Predicting the reliability of components produced in an improving production process.

Donal Moore, S.S. Mathematics, Trinity College Dublin

29/03/2006

Abstract

This project proposes a statistical method to model and predict the number of failures per month of electronic components. In particular, it provides a method to model the failures subject to the components being produced where there is an improving production process present. The inference and prediction are created by working in the Bayesian framework and implemented using Monte Carlo methods. The project begins by explaining how improving production processes can arise. The next section shows that the failures per month are approximately Normal and gives expressions for the mean and variance, with examples given. Next, the motivation for using Bayesian inference and Monte Carlo methods rather than classical methods, is presented. Then, the following sections outline the method for creating the model in the Bayesian framework such as the choice of prior distributions. Importance sampling (a Monte Carlo technique) is used to estimate the posterior distribution. The model is then tested by predicting the failures in month 7 and comparing it to the actual value. Cross validation is also used to test the model to an idea of the replicative power of the model.

Acronym and abbreviations

N	Number of components produced each month
K_{ij}	Number of components produced in month i that fail in month j
K_j	Number of components that fail in month j, irrespective of the month of production
$\underline{\mathbf{K}}$	vector of number of fails observed in month $1, 2, \ldots, J$
n_{ij}	Number of components produced in month i that are working in month j
pf_{ij}	Probability that a component produced in month i fails in month j
$\Lambda(i)$	Failure rate of a component produced at month i
$p.d.f.$	Probability Density Function
$p.m.f.$	Probability Mass Function
$c.d.f.$	Cumulative Density function

1 Introduction

The aim of the project is to predict the reliability of components by analysing the field return data for the first J month, i.e. how many components fail

Figure 1: Plot of observations available

in the first J months of production. The only information available is the number of components that fail in these months. It is not known when the components were produced. It is known that once failures are observed, the company strive to make improvements to the production process, hoping to reduce the number of failures to an acceptable level. The process of observing and improving occurs every month and, the company hopes, the improvement result components being produced that have a longer expected lifetime. The term "Improving Production Process" (I.P.P.) will be used to describe this scenario. This situation is encountered in many manufacturing processes but is of particular relevance to the telecoms industry where many companies conform to a quality standard TL9000. One aspect of this standard is monthly reporting of number of components returned because of failure. I.P.P. could be present if, at the beginning of the production process a flaw in the design is overlooked and is only identified after the components are returned. The company will try to fix this problem and produce components without this flaw, which should result in components being produced with a longer expected lifetime. The expected lifetime of a components reflects how reliable the component is. Therefore, the reliability of a component depends on when it was produced. Roughly speak, the earlier a component is produced, the less reliable it will be. However, for the purposes of this project, it is now know when the failed components were produced. Figure (**??**) shows the data available.

2 Modelling the reliability.

A common method of modelling the reliability of a component is to use the idea of "failure rate." The failure rate can be interpreted as the rate at which working components fail. To define the failure rate of a component, let T_p denote the time until failure of a component produced at time p. T_p is a random variable. Let T_p have c.d.f.

$$F(t) = \mathbb{P}(T_p \leq t)$$

for $t > 0$ and have p.d.f.

$$f(t) = \frac{d}{dt}F(t).$$

for $t > 0$.

Define the reliability function $R(t) = 1 - F(t) = \mathbb{P}(T_p > t)$. The failure rate is defined as

$$\Lambda(t) = \frac{f(t)}{R(t)} = -\frac{\frac{d}{dt}R(t)}{R(t)} = -log(R(t)). \quad (1)$$

For the purposes of this project, it is assumed that the failure rate of a component is constant. This means that the failure rate of a component just after it is produced is the same as at any time in the future, given that it has survived until then. This is not unreasonable for components or systems that are not subject to "ageing" (rust/wearing etc) such as light bulbs.

2.1 Effect of an improving production process on reliability

The time when a component is produced affects its reliability. Thus, the value of a component's (constant) failure rate depends on when it was produced. Let $\Lambda(p_x)$ and $\Lambda(p_y)$ denote the failure rates of components x and y produced at times p_x and p_y. Assume that $p_y \geq p_x$. Note that both $\Lambda(p_x)$ and $\Lambda(p_y)$ are both constant and that they are determined by the time of production, p_x and p_y.

Another assumption that is made is that, eventually, as many improvements as possible will be made to the process. This means that eventually components will be produced with negligible difference in failure rates. Let the initial failure rate be $\lambda_1 + \lambda_2$ and assume that it decreases exponentially to λ_2 at a rate α.

The relative decrease in failure rate between components produced at time p can be summarised as follows:

$$\Lambda(p) = \lambda_2 + \lambda_1 e^{(-\alpha p)} \tag{2}$$

Figure (**??**) is an example of a plot of the failure rate for $\lambda_1 = .3$, $\lambda_2 = .1$ and $\alpha = .5$.

Figure 2: Example of failure rate

It should be noted that this curve does not describe the failure rate of an individual component. The failure rate of each individual component is constant in time. The figure describes how the failure rate of components produced at different times is different. It describes how a component produced early in the production process will have a higher failure rate than one produced at a later stage. The failure rate of each individual component is constant and is given by the value of the curve in figure (**??**) at the time of production.

This can be interpreted as the failure rate of components produced at the beginning of the production process (p=0) having failure rate of $\lambda_1 + \lambda_2$. Then, as the production process improves, components are being produced that have a lower failure rate. Eventually, the failure rates of the components being produced converge to λ_2.

The aim of this project is to estimate λ_1 λ_2 and α based on observations of the number of failures. Failures are observed at the end of each month. At this point, the following assumptions will be made concerning when components were produced and when the observed failures occurred:

1. Components are produced at the beginning of each month and only at the beginning of each month. This means every component produced in that month has the same failure rate.
2. The number of components produced each month is the same for every month and this is known.
3. Components are working when they are produced.

Denote the month of production by $i = 0, 1, 2, \ldots$ and the month of failure by $j = 1, 2, \ldots$. The failure rate of component produced in month i is given by $\Lambda(i)$. We will now examine properties of the number of failures per month to try and create a model to describe how many fail per month.

3 Distribution of the number of failures in a month

Let K_{ij} be the number of components produced in month i that fail in month j. Let K_j be the number of components that fail in month j irrespective of the month of production. Let N be the number of components that are produced each month. Let n_{ij} be the number of components produced in month i that are working (at the start of) month j. Let pf_{ij} be the probability that a component produced in month i fails in month j.

From the above definitions it is clear that $n_{ij} = n_{i,j-1} - K_{i,j-1}$, i.e. the number of components produced in month i that are working in month j is the number working the month $(j-1)$- number that failed in month $(j-1)$.

The number of components produced each month is $N = n_{ii}$ for $i = 0, 1, \ldots$ The number of components that fail in month j, K_{ij} depend on the number of components working in the previous month, month $(j-1)$, n_{ij}. Since each of the K_{ij} components were produced in the same month, they have the same failure rate. As we shall show, this means that the probability of any one of them failing in month j is the same. Since K_{ij} is a random variable, it can be assigned a probability distribution. The Binomial distribution can be used to describe K_{ij}, thus we can write

$$K_{ij} \sim Binomial(n_{i,j-1}, pf_{ij}).$$

To calculate pf_{ij}, consider equation (**??**) for the failure rate of a component produced in month i.

Using equations **??** and **??**

$$\Lambda(i) = -\frac{d}{dt} log(R(t))$$

Integrating and using the definition of R(t) gives

$$\Lambda(i)s = -logR(s) + logR(0) = log(1 - F(t)))$$

which gives

$$F(s) = 1 - e^{-\Lambda(i)s} \tag{3}$$

This is the c.d.f. for an exponential distribution with rate $\Lambda(i)$. Thus, the time until failure,T_i, of a component produced at month i is exponentially distributed.

The probability that a component produced in month i survives longer than month j, given that it survived until month $(j-1)$ is

$$\begin{aligned}\mathbb{P}(T_i > j | T_i > (j-1)) &= \frac{\mathbb{P}(T_i > j)}{\mathbb{P}(T_i > j-1)} \\ &= \frac{e^{-\Lambda(i)j}}{e^{-\Lambda(i)(j-1)}} \\ &= e^{-\Lambda(i)} \\ &= \mathbb{P}(T_i > 1) \end{aligned} \tag{4}$$

Thus,

$$\begin{aligned} pf_{ij} &= \mathbb{P}(T_i \leq j | T_i > (j-1)) = 1 - \mathbb{P}(T_i > j | T_i > (j-1)) \\ &= 1 - \mathbb{P}(T_i > 1) = 1 - e^{-\Lambda(i)} \end{aligned} \tag{5}$$

Equation (**??**) shows that the exponential distribution is memory-less.This means that the probability of a component which has survived until month j, surviving at least another month is the same as a component surviving longer than 1 month just after it has being produced. In general, the probability that a component will survive another K months, given that it has survived J months is the same as it surviving longer than K months just after being produced. This means that the probability of failing, pf_{ij} does not depend on the month of failure j. Therefore, the probability of failing in month j will be written as pf_i.

Thus,

$$K_{ij} \sim Binomial(n_{i,j-1}, (1 - e^{\Lambda(i)}))$$

Let $f_{ij}(k_{ij})$ be the p.m.f. of K_{ij}

The p.m.f. of $K_j = \sum_{i=0}^{j-1} K_{ij}$. is given by the convolution $f_{0j} * f_{1j} * \ldots * f_{j-1,j}$. The convolution of two random variables if defined as follows. If: X_1 and X_2 are random variables with (discrete) p.m.f. $f_1(x_1)$ and $f_2(x_2)$ and $S = X_1 + X_2$, then the p.m.f of S is given by the *convolution* of $f_1(x_1)$ and $f_2(x_2)$.

$$f_s(s) = f_1(x_1) * f_2(x_2) = \sum_{\forall s} f_1(x_1) f_2(s - x_1).$$

For example, for $j = 1$,

$$K_1 = \sum_{i=0}^{0} K_{i1} = K_{01}$$

is distributed $Binomial(N, pf_0)$. The p.m.f. of K_1 is therefore:

$$f_1(k_1) = \binom{N}{k_1} {pf_1}^{k_1} (1 - pf_1)^{N-k_1}$$

For j=2,

$$K_2 = \sum_{i=0}^{1} K_{i2} = K_{02} + K_{12}.$$

The p.m.f. of K_2 is therefore:

$$f_2(k_2) = f_{02}(k_{02}) * f_{01}(k_{01}),$$

where $f_{ij}(k_{ij})$ is the p.m.f. of the number of components produced in month i that fail in month j.

$$f_{ij}(k_{ij}) = \binom{n_{ij}}{k_{ij}} p{f_i}^{k_{ij}} (1 - pf_i)^{n_{ij} - k_{ij}}.$$

Thus,

$$\begin{aligned} f_2(k_2) &= \sum_{x=0}^{N} f_{02}(x) f_{12}(k_2 - x) \\ &= \sum_{x=0}^{N} \binom{n_{02}}{x} p{f_0}^{x} (1 - pf_0)^{n_{02} - x} \binom{n_{12}}{k_2 - x} p{f_1}^{k_2 - x} (1 - pf_1)^{n_{12} - (k_2 - x)} \end{aligned}$$

This is an expression for the simple case when $j = 2$. It does not simplify easily and becomes extremely difficult to work with as j increases.

Fortunately, as we show next, K_j is approximately Normal. This approximation makes statistical analysis and prediction much easier.

4 Normal approximation of K_j

It should be noted here that the normal distribution was first developed in 1733 by Abraham de Moivre as a way to approximate a binomial distribution. ***YOU NEED A CITATION*** The number of components produced in month i_1 that fail in month j are independent of the number of components that are produced in month $i_2 \neq i_1$ that fail in the same month, i.e. $K_{i_1 j}$ and $K_{i_2 j}$ are independent. As we showed in the last section, these are distributed $Binomial(n_{i_1 j}, pf_{i_1})$ and $Binomial(n_{i_2 j}, pf_{i_2})$, respectively. Therefore, the number of components that fail in month j, $K_j = \sum_{i=0}^{j-1} K_{ij}$ is the sum of independent binomials.

Since they are independent, the mean and variance of K_j can be easily calculated.(see appendix SOMETHING for details)

$$\mu_j = \mathbb{E}[K_j] = \mathbb{E}[\sum_{i=0}^{j-1} K_{ij}] = \sum_{i=0}^{j-1} \mathbb{E}[K_{ij}] = Npf_i(1 - pf_i)^{j-i-1} \tag{6}$$

$$\begin{aligned} \sigma_j^2 &= \mathbb{V}\mathbf{ar}[K_j] = \sum_{i=0}^{j-1} \mathbb{V}\mathbf{ar}[K_{ij}] \\ &= Npf_i(1 - pf_i)^j + N p{f_i}^2 (1 - pf_i)^{j-i-1} - N p{f_i}^2 (1 - pf_i)^{2(j-i-1)} \end{aligned} \tag{7}$$

Since K_j is the sum of independent random variables, we look to the Central Limit Theorem to give us an approximation for its distribution.

The Central Limit Theorem states:

> Let $X_1, X_2, X_3, \ldots,$ be a sequence of random variables which are defined on the same probability space, share the same probability distribution D and are independent. Assume that both the expected value μ and the standard deviation σ of D exist and are finite.
>
> Consider the sum $S_n = X_1 + \ldots + X_n$. Then the expected value of S_n is $n\mu$ and its standard deviation is $n^{\frac{1}{2}}\sigma$. Furthermore, the distribution of S_n approaches the Normal distribution $N(n\mu, n\sigma^2)$ as $n \to \infty$.

In this context, $X_i = K_{ij}$ and $S_n = K_j$. The random variables K_{ij} are defined on the same probability space, their expected value and standard deviation are known and finite but K_{ij} don't have the same distribution.This means the Central Limit Theorem cannot be applied.

However, a version of the central limit theorem, known as Lyapunov's Central Limit Theorem, can be applied. It states:[**?**]

> Suppose that for each n the sequence $X_{n1}, X_{n2}, \ldots X_{nr_n}$ is independent. If:
>
> 1. $\mathbb{E}[X_{nk}] = 0$
> $\sigma_{nk}^2 = \mathbb{E}[X_{nk}^2]$
> $s_n^2 = \sum_{k=1}^{r_n} \sigma_{nk}^2$
>   where σ_{nk} is the variance and is finite,
> 2. for some $\delta > 0$, $\mathbb{E}[|X_{nk}^{2+\delta}|]$ exists and is finite for all n and
> 3. $\lim_{n\to\infty} \sum_{k=1}^{r_n} \frac{1}{s_n^{2+\delta}} \mathbb{E}|X_{nk}|^{2+\delta} = 0,$
>
> then, $\frac{S_n}{s_n}$ converges to a standard normal distribution.

The requirement for $\mathbb{E}[X_{nk}] = 0$ entails no loss of generality. The theorem can be applied the number of failures per month by:

1. Let $X_{nk} = K_{ij} - \mathbb{E}[K_{ij}]$
2. Let $s_n^2 = \sum_{i=0}^{j-1} \mathbb{V}\mathbf{ar}[K_{ij}]$
3. Let $r_j^{2+\delta} = \sum_{i=0}^{j-1} \mathbb{E}[|K_{ij} - \mathbb{E}[K_{ij}]|^{2+\delta}]$
4. Let $n = j$

It is easy to show that for $\delta = 1$, $\mathbb{E}[|K_{ij} - \mathbb{E}[K_{ij}]|^3]$ exists for all j. This is simply because the K_{ij} and therefore $\mathbb{E}[K_{ij}]$ is bounded above by N, the number of components produced each month. A similar argument can be used to show that Lyapunov's condition holds. (See appendix for details.)

It can be shown that (for $\delta = 1$):

$$\frac{r_j}{s_j} \leq \frac{(j-1)^{-\frac{1}{6}} N^{-\frac{1}{6}} (8N^2 + 6N + 3 - 3p_l + 3p_l p_u - 3p_l(1-p_u)^{j-1})^{\frac{1}{3}}}{2p_l^2(e^{-j(\lambda_1+\lambda_2)}) - p_u^2} = C(j) \quad (8)$$

$$\lim_{j \to \infty} C(j) = 0 \Rightarrow \lim_{j\to\infty} \frac{r_j}{s_j} = 0 \text{ since } \frac{r_n}{s_n} \geq 0 \text{ and } \lim_{j\to\infty} \frac{r_n}{s_n} \leq C(j)$$

Therefore, Lyapunov's condition holds and so $K_j \sim Normal(\mu_j, \sigma_j^2)$, where μ_j and σ_j^2 are defined by equations (**??**) and (**??**). Appendix (**??**) shows simulations of K_j compared to their Normal approximation.

This Normal approximation can be improved slightly. The Normal distribution is defined for all $x \in (-\infty, \infty)$. However, K_j is bounded. It cannot be negative and it has an upper bound, since the number of components working in the month before, $n_{i,j-1}$ is less than or equal to jN. For example, in the 3^{rd} month, no more than $3N$ components can fail since only $3N$ components have been produced. See Appendix (**??**) for more details. Let u be the upper bound for K_j. We can now truncate the Normal distribution so that the the probability of observing $K_j = K$ is zero outside the interval $[0, u]$. *put into*

appendix or give reference to appendix?

This is because the number of failures cannot exceed the number that are working the month before, $n_{i,j-1}$ and K_j clearly must be greater or equal to 0. Therefore, we can take a Normal distribution, truncated at 0 and jN as the approximation of the distribution of K_j. The p.d.f. of this truncated Normal distribution is

$$\mathbb{P}(K_j = k) = \frac{\exp\left[\frac{-(k-\mu_j)^2}{2\sigma^2}\right]}{\int_0^{jN} \exp\left[\frac{-(k-\mu_j)^2}{2\sigma^2}\right] dk} \quad (9)$$

We have now shown that the number of fails per month are approximately Normal and further more, we have given expression for the mean and variance. This is the model for the data. We can now begin the statistical inference.

5 Statistical Inference for Field Return Data

Statistical inference is the process of acquiring knowledge or learning about quantities that interest us from data that we have observed. Here, we try to learn about the failure rate of components by observing the number that fail each month. This section will describe the statistical inference for the model outlined in the last section for the field return data. The Bayesian approach to statistical inference will be described first.

5.1 Summary of the Bayesian approach to statistical inference

The number of fails K_j in each month $j = 1, 2, \ldots, J$, are unknown quantities. Let $\underline{\mathbf{K}} = K_1, K_2, \ldots, K_J$ denote the vector of the number of fails in each month. The field return data is $\underline{\mathbf{K}}$. The uncertainty about $\underline{\mathbf{K}}$ can be quantified using probability. There are two dominant interpretations for the meaning of probability; Physical probability and Psychological probability.

The physical interpretation probability is that it is a property of the physical world, like mass or temperature, irrespective of people and logic. There two ways to define physical probability.

1. By the classical definition of physical probability, if X is an event which occurs in k specific outcomes from a possible N outcomes occur, then $\mathbb{P}(X) \equiv \frac{k}{N}$.
2. By the relative frequency definition of physical probability, $P(X)$ is defined by the proportion of times that$X$ occurs in a long sequence of essentially identical experiments.

In both cases, probability is defined irrespective of the individual who is interested in event X. It is purely a property of the physical world.

The psychological interpretation of probability is that, if X is an event, then $\mathbb{P}(X)$ is a degree of belief, or intensity of conviction about X that an individual holds. Under the psychological definition of probability, different people may have different $\mathbb{P}(X)$, which is simply saying that different people may hold different beliefs. Psychological probability does not have to follow the laws of probability and therefore does not have to be consistent. When it strictly obeys the laws of probability and is consistent, it is called Subjective Probability. Bayesian statistical inference is based on subjective probability. This is one of the key differences between Classical statistical inference (which includes ideas such as t-tests and confidence intervals) and Bayesian statistical inference.

To learn about $\underline{\mathbf{K}}$ using the Bayesian approach, we first adopt subjective probability to describe the uncertainty about $\underline{\mathbf{K}}$. This uncertainty is quantified by $\mathbb{P}(\ \underline{\mathbf{K}} \mid \mathbb{H}\ )$, where $\mathbb{H}$ contains all the information available to us before the field return data is observed. The uncertainty about $\underline{\mathbf{K}}$, that is $\mathbb{P}(\ \underline{\mathbf{K}} \quad \mathbb{H}\ )$, is a function of our beliefs. The set $\mathbb{H}$ contains all of our knowledge, experience and background information. It is usually very large, of high dimension and a lot of the information is not relevant to $\underline{\mathbf{K}}$. Thus, it is not very efficient to work with $P(\ \underline{\mathbf{K}} \mid \mathbb{H}\ )$ directly. We wish to summarise $\mathbb{H}$ in such a way so that $\mathbb{P}(\ \underline{\mathbf{K}} \mid \mathbb{H}\ )$ is more manageable. This introduces the idea of a parameter and a parametric model. Assume that there exists a random quantity θ which summarises $\mathbb{H}$ such that, given θ, $\underline{\mathbf{K}}$ is independent of $\mathbb{H}$. This means that if we know θ, we will not get any extra information about $\underline{\mathbf{K}}$ from $\mathbb{H}$ that we have not already gotten from θ. To try and find a suitable θ to help us learn about $\underline{\mathbf{K}}$, note that, it is assumed that the number of fails in month j, K_j, depends only on the failure rate $\Lambda(i)$ for $i = 0, 1, \ldots, j-1$. The failure rate for each month i, $\Lambda(i)$, is determined by the parameters $\lambda_1, \lambda_2, \alpha$. Therefore, if we know the values of $\lambda_1, \lambda_2, \alpha$, we will not get extra information about $\underline{\mathbf{K}}$ from $\mathbb{H}$. Thus, given $\lambda_1, \lambda_2, \alpha$, the field return data $\underline{\mathbf{K}}$ is independent of $\mathbb{H}$. We have now found a suitable expression for θ. If $\theta = (\lambda_1, \lambda_2, \alpha)$, then given θ, $\underline{\mathbf{K}}$ is independent of $\mathbb{H}$.

By the partition law, we can write:

$$\mathbb{P}(\underline{\mathbf{K}} \mid \mathbb{H}) = \int_{\forall\theta} \mathbb{P}(\underline{\mathbf{K}} \mid \theta, \mathbb{H})\mathbb{P}(\theta \mid \mathbb{H})d\theta = \int_{\forall\theta} \mathbb{P}(\underline{\mathbf{K}} \mid \theta)\mathbb{P}(\theta \mid \mathbb{H})d\theta \tag{10}$$

The last equality follows since $\underline{\mathbf{K}}$ is independent of $\mathbb{H}$ given θ. In equation (**??**), θ is a parameter and $\mathbb{P}(\underline{\mathbf{K}} \mid \theta)$ is the parametric model for the data. In the last section, we showed how K_j is Normal with mean and variance depending on $\theta = (\lambda_1, \lambda_2, \alpha)$. The number of components that fail in month j_1 and month j_2 are assumed to be independent for $j_1 \neq j_2$. This assumption is not too unreasonable since the numbers of components failing each month is expected to be small.(See appendix **??** for details.) Therefore, we can write

$$\mathbb{P}(\underline{\mathbf{K}} \mid \theta) = \mathbb{P}(K_1, K_2, \ldots, K_J \mid \theta) = \prod_{j=1}^{J} \mathbb{P}(K_j \mid \theta) \tag{11}$$

Since $\underline{\mathbf{K}}$ is observed, $\mathbb{P}(\underline{\mathbf{K}} \mid \theta)$ is called the likelihood of θ (given the data). Rewriting equation (**??**) gives

$$\mathbb{P}(\underline{\mathbf{K}} \mid \mathbb{H}) = \left(\prod_{j=1}^{J} \mathbb{P}(K_j \mid \theta) \right) \mathbb{P}(\theta \mid \mathbb{H})$$

The random quantity θ determines the failure rate and the failure rate allows us to model the failures. Then we will be able to predict the number of components that will fail in future months. Therefore, we would like to be able to describe θ by making use of the observations (and the model for the observations) and prior information that is available to us. The expression which describes the uncertainty of θ in terms of prior knowledge and observed data is known as the Posterior Distribution for θ given the data. Using Bayes Law and equations (**??**) and (**??**), we can write

$$\mathbb{P}(\theta \mid \underline{\mathbf{K}} , \mathbb{H}) = \frac{\mathbb{P}(\underline{\mathbf{K}} \mid \theta)\mathbb{P}(\theta \mid \mathbb{H})}{\int_{\forall\theta} \left(\mathbb{P}(\underline{\mathbf{K}} \mid \theta)\right) \mathbb{P}(\theta \mid \mathbb{H})d\theta} \tag{12}$$

$$= \frac{\prod_{j=1}^{J} \mathbb{P}(K_j \mid \theta)\mathbb{P}(\theta \mid \mathbb{H})}{\int_{\forall\theta} \left(\prod_{j=1}^{J} \mathbb{P}(K_j \mid \theta) \right) \mathbb{P}(\theta \mid \mathbb{H})} \tag{13}$$

Since $\int_{\forall\theta} \left(\prod_{j=1}^{J} \mathbb{P}(K_j \mid \theta) \right) \mathbb{P}(\theta \mid \mathbb{H})d\theta$ is constant in θ, we can write

$$\mathbb{P}(\theta \mid \underline{\mathbf{K}} , \mathbb{H}) \propto \prod_{j=1}^{J} \mathbb{P}(K_j \mid \theta)\mathbb{P}(\theta \mid \mathbb{H})$$

or

$$posterior \propto likelihood \times prior$$

The posterior distribution is the distribution of the parameters obtained by updating the prior distribution by using observed data. In this case, $\theta = (\lambda_1, \lambda_2, \alpha)$. To make use of the posterior distribution, an expression for the likelihood and for the prior distribution must be found.

Another distribution we may be interested in is the distribution for the number of components that will fail in the next month (in month 7) given the observations $\underline{\mathbf{K}}$. This is known as the posterior predictive distribution.

$$\mathbb{P}(K_7| \underline{\mathbf{K}}) = \int_{\forall\theta} \mathbb{P}(K_7|\theta)\mathbb{P}(\theta| \underline{\mathbf{K}}, \mathbb{H})d\theta \tag{14}$$

This is similar to the posterior distribution for θ but the prior distribution $\mathbb{P}(\theta| \mathbb{H})$ for θ that appeared in the expression for the posterior distribution is replaced by the posterior distribution $\mathbb{P}(\theta| \underline{\mathbf{K}}, \mathbb{H})$.

5.2 The Likelihood

The likelihood, $\mathbb{P}(K_1, K_2, \ldots, K_J|\lambda_1, \lambda_2, \alpha)$. It is the joint probability of observing the data given the parameters. Section (??) showed how the data could be modelled with a normal distribution and it is assumed that the observations are independent. [1] Therefore,

$$\mathbb{P}(K_1, K_2, \ldots, K_J|\lambda_1, \lambda_2, \alpha) = \prod_{j=1}^{J} \mathbb{P}(K_j|\lambda_1, \lambda_2, \alpha) = \prod_{j=1}^{J} \frac{1}{\sqrt{2\pi\mathbb{V}\mathbf{ar}[K_j]}} \exp\left(\frac{K_j - \mathbb{E}[K_j]}{2\mathbb{V}\mathbf{ar}[K_j]}\right) \tag{15}$$

where $\mathbb{P}(K_j|\lambda_1, \lambda_2, \alpha)$ is the p.d.f. for the Normal distribution with mean $\mathbb{E}[K_j]$, from equation (??) and variance $\mathbb{V}\mathbf{ar}[K_j]$ from equation (??)

5.3 Prior distributions

The prior distribution $\mathbb{P}(\theta| \mathbb{H}) = \mathbb{P}(\lambda_1, \lambda_2, \alpha| \mathbb{H})$ describes the uncertainty about $(\lambda_1, \lambda_2, \alpha)$ before the data is observed. It is assumed that α is independent of λ_1 and λ_2. This makes it easier to find an expression for the prior distribution.

5.3.1 Prior distribution for λ_1 and λ_2

First, λ_1 and λ_2 will be examined. These values determine the initial failure rate $(\lambda_1 + \lambda_2)$ and the failure rate to which the production process goes to (λ_2). It is possible that λ_1 and λ_2 could take on any positive real number. However, the components in question have already been sold and past experience tells us that it is very unlikely that almost all of the components will fail before they are 1 month old. The probability that a component produced in month i fails before it is 1 month old, pf_i is highest at the beginning of the production

[1] see appendix or include more here on independence

process. This corresponds to when $i = 0$. At $i = 0$, $pf_0 = 1 - e^{\lambda_1+\lambda_2}$ This how our prior knowledge allows us to put an upper bound on $\lambda_1 + \lambda_2$

Figure (**??**) shows the probability of a component failing before the end of the first month for different values of $\lambda_1 + \lambda_2$. For example, if $\lambda_1 + \lambda_2 = 1$, pf_0 is approximately .63. .

Figure 3: Plot of pf_0 against $(\lambda_1 + \lambda_2)$

From this plot, it is clear that for $\lambda_1 + \lambda_2 \geq 5$, pf_0 is close to 1, or equivalently, if $\lambda_1 + \lambda_2 \geq 5$, it is very likely that every component produced in the first month will fail before it is 1 month old. For $\lambda_1 + \lambda_2 = 5$, $pf_0 > .99$. Even though we strongly believe that the probability that every component produced in the first month will fail before it is one month old, we should still admit this possibility. It would be far worse to dismiss it. Appendix (**??**) shows examples of what will happen if we chose a different value for the uper bound of $\lambda_1 + \lambda_2$. Apart from the restriction on $\lambda_1 + \lambda_2$, there is no other restriction that will be put on λ_1 or λ_2. Therefore, a uniform distribution will be used for $\mathbb{P}(\lambda_1, \lambda_2)$

The joint distribution for λ_1, λ_2 that suitably reflects the prior information can now be constructed. The pdf for the distribution of λ_1, λ_2 is given by

$$\mathbb{P}(\lambda_1, \lambda_2) = \mathbb{P}(\lambda_1|\lambda_2)\mathbb{P}(\lambda_2)I_{\{\lambda_1+\lambda_2\leq 5\}} \tag{16}$$

where $\mathbb{P}(\lambda_1|\lambda_2)$ is a uniform distribution on $[0, 5 - \lambda_2]$ and $\mathbb{P}(\lambda_2)$ is a uniform distribution on $[0, 5]$ and

$$\begin{aligned} I_{\{\lambda_1+\lambda_2\leq 5\}} &= 1 \quad \ldots \quad \{\lambda_1 + \lambda_2 \leq 5\} \\ &= 0 \quad \ldots \quad \{\lambda_1 + \lambda_2 > 5\} \end{aligned}$$

This leads to the following expressions for $\mathbb{P}(\lambda_1, \lambda_2)$

$$\mathbb{P}(\lambda_1, \lambda_2) = \frac{1}{12.5}$$

$$\mathbb{P}(\lambda_2) = \frac{5 - \lambda_2}{12.5}$$

$$\mathbb{P}(\lambda_1|\lambda_2) = \frac{1}{5 - \lambda_2} \tag{17}$$

5.3.2 Prior distribution for α

The failure rate of a component produced in month i is given by $\Lambda(i) = \lambda_2 + \lambda_1 e^{-\alpha i}$ The parameter α controls how quickly the failure rate decreases to λ_2. This is how quickly the production process achieves the best way of producing components. This occurs when $e^{-\alpha i}$ is close to 0. Figure (**??**) shows 3 plots of $e^{-\alpha t}$ for $t \geq 0$, with $\alpha = 1, 5, 10$. When $\alpha = 10$ (black line), $e^{-\alpha t}$ is close to 0 before $t = 1$. This means that the production process will achieve the best way

Figure 4: Plot of $e^{-\alpha t}$ against $t > 0$

of producing components well before the end of the first month. When $\alpha = 5$, (red line), $e^{-\alpha t}$ is close to 0 at $t = 1$. The blue line corresponds to $\alpha = 1$. Previous experience suggests that it is likely to take longer than one month to achieve the best way of producing components. This means that α is unlikely to be 10. Prior knowledge also suggests that it is more likely to take longer than 4 months to achieve the the best production process. This corresponds to $\alpha \leq 1$. However, we are not willing to forget about the possibility that the production process could achieves the best method of producing components very quickly. A gamma distribution with shape parameter a and rate parameter b is used to describe this. The p.d.f. of the gamma distribution is:

$$f_\alpha(x) = \frac{b^a}{\Gamma(a)} x^{a-1} \exp -xb$$

with expected value

$$\mathbb{E}[x] = \frac{a}{b}$$

and variance

$$\mathbb{V}\mathbf{ar}[x] = \frac{a}{b^2}$$

Since we believe that it is more likely for $\alpha \leq 1$ than $\alpha \geq 1$, a and b are chosen so that the mode of the distribution is less than 1. A gamma distribution with shape $a = .5$ and rate $b = .5$ was chosen as a prior for α

Therefore the prior distribution for α is

$$\mathbb{P}(\alpha| \mathbb{H}) == \frac{.5^{.5}}{\Gamma .5} \alpha^{-.5} \exp(-2\alpha) \tag{18}$$

Here is plot of the pdf for $\mathbb{P}(\alpha| \mathbb{H})$

Figure (**??**) is a histogram of the $\mathbb{P}(\alpha)$

Figure 5: histogram of 1e6 deviates from Gamma(.6,.4)

The prior distribution for λ_1, λ_2 and α can now be fully written as

$$\mathbb{P}(\lambda_1, \lambda_2, \alpha) = \mathbb{P}(\lambda_1, \lambda_2| \mathbb{H})\mathbb{P}(\alpha| \mathbb{H}) = \frac{1}{50} \frac{.5^{.5}}{\Gamma(.5)} \alpha^{-.5} \exp(-2\alpha) = \frac{\exp(-2\alpha)}{50\sqrt{2\pi\alpha}} \tag{19}$$

since $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ With expressions for both the likelihood and the prior distributions, the posterior distribution can be expressed by using equations (**??**), (**??**) and (**??**).

$$\mathbb{P}(\lambda_1, \lambda_2, \alpha | K_1, K_2, \ldots, K_J) \propto \left(\prod_{j=1}^{j=J} \frac{1}{\sqrt{\mathbb{V}\mathbf{ar}[K_j]}} \exp\left[-\frac{(K_j - \mathbb{E}[K_j])^2}{2\mathbb{V}\mathbf{ar}[K_j]} \right] \right) \frac{\exp(-2\alpha)}{\sqrt{\alpha}}$$

$$= \frac{\exp\left[-\left(\sum_{j=1}^{J} \left(\frac{(K_j - \mathbb{E}[K_j])^2}{2\mathbb{V}\mathbf{ar}[K_j]} \right) + 2\alpha \right) \right]}{\sqrt{\alpha} \prod_{j=1}^{J} \sqrt{\mathbb{V}\mathbf{ar}[K_j]}} \quad (20)$$

This is very difficult to work with explicitly since $\mathbb{E}[K_j]$ and $\mathbb{V}\mathbf{ar}[K_j]$ depend non trivially on λ_1, λ_2 and α, as we have seen in section (**??**). A useful closed form expression for the posterior distribution is not available so Monte Carlo methods will be used to approximate it.

6 Monte Carlo methods to estimate the posterior distribution.

The Monte Carlo method for estimating a p.d.f. is the process of simulating values from the distribution repeatedly and using the simulated values to estimate properties of the distribution, such as the mean and variance. Importance sampling is such a method. With importance sampling, a large number of samples are simulated from the prior distribution, say $K = 10^5$. Each sample of $\lambda_1^{(k)}, \lambda_2^{(k)}$ and $\alpha^{(k)}$ is given a weight, w_k. The weight is the probability of observing the data, given $\lambda_1^{(k)}, \lambda_2^{(k)}$ and $\alpha^{(k)}$ (i.e. the likelihood of $\lambda_1^{(k)}, \lambda_2^{(k)}$ and $\alpha^{(k)}$) The weight is then normalised so that $\sum_{k=1}^{K} w_k = 1$. The normalised weights w_k define a distribution on the K samples. Now, $M \ll K$, say $M = 10^4$, samples are drawn, with replacement, from the K samples according to the distribution w_k on the K samples. Sampling via importance sampling in this way is equivalent to sampling from the posterior distribution.[**?**] ***read before you submit....*** Therefore, selecting the M values in this way is equivalent to simulating M values from the posterior distribution. Using the samples from the posterior distribution, it is possible to estimate many quantities of interest that are defined by the posterior p.d.f., such as the expected value, variance and highest posterior densities (the Bayesian equivalent to confidence intervals.) These quantities are estimated by using Monte Carlo Integration.

6.1 Monte Carlo Integration

Let $\theta = (\lambda_1, \lambda_2, \alpha)$ To illustrate the method of Monte Carlo Integration, consider the posterior mean, $\mathbb{E}[\theta| \underline{\mathbf{K}}]$

$$\mathbb{E}[\theta| \underline{\mathbf{K}}] = \int_{\forall\theta} \theta \mathbb{P}(\theta| \underline{\mathbf{K}}) d\theta$$

where $\mathbb{P}(\theta|\ \underline{\mathbf{K}}\ )$ is the posterior distribution of θ. It is possible to approximate $\int_{\forall\theta} \theta\mathbb{P}(\theta|\ \underline{\mathbf{K}}\ )d\theta$ by $\frac{1}{M}\sum_{m=1} M\theta^{(m)}$, where $\theta^{(m)}$ are samples from the posterior distribution. In the previous section, we showed how to sample from the posterior distribution using importance sampling. Now, using these samples, we can approximate the mean of the posterior distribution.

$$\mathbb{E}[\theta|\ \underline{\mathbf{K}}\ ] \approx \frac{1}{M}\sum_{m=1}^{M} \theta^{(m)} \tag{21}$$

In general, it is possible to find the expected value of a function $h(\theta)$ in this way.

$$\mathbb{E}[h(\theta)|\ \underline{\mathbf{K}}\ ] \approx \frac{1}{M}\sum_{m=1}^{M} h(\theta^{(m)}) \tag{22}$$

For example, if $h(\theta|\ \mathbb{H}\ ) = (\theta - \mathbb{E}(\theta))^2$, then equation (**??**) allows us to estimate the variance. We may also be interested in the highest posterior density of θ, that is the region where 95%, say, of the probability of θ lies in the posterior distribution. The upper and lower bound, u_l, u_u, for the HPD interval is found by solving the following equations for b_{lower} and b_{upper}:

$$\int_{b_l}^{b_u} \mathbb{P}(\theta|\ \underline{\mathbf{K}}\ ) = .95$$

Or equivalently,

$$\int_{0}^{b_l} \mathbb{P}(\theta|\ \underline{\mathbf{K}}\ ) = .05 \tag{23}$$

$$\int_{b_l}^{\infty} \mathbb{P}(\theta|\ \underline{\mathbf{K}}\ ) = .05 \tag{24}$$

Equations (**??**) and (**??**) can be re written as

$$\int_{0}^{\infty} I_{(\theta\in(0,b_l))}\mathbb{P}(\theta|\ \underline{\mathbf{K}}\ ) = \mathbb{E}[I_{(\theta\in(0,b_l))}] \text{ w.r.t. } \mathbb{P}(\theta|\ \underline{\mathbf{K}}\ ) \approx \frac{1}{M}\sum_{m=1}^{M} I_{(\theta^{(m)}\in(0,b_l))} \tag{25}$$

Similarly,

$$\int_{b_l}^{\infty} \mathbb{P}(\theta|\ \underline{\mathbf{K}}\ ) \approx \sum_{m=1}^{M} I_{(\theta^{(m)}\in(b_u,\infty))}, \tag{26}$$

where in both cases, $\theta^{(m)}$ are sampled from the posterior distribution. A simple way to compute $\sum_{m=1}^{M} I_{(\theta^{(m)}\in(b_u,\infty))}$ is to set b_l and b_u as the $.025M^{th}$ and $.975M^{th}$ biggest value of θ, respectively, from the set $\{\theta^{(m)} : m = 1, 2, \ldots, M\}$.

It is also possible to compute the normalising constant for the posterior distribution (the constant that ensures the distribution integrates to 1) in this way. From equation(**??**):

$$\mathbb{P}(\theta|\ \underline{\mathbf{K}}\ ,\ \mathbb{H}\ ) \propto \prod_{j=1}^{J} \mathbb{P}(K_j|\theta)\mathbb{P}(\theta|\ \mathbb{H}\ ).$$

The normalising constant is therefore

$$\int_{\forall\theta} \prod_{j=1}^{J} \mathbb{P}(K_j|\theta)\mathbb{P}(\theta|\ \mathbb{H}\ ) \approx \frac{1}{M}\sum_{m=1}^{M}\prod_{j=1}^{J}\mathbb{P}(K_j|\theta) \text{ w.r.t. } \mathbb{P}(\theta|\ \mathbb{H}\ ) \tag{27}$$

It may also be of interest to us to compute the p.d.f. for the posterior distribution. This is done in the following way: First, note that for small $\delta\theta$,

$$\begin{aligned} \mathbb{P}(\theta = \theta^*|\ \underline{\mathbf{K}}\ )\delta\theta &\approx \mathbb{P}(\theta^* \leq \theta \leq \theta^* + \delta\theta|\ \underline{\mathbf{K}}\ ) \\ &\approx \frac{1}{M}\sum_{m=1}^{M} I_{\theta^{(m)}\in(\theta,\theta+\delta\theta)} \end{aligned} \tag{28}$$

Therefore,

$$\mathbb{P}(\theta|\ \underline{\mathbf{K}}\ ) \approx \frac{1}{\delta\theta M}\sum_{m=1}^{M} I_{\theta^{(m)}\in(\theta,\theta+\delta\theta)}$$

7 Impelemting and Testing the model

The methods presented in the previous sections is a general method which (we hope) will work with any number of observations J and will be able to capture the relative change between failure rate between components produced in different months, no matter how dramtic the change it.

The data that will be used to to test the model s generated by simulating the number of failures per month with $N = 1000$, $\lambda_1 = 0.09$, $\lambda_2 = 0.01$ and $\alpha = 0.5$. This corresponds to the failure rate beginning at 0.1 and decreasing exponentially to 0.01 at a rate 0.5, or equvilently, the probability that a component fails before the end of the first month, pf_0, begins are $1 - e^{-(.1} \approx 0.095$ and decreases to about 0.01 after aprroximately . Since the valus of these parameters are known, the accuracy of the model can be crudely estimated. Observations for the number of failures per month for the first 24 months were simulated.

First, we will consider the case when only 6 observations are available. Importance sampling was implemented with $K = 10000$ values sampled from the prior distribution and $M = 1000$ values sampled from the posterior.

Figure (**??**) are histograms of samples of λ_1, λ_2 and α from the prior distribution. Figure (**??**) show a plot of the likelihood of λ_1, λ_2 and α given the

first 6 observations $K_1, K_2, \ldots, K_6$. We can no examine a plot of the weight

given to each of the K samples from the prior distribution. Since alot of the weights are very close to 0, it is natural to look at the (non-zero) weights on the log scale. Figure (**??**) shows this. From examining Figure (**??**), we can see

the number of prior samples that are given small weight (weight close to zero, or equivaltenly, log(weight) very large and negative.

We will now implement the methods as described in the previous section to the data that is available. Figure (**??**) shows the data available to us. Using importance sampling with $K = 10^5$ and $M = 10^4$, the following was obtained.

Now that an estimate for λ_1, λ_2 and α have been obtained, it is possible to check whether they effectively model the failure rates. Let $\hat{\lambda_1}, \hat{\lambda_2}$ and $\hat{\alpha}$ denote the estimates of λ_1, λ_2 and α

First, the number of components that will fail in month 7 is predicted. this is done by finding the expected value of K_7, given $\hat{\lambda_1}, \hat{\lambda_2}$ and $\hat{\alpha}$. This is equivalent to computing $\mathbb{E}[K_7]$ using equation (**??**). The 95% confidence interval can also be constructed by estimating the variance using equation (**??**). The expected value and the 95% confidence interval can be then be compared to a simulated value for K_7. Here are the results:

True value of K_7	Expected value	95% lower bound	95% Upper bound
702	517.4653	474.1273	560.7961

Table (**??**) shows the result of using importance sampling with $K = 10^4$, $M = 10^3$ and with the observations $K_1 \ldots K_J = 0, 300, 494, 562, 700, 701$.

8 results of further simulations

the results, in the form of plots and table will be outlined here. The weakness and strengths of the model will be illustrated.

Appendix

A Lyapunov's Condition

Let $\mathbb{E}[K_{ij}] = \mu_{ij}$

$$\mathbb{E}[|K_{ij} - \mu_{ij}|^3] \leq \mathbb{E}[|K_{ij} + \mu_{ij}|^3]$$

since K_{ij} and μ_{ij} are both positive. Therefore

$$\begin{aligned}
\mathbb{E}[|K_{ij} - \mu_{ij}|^3] \leq E[|K_{ij} + \mu_{ij}|^3] &= \mathbb{E}[K_{ij}^3] + 3\mathbb{E}[K_{ij}^2]\mu_{ij} + 3\mathbb{E}[K_{ij}]\mu_{ij}^2 + \mu_j^3 \\
&= \mathbb{E}[K_{ij}^3] + 3\mathbb{E}[K_{ij}^2]\mu_{ij} + 4\mu_{ij}^3 \qquad (29)
\end{aligned}$$

The number of components produced in month i that fail in j $K_{ij} \sim Binomial(n_{i,j-1}, pf_i)$ Using the iterated laws of expectation:

$$\mathbb{E}[K_{ij}^3] = \mathbb{E}[\mathbb{E}[K_{ij}^3 | n_{i,j-1}]]$$

The moment generating function for a $Binomial(n, p)$ is:

$$\phi(t) = \mathbb{E}[e^{tx}] = (q + pe^t)^n$$

where $q = 1 - p$; Using this we can find $\mathbb{E}[K_{ij}^3 | n_{i,j-1}]$ by taking the finding the 3^{rd} derivative at 0.

$$\begin{aligned}
\frac{d\phi}{dt}|t = 1 &= \mathbb{E}[K_{ij}^3 | n_{i,j-1}] \\
&= pf_i^3 n_{i,j-1}(n_{i,j-1} - 1)(n_{i,j-1} - 2) \\
&+ 3pf_i^2 n_{i,j-1}(n_{i,j-1} - 1) + n_{i,j-1}pf_i \qquad (30)
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathbb{E}[K_{ij}^3] &= \mathbb{E}[\mathbb{E}[K_{ij}^3 | n_{i,j-1}]] \\
&= pf_i^3 \mathbb{E}[n_{i,j-1}(n_{i,j-1} - 1)(n_{i,j-1} - 2)] \\
&\quad +3pf_i^2 \mathbb{E}[n_{i,j-1}(n_{i,j-1} - 1)] + \mathbb{E}[n_{i,j-1}]pf_i \qquad (31)
\end{aligned}$$

The factorial moment generating function for a $Binomial(n, p)$ is $\mathrm{F}(t) = \mathbb{E}[t^x] = (q + tp)^n$ where $q = 1 - p$. We will be able to find $\mathbb{E}[(n_{i,j-1}(n_{i,j-1} - 1)]$ and $\mathbb{E}[(n_{i,j-1}(n_{i,j-1} - 1)(n_{i,j-1} - 2)]$ by taking the 2^{nd} and 3^{rd} derivative respectiely and setting them equal to 1.

$$\begin{aligned}
\mathbb{E}[n_{i,j-1}(n_{i,j-1} - 1)] &= N(N - 1)(1 - pf_i)^{(2-j-i)} \qquad (32) \\
\mathbb{E}[(n_{i,j-1}(n_{i,j-1} - 1)(n_{i,j-1} - 2)] &= N(N - 1)(N - 2)(1 - pf_i)^{3(j-i-1)} \qquad (33)
\end{aligned}$$

Plugging equations (??), (??), (??) and (from section (??)) equation (??) into equation (??), gives

$$\begin{aligned}
\mathbb{E}[K_{ij}^3] &= N(N - 1)(N - 2)(1 - pf_i)^{3(j-i-1)} pf_i^3 \\
&+ 3pf_i N(N - 1)(1 - pf_i)^{2(j-i-1)} + N(1 - pf_i)^{(j-i-1)} pf_i \qquad (34)
\end{aligned}$$

To find $\mathbb{E}[K_{ij}^2]$, note simply that $\mathbb{E}[K_{ij}^2] - \mathbb{E}[K_{ij}]^2 = \mathbb{V}\mathbf{ar}[K_{ij}]$. Using equations (**??**) and (**??**) we find that

$$\begin{aligned}\mathbb{E}[K_{ij}^2] &= Npf_i(1-pf_i)^{(j-i)} + Npf_i^2(1-pf_i)^{(j-i-1)} + N^2pf_i^2(1-pf_i)^{2(j-i-1)} \\ &+ Npf_i^2(1-pf_i)^{2(j-i-1)} \end{aligned} \tag{35}$$

By plugging (**??**), (**??**) and (**??**) into equation (**??**), we find that

$$\begin{aligned}\mathbb{E}[|K_{ij}-\mu_{ij}|^3] &\leq \mathbb{E}[K_{ij}^3] + 3\mathbb{E}[K_{ij}^2]\mu_{ij} + 4\mu_{ij}^3 \\ &= N(8N^3 - 3N + 2) + 3N^2(3 - pf_i(1-pf_i)^{(j-i-1)}) \\ &+ N(1 - 3pf_i(1-pf_i)) \end{aligned} \tag{36}$$

B Plots of simulated fails per month vs normal approximation

(a) Here is the description of the figure in left (b) Here is the description of the figure in right

Figure 6: Here is the caption of the entire figure.

plots like figure (**??**) will be put in here

C independence of K_{ij} and calculating the mean and variance for the normal approximation to $f_j(k_j)$

The observations K_j are not independent of each other. The number of components that fail in month j depends on the $j-1$ months before. This is because, at month j, the number of components that have been produced up to then is jN, number of components that fail (cumulatively) before month j is $\sum_{m=1}^{j-1} K_m$, which gives the number of working components in month $j-1$, as

$$(j-1)N - \sum_{m=1}^{j-1} K_m$$

, where K_m is the number of components that fail in month m.

We have already shown in section (**??**) that each month can be approximated by a truncated normal distribution. The upper bound for each month j should be

$$u_j = jN - \sum_{m=1}^{j-1} K_m$$

Clearly, u_j depends on the number of components that have already failed, that is on $K_1, K_2, \ldots, K_{j-1}$. The number of components that fail in month j cannot exceed u_j This means that K_j depends on $K_1, K_2, \ldots, K_{j-1}$. However, it is assumed that the number of components that fail each month is small compare to N. This is reasonable to assume since the manufacturing company will not sell their components if they expect a large number of them to fail. This means that the probability that K_j will be close to the upper bound is very small. Therefore, the effect of truncation is negligable.

References

[1] M. A. Tanner. *Tools for statistical inference.* Springer, New York, Tthird edition, 1997.

[2] S. M. Ross. *Simulation.* Academic Press, London, Third edition, 2002.

[3] Patrick Billingsley *Probability and Measure* Wiley, New York, Third edition, 1995

[4] Mood, Alexander McFarlane *Introduction to the theory of statistics*