batch decreases as the batch size (Q) increases. This implies that the smaller the lot size, the better the outgoing quality. Finally, two special cases are:

$$Q = 1: \quad \phi(Q) = \phi_0$$
$$Q = \infty: \quad \phi(Q) = \phi_1$$

[5] There is a vast literature dealing with the determination of optimal rules that take into account the economic consequences of these two types of errors. These can be found in most books on statistical quality control.

[6] For details of the derivation, see Chapter 13 of Blischke and Murthy (2000).

In the latter case, only a finite number of items are produced with the process being in control and an infinite number produced with the state being out of control.

8.6.3 Weeding Out Non-conforming Components

The aim of weeding is to essentially detect non-conforming items through inspection and testing of each item. Testing can be done either at the component level, product level and/or at one or more of the intermediate levels of the production process. Detection of non-conformance at the earliest possible instant is desirable, as this allows for immediate corrective action. Some times a non-conforming item can be transformed into a conforming item by reworking it, and at other times, the item needs to be scrapped.

The quality of inspection and testing is another issue that needs to be considered. If inspection and testing are perfect, then every non-conforming item tested is detected. With imperfect testing and inspection, not only may a non-conforming item not be detected, but a conforming item may be classified as non-conforming. As a result, the outgoing quality (the fraction of conforming items) depends on the level of testing and the quality of testing.

Component Level Weeding

When there is quality variation at the component level, the actual failure distribution of component lifetimes is given by (8.1). Weeding involves putting each component on a test bed so that it becomes operational. The duration of the test is τ. Components that fail during testing are scrapped. The rationale for this is that non-conforming items are more likely to fail than conforming items and hence are weeded out.

The probability that a conforming (non-conforming) component will fail during testing for a period τ is given by $F_c(\tau)$ $[F_n(\tau)]$. As a result, the probability that a component that survives the test is non-conforming is given by

$$p_1 = \frac{pR_n(\tau)}{(1-p)R_c(\tau) + pR_n(\tau)} \tag{8.5}$$

where $R_c(\tau) = 1 - F_c(\tau)$ and $R_n(\tau) = 1 - F_n(\tau)$ denote the survivor functions of conforming and non-conforming components, respectively. Since $R_c(\tau) > R_n(\tau)$, we have $p_1 < p$. The failure distribution of an item that survives the test is given by

$$\widetilde{F}_c(t) = (1-p_1)\widetilde{F}_c(t) + p_1\widetilde{F}_n(t) \tag{8.6}$$

where $\widetilde{F}_c(t)$ and $\widetilde{F}_n(t)$ are given by

$$\widetilde{F}_c(t) = \frac{F_c(t+\tau) - F_c(\tau)}{1 - F_c(\tau)} \tag{8.7}$$

and

$$\widetilde{F}_n(t) = \frac{F_n(t+\tau) - F_n(\tau)}{1 - F_n(\tau)} \tag{8.8}$$

for $t \geq 0$. Note that as τ increases, p_1 (the probability that an item released is non-conforming) decreases, and hence the outgoing quality is improved. However, this is achieved at the expense of the useful life of conforming items released being reduced by an amount τ.

8.6.4 Acceptance Sampling

The input material (raw materials and components) is obtained from external suppliers in batches. The quality for raw material is defined through some characteristics (e.g., strength, chemical composition) and for components it is the reliability. The quality of input material can vary from batch to batch. A batch is defined to be unacceptable if the quality does not meet the specified value (e.g., mean time to failure, or the fraction or number of conforming items in the batch, is below some specified value). Such batches need to be rejected. Batches for which the quality meets or exceeds the specified value (e.g., mean time to failure, or the fraction or number of conforming items, is above some specified value) are to be accepted.

The decision to accept or reject a batch is based on testing a small sample from the batch. This is known as acceptance sampling. A variety of acceptance sampling schemes for attribute (integer valued measurements) and variables (continuous-valued measurements) can be found in the literature. They can be divided into three groups: (i) single-sampling, (ii) multiple-sampling, and (iii) sequential-sampling.[7]

The single sampling by attribute plan involves taking a random sample of n items from a lot of size N. Let d (called the sample number) denote the number of items that are non-conforming (e.g., have a defect or fail during the test period). This is compared with a pre-specified number c (called the acceptance number) to decide whether to accept or reject a batch. If $d \leq c$, the batch is accepted and if $d > c$, the batch is rejected.

In a double-sampling scheme, a first sample of size n_1 is drawn from the batch and tested. Let d_1 denote the number of non-conforming items. The outcome action is as follows:

If $d_1 \leq a_1$, the batch is accepted.
If $d_1 > r_1$, the batch is rejected.
If $a_1 < d_1 \leq r_1$, a second sample is drawn.

The second sample involves drawing randomly n_2 items from the batch. Let d_2 denote the number of non-conforming items in the sample. The outcome action is as follows:

If $d_1 + d_2 \leq a_2$, the batch is accepted.
If $d_1 + d_2 > a_2$, the batch is rejected.

The extension to multiple sampling is a natural extension of this and can involve drawing more than two samples.

[7] Most books on quality control discuss some of the sampling schemes. A detailed discussion of the different schemes can be found in Schilling (1982)

In sequential sampling, the sample size is one and the outcome (accept, reject or continue sampling) is decided after each sample is tested.

As in the case of control charts, we can make two types of errors.

Type 1 error: Rejecting a batch that should have been accepted.
Type 2 error: Accepting a batch that should have been rejected.

Ideally, we would like to have the probabilities of both of these wrong decisions to be zero. Unfortunately, this is not possible. In the case of single-sample scheme, the probabilities depend on the parameters d and c (and the duration of test in the case of life testing).

8.6.5 Sub-set Selection

A manufacturer can often select the component supplier from several component manufacturers. The reliability of the components differs across component manufacturers and the problem facing the manufacturer is to select the best component supplier. This problem can be posed as selecting the best population from a collection of populations and is called the sub-set selection problem. In order to do this, we need to define more precisely the notion of "best" and several different notions have been proposed and studied.[8]

8.6.6 Burn-in

When variations in assembly operations are significant, the ROCOF function for the product has the bathtub shape (see Figure 8.6). Let $\lambda(t)$ denote this bathtub function with $\lambda(t)$ decreasing for $0 \leq t \leq t_1$. As a result, if produced items are released without any further action, a high fraction would fail in the early period, leading to high warranty costs and loss of customer goodwill. In this case, burn-in can be used to improve product reliability by consuming a part of the lifetime. The approach is to test each item for a period τ prior to its sale. Any items that fail within this period are minimally repaired. If the time to repair is small (in relation to τ), so that it can be ignored, then the ROCOF function is unaffected by failure and repair action.

Let the ROCOF function after burn-in be given by $\widetilde{\lambda}(t)$. Then $\widetilde{\lambda}(t) = \lambda(t + \tau)$ for $t \geq 0$. By choosing $\tau = t_1$, $\widetilde{\lambda}(t)$ is no longer bathtub shaped and has a shape similar to that shown in Figure 8.4. For more on burn-in, see Jensen and Peterson (1982).

Burn-in results in additional costs due to (i) fixed set-up cost of the burn-in facility, (ii) variable cost (which increases with τ) for testing each item, and (iii) rectification cost for failures during burn-in. Hence, burn-in is worthwhile only if its benefits (measured in terms of the improvements in reliability) exceed the cost of burn-in.

[8] There is an extensive literature on this topic. The earlies paper is by Rademaker and Antle (1975) and looks at the optimal sample size to decide which of the two populations have the larger reliability. Kingston and Patel (1980a,b) focus on selecting the population with the largest reliability based on type II censoring. The shape parameter can be the same or different and needs to be estimated. Later papers include Hsu (1982), Sirvanci (1986), Gupta and Miescke (1987), and Gill and Mehta (1994).

8.7 Optimal Quality Control Effort

The quality (defined in terms of conformance) of items produced depends on the quality control effort. The quality improves as the effort is increased, but this is achieved at the expense of increased quality control cost.

In the case of batch production, the quality $\phi(Q)$ increases as the batch size Q increases. This implies that the smaller the lot size, the better the outgoing quality. On the other hand, the size of the lot has implications with regard to unit manufacturing cost, since each batch production results in a fixed set-up cost. This implies that it is necessary to determine the optimal lot size by a proper trade-off between this cost and the benefits derived through better outgoing quality.

In the case of weeding, if a non-conformance is not detected at the earliest possible instant, the effort involved until it is detected is either wasted (if the item has to be scrapped) or the amount of rework required to make it conforming increases (if the non-conforming item can be fixed). Both result in extra cost. On the other hand, testing and inspection also cost money and it is necessary to achieve a suitable balance between these two costs. This implies that the location of inspection and testing stations in a production process needs to be optimally selected.

In the case of burn-in, it is necessary to determine τ optimally so that a sensible trade-off is achieved between the cost of testing and the benefits derived through improvements in the quality of outgoing products.

Determining the optimal level of quality control effort must take into account the trade-off between the cost of quality control and the benefits derived. In the context of reliability, the benefits of higher quality (conformance) are higher customer satisfaction and lower warranty costs. Figure 8.9 shows the trade-off between warranty costs and quality control costs and we need to build models to determine the optimal quality effort. There is a vast literature on this topic.[9]

8.8 Case Study: Cellular Phone

In this section, we discuss the effect of manufacturing on product reliability by looking at manufacturing of chips and shell (outer casing) of a cellular phone.

Chips

The chip manufacturing involves several stages. The first stage is to produce the wafers. These are then cut to produce dies and these are packaged to produce chips. The chip reliability is strongly influenced by the production process. Kuo and Kim (1999) classify failures of chips into the following three categories:

[9] – Optimal batch size: Djamaludin et al. (1994, 1995, 1997); Chen et al. (1998); Yeh and Lo (1998); Yeh et al. (2000)
– Acceptance sampling: Schneider (1989); Kwon (1996); Hisada and Arizino (2002); Huei (1999)
– Burn-in: Murthy et al. (1993); Blischke and Murthy (1994); Murthy (1996); Mi (1997); Kar and Nachlas (1997)

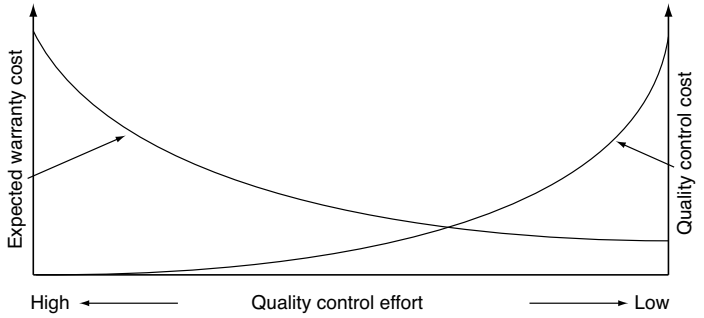


Figure 8.9. Investment in quality control effort

1. Electrical stress failures (e.g., electrical overstress, electrostatic discharge) are due to design and/or misuse of the chips.
2. Inherent failures (e.g., crystal defects, dislocations and processing defects, gate oxide breakdown, ionic contamination, surface charge spreading) tend to be the result of wafer production.
3. Extrinsic failures (e.g., die attachment failure, particulate contamination) tend to be the result of operations during device packaging.

Early failures are a result of poor design, improper manufacturing and/or incorrect use. This manifests as a failure rate which starts with a high value and decreases over the initial period (referred to as the infant mortality period). This period is typically around one year, after which the failure rate is constant over a fairly long period, which is around 40–50 years.

Burn-in

During manufacturing, infant mortality failures are removed through an accelerated burn-in process. Since the majority of chip failures is due to temperature, the burn-in involves applying high temperature and voltage to weed out items prone to early failure. The burn-in (BI) can be done at wafer lever (WLBI), die level (DLBI) or package level (PLBI).

Several questions need to be addressed in deciding on the burn-in strategy to weed out infant mortality failures. These include:

1. What should be the duration of burn-in?
2. What levels of temperature and voltage should be used in burn-in?
3. Should burn-in be done at wafer, die and/or package levels?

A proper cost-benefit analysis is needed to find answers to these questions through use of models. Kuo and Kim (1999) discuss burn-in conditions and types for semiconductor products and compare three types of burn-in.

Yield

Yield is defined as the ratio of the usable (or conforming) items to the number of items produced. Kuo and Kim (1999) characterize the semiconductor process in terms of the following sub-processes:

- Crystal growth process
- Front-end fabrication process
- Wafer probe
- Assembly and packaging
- Final testing

The wafer process yield is the yield of the first two sub-processes, wafer probe yield, assembly yield and final test yield refers to the yields of the remaining three sub-processes. The overall yield is the product of these four yields.[10] Typical average figures for the various yields (as reported in Kuo and Kim (1999)) are as given below:

Wafer process yield	94%
Wafer probe yield	50%
Assembly yield	96%
Final test yield	90%

Yield and reliability are closely linked as total wafer yield is a measure of good chips per wafer normalized by the number of chip sites per wafer. The yield is assessed through quality control scheme (involving 100% inspection and testing) and the yield-reliability model helps in assessing the reliability of the end product.[11]

[10] Cunningham et al. (1995) suggest line yield, die yield and final test yield to obtain the overall yield.

[11] See Ferris-Prabhu (1992); Hnatek (1995) for more details of yield and modelling yield in semiconductor devices.