ESTIMATING THE NUMBER OF PRODUCT FAILURES: A THEORETICAL APPROACH

by
AYDA AMNIATTALAB

Submitted to Faculty of Engineering and Natural Sciences
in partial fulfillment of the requirements for the degree of
Master of Science

Sabanci University
July 2018

ESTIMATING THE NUMBER OF PRODUCT FAILURES: A THEORETICAL APPROACH

APPROVED BY

Assist. Prof. Dr. Murat Kaya (Thesis Supervisor) ....

Prof. Dr. J.B.G. Frenk (Thesis Co-supervisor) ....

Assoc. Prof. Dr. Kemal Kılıç ....

Assoc. Prof. Dr. Semih Onur Sezer ....

Assoc. Prof. Dr. Esma Nur Çinicioğlu ....

DATE OF APPROVAL: 30.07.2018

ACKNOWLEDGMENTS

First of all, I would like to express my gratitude to my co-supervisor Prof. Dr. Murat Kaya for his guidance, continued support, constant presence and constructive suggestions which were determinant for the accomplishment of the work presented in the thesis.

I also would like to thank to my co-supervisor Prof. Dr. Hans Frenk for his guidance and unwavering support throughout my M.Sc. research studies. It was an honor for me to be a student of him and share his exceptional scientific knowledge. As what Karl Menninger says: “What a teacher is, is more important than what he teaches.”

I am also thankful to Professors Kemal Kılıç and Raha Akhavan for their valuable assistance and suggestions.

Finally, I offer my special regards and blessings to my family for their concern, love and unconditional support throughout my life.

ESTIMATING THE NUMBER OF PRODUCT FAILURES:
A THEORETICAL APPROACH

AYDA AMNIATTALAB

Industrial Engineering Department, Master of Science, 2018

Thesis supervisor: Assist. Prof. Dr. Murat Kaya
Thesis co-supervisor: Prof. Dr. J.B.G. Frenk

Key words: Stochastic Point Processes, Time Series, Linear Regression, Maximum Likelihood, Censored Data.

ABSTRACT

In this thesis, we propose a stochastic process describing the total number of failed items under warranty over time. This stochastic process consists of a sales process represented by a stochastic point process and a process counting the total random number of repairs applied to an arbitrary item of this product. Combining these two stochastic processes yields a representation of the counting process of the total random number of failed items returned to the manufacturer within their warranty period. To fit the proposed parametric model to a large data set we need to estimate separately the intensity measure of both the failure and sales process. To estimate the intensity measure of the cumulative sales process we use some well known parametric functions and apply linear regression techniques. Also, under the assumption that a repair does not change the age of the particular item of the product it can be shown that the counting process of failures is a non-homogenous Poisson process and so we need to estimate the cdf of the time to the first failure. Since our data set is censored we apply the Maximum Likelihood principle for censored data and use as a parametric class the class of Weibull distributions. Our approach serves as an alternative to the time series based approaches for cases where item tracking information is available.

ARIZALI ÜRÜNLERIN SAYIM TAHMİNİ: TEORİK YÖNTEM

AYDA AMNIATTALAB

Endüstri Mühendisliği, Yüksek Lisans Tezi, 2018

Tez Danışmanı: Dr. Öğr. Üyesi Murat Kaya
Yardımcı Tez Danışmanı: Prof. Dr. J. B. G. Frenk

Anahtar kelimeler: Stokastik Nokta Süreçleri, Zaman Serileri, Doğrusal Regresyon, Maksimum Olabilirlik, Sansürlü Veri.

ÖZET

Bu tezde, garanti süresi içerisinde arıza çıkaran ürünlerin sayısı üzerine bir model önerilmiştir. Bu model, satış sürecini genel nokta süreçleri (sürekli zamanlı model) ya da zaman serileri (ayrık zamanlı model) aracılığı ile göstermenin yanı sıra garanti kapsamında geçekleşen toplam onarım sayısını da açıklamaktadır. Bu iki stokastik sürecin birleşmesi, üretici tarafından tamir edilecek toplam arızalı ürün sayısını sayan bir stokastik sürecin tam bir temsilini verir. Geliştirilen kuramsal model ve onun türetilmiş özellikleri, parametrik bir model olarak büyük bir veri kümesine uygulanmıştır. İlk olarak, iki ayrı ürünün gelecekteki satışlarının bazı iyi bilinen parametrik fonksiyonlar kullanılarak doğrusal regresyon vasıtasıyla nasıl tahmin edilebileceği incelenmiştir. İkinci olarak, onarımın ürünün yaşını değiştirmediği varsayımı altında, arıza süreleri Maksimum Olabilirlik Tahmin Edicisi yönteminin sansürlü veriler üzerinde kullanımı yoluyla tahmin edilmiştir. Bu yöntemde, sınıf olarak sık kullanılan Weibull dağılım sınıfı kullanılmıştır. Önerilen yaklaşım, ürün bazında takibin mümkün olduğu durumlarda zaman serilerini kullanan yaklaşımlara bir alternatif oluşturmaktadır.

Contents

Abstract v
Özet vi

1 INTRODUCTION **10**

2 LITERATURE REVIEW **12**
2.1 Sales 12
2.2 Failures 13

3 A MATHEMATICAL MODEL FOR THE TOTAL NUMBER OF FAILURES AND THE COST PROCESS **16**
3.1 A continuous time model for the total cost of repairing failed items and the total number of failed items 16
3.2 On non-homogeneous Poisson cumulative sales processes. . . . 32

4 ESTIMATING THE INTENSITY MEASURE OF THE SALES PROCESS **54**
4.1 Sales data 54
4.2 Methodology for estimating sales 56
4.2.1 The additive model 58
4.2.2 The multiplicative model 60
4.3 Fitting the additive model to daily sales 61
4.4 Fitting the multiplicative model to daily sales 62
4.5 Fitting the multiplicative model using a smoothed version of the data 65
4.6 Model selection 67

5 ESTIMATING THE INTENSITY MEASURE OF THE FAILURE PROCESS **69**
5.1 Failure data 69
5.2 Methodology 73
5.2.1 Basic concepts 73
5.2.2 Maximum Likelihood Estimation for censored data . . . 76
5.2.3 Weibull distribution 80
5.2.4 MLE optimization for Weibull distribution 82
5.3 Estimation of Weibull parameters 85

6 CONCLUSION AND DISCUSSION **88**
Appendix A 90
Appendix B 93
BIBLIOGRAPGY 94

List of Figures

1.1 Life flow 11

2.1 Censoring within the failure times data 15

4.1 Sales data in raw form 55
4.2 Daily sales data of product A (2013-2017) 55
4.3 Daily sales data of product B (2013-2017) 56
4.4 Average daily sales data of product A 56
4.5 Average daily sales data of product B 56
4.6 The additive models fitted over average daily sales data for product A (a) and product B (b) 62
4.7 Brockhoff's additive model 63
4.8 Brockhoff's multiplicative model 63
4.9 The multiplicative models fitted over average daily sales data of product A (a) and product B (b) 64
4.10 The multiplicative model fitted over weekly sales of product A (a) and product B (b) in 2017 65
4.11 The multiplicative model fitted over daily sales of product A (a) and product B (b) in 2017. Left: Product A, right: product B. . . 67
4.12 Residuals vs. time (Left: product A , right: product B) 67

5.1 Failure data in raw form 70
5.2 Number of failures by month (product A) 70
5.3 Number of failures by month (product B 71
5.4 Lifetimes of uncensored Product A items sold in 2013 72
5.5 Lifetimes of uncensored Product B items sold in 2013 72
5.6 Hazard rate function of lifetimes of items sold in 2013 (Product A) 72
5.7 Hazard rate function of lifetimes of items sold in 2013 (Product B) 72
5.8 One-dimensional log-likelihood function and its gradient (Product A) 86
5.9 One-dimensional log-likelihood function and its gradient (Product B) 86
5.10 Empirical cdf and Weibull distribution with the estimated parameters for product A (a) and product B (b) 87

List of Tables

4.1 Goodness-of-fit measurements for the additive model 61
4.2 Goodness-of-fit measurements for the multiplicative model . . . 63
4.3 Weights for each day of the week 66

5.1 Failure data statistics 71
5.2 Frequency of failures during the warranty period 71

Chapter 1

INTRODUCTION

In this thesis, we develop a model to predict the number of failures (breakdowns) that the units of a consumer product experience in the future. This is an important problem for manufacturers as these failed items return to the firm for repairs, especially if the product is under warranty. The firm uses predictions of failures for various purposes including spare parts inventory planning and repair facility resource planning. These predictions also allow the firm to detect anomalies in product returns, which may indicate quality issues in the production process or with component suppliers.

A straightforward way to predict the number of failures would be using a time series analysis based on historical failure data. This approach, while practical, ignores the effect of the *installed product base*, that is the number of items sold and in use by customers. The inclusion of the installed product base information can enhance failure rate predictions as higher number of items in use would naturally increase the rate of failure observations. The installed base, in turn, is determined by the rate of product sales, which is shaped by product life cycle phases as well as seasonality.

In this thesis, we suggest a novel stochastic model of product failures that is based on two sub-models for the counting process of item sales, and failures of each individual item. Note that these sub-models provide value on their own as well, for sales forecasting and item-level failure analysis.

We test our model using the sales and failure data of two products (to which we refer to as product A and product B) of a major household manufacturing firm in Turkey. The data is item-level, that is, it provides the sales and failure date(s) of each individual item. The sales data contains the dates on which items are installed. Since the usage time of the item starts at the day of its installation, we assume the sales date of the item to be the same as its installation date. The failure data contains the dates on which item was returned to the firm in order to be repaired for any reason that is covered under the warranty policy. Assuming that a failed item is returned for repair immediately and that the repair takes no time, failure and repair times becomes the same. Therefore, we can model the failure process by using the repair data of the firm.

The life flow of the products we consider can be summarized as in Figure 1.1. Throughout the *warranty period*, customers demand repair services when their item fails. The repair performed can be either *minimal* or *major*. A minimal repair restores the item to the previous deterioration stage, while a major repair restores the item to good-as-new [21]. If an item is not returned to the firm during the warranty period, the firm does not know if and when it failed. Thus, the only information about the *lifetime* of that item will be that it exceeds the warranty length. Such data is referred to as *right-censored*, which is a typical situation with lifetime data. Developing a methodology that deals with censoring is a major challenge of lifetime data analysis.

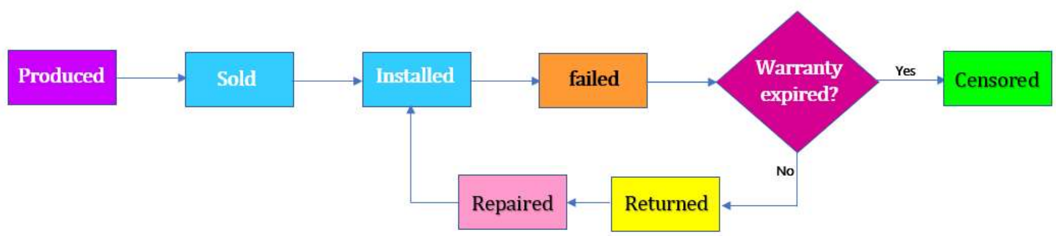


Figure 1.1: Life flow

While there has been numerous research on forecasting product failure rates [25, 19, 26, 49], only few researchers have considered failures as a collection of random variables and suggested stochastic processes for their analysis. The common approach is to use either a simple empirical forecasting method or a Box-Jenkins transfer function model to capture the dependence of the item returns (or the demand for spare parts) to the product installed base (or sales quantities). We contribute to literature by generating a stochastic model for both the total number of failures, and the accumulated repair cost over time.

The thesis is structured as follows. In Chapter 2, a review of the existing research on lifetime analysis is provided. Next, in Chapter 3, the mathematical model is presented. After the general continuous model for estimating the total repairs and their costs is provided, a section on the cumulative sales process is included. In Chapter 4, the sales data is analyzed and a proper parametric function is provided for its estimation. In Chapter 5, the failure data is analyzed and a Weibull distribution is fitted over the lifetime data. Finally, in Chapter 6, the results of the analysis are discussed thoroughly.

Chapter 2

LITERATURE REVIEW

Managing the return flow of failed items has added to the complexity of reverse logistics operations. Predicting the number of items that will be returned in some future period in advance will help maintain proper levels of inventory and resource allocation. Accordingly, predicting the failure rate of different products has been the focus of numerous works in literature [50, 49, 36, 26]. In one basic approach, to which Toktay et al. [49] refer to as "Naive estimation", the failure probability is estimated merely by dividing the cumulative failures to cumulative sales [50, 19]. In this case, the only information available is the total percentage of failures and not their timing. Later, data-driven regression models which capture the dependence of failures to some explanatory variables such as price, product category and the reason for return became widely accepted [20]. Certain studies that use this approach exploit the fact that current failures are a function of previous sales [49].

In this thesis, we propose a model based on point processes to estimate the number of failed items by a specific point in time. Our model can be decomposed into two parts: the estimation of sales and the estimation of failures by time t. In the following two sections, we provide an overview of the research conducted on these two areas.

2.1 Sales

A large number of works in literature estimate the sales quantities of a product, using mainly simple time series forecasting methods [44, 34]. Very few of these, however, consider the product life cycle, which provides a shape for the sales function of the product in its different life stages. With this approach, the future sales is not merely an extrapolation of the most recent sales trend. Instead, the change in the sales growth at different life stages is considered.

One of the works done in this area is of Brockhoff's [12]. This author assumes the life cycle to consist of three parts which respectively represent the sales growth during the initial periods, the sales decline in the later stages and overlapping of these two phases with the impact of sales of the firm's complementary or substitute products. To estimate the shape of the life cycle, the author proposes two models that respectively depend on the additive and the multiplicative combination of certain functions. A variant of the least squares method is used for calculating the standard error and the parameters of the life cycle.

Gomez et al. [33] identify three stages over the life cycle: sharply growing, stagnating in one level and slowly declining. Then, considering this asymmetry over the life cycle, they propose to use the Weibull distribution with two parameters to describe the appearance of the curve on the left and right side of the life cycle plot. Minner [35] proposes a logistic growth function with three parameters. The author uses the estimated sales amounts to calculate the installed base measure; i.e., the number of items that are still in use by the customers.

2.2 Failures

In this thesis we focus on the failures of durable products which are also repairable. Therefore, we use the terms failure and return interchangeably. Forecasting can either be done on period-level information; i.e., based on sales and failures volumes in each period, or on item-level information where the timing of sales and failures are tracked on an individual basis [49]. Toktay et al. [50] studied forecasting of product returns for disposable cameras. To model the returns flow, they assume the returns to be dependent on sales through an unknown return probability and delay distribution. Kelle and Silver [25], and Goh and Varaprasad [19] studied forecasting of returns for reusable containers. In all these environments, the available data are in period-level. The techniques used for such data types are usually based on the time series modeling methodology, which relates past return volumes to future return volumes. However, this approach ignores the effect of past sales data and therefore can result in inaccurate estimations. Goh and Varaprasad [19] propose to use the transfer function model of Box and Jenkins [9] which relates current failures to previous sales. Yet, this model neglects the fact that data is augmented in each period as new sales and failures data are revealed.

The distributed lag model of Toktay et al. [50] can conduct this augmentation. If we denote the number of realized sales at time t by N_t, $t = 1, ..., T$ and the random number of failures at time t by the random variable M_t, $t = 1, ..., T$

the functional form of a distributed lag model is given by

$$\mathbf{M}_t = \sum_{k=1}^{t-1} \beta_k N_{t-k} + \varepsilon_t \qquad t = 2, 3, ..., T \tag{2.2.1}$$

where β_k is the k^{th} reaction coefficient which represents the proportion of the realized sales N_{t-k} in period $t - k$ that contributes towards the number of failures in period t. The sequence of random variables ε_t, $t = 2, ..., T$ denote the sequence of independent measurement errors and T is the number of periods of data available for estimation. Toktay et al. [50] used geometric and negative binomial delay functions to represent the coefficients. Clottey et al. [17] employ a Bayesian approach with a distributed lag model using the continuous analog of the geometric function used by Toktay et al. [50]. Employing an exponential delay function, they obtained accurate estimates of the number of failures in future.

Black box methods such as Box-Jenkins forecast the future failures as a simple extrapolation of observed failures during the initial and mature phases [26]. This approach, because of its simplicity, became very popular; however it neglects the decrease of failure during the end-of-life phase of the product life cycle, leading to overestimation of the actual failures. To prevent this, Dekker et al. [26] use the installed base of the products at time t to forecast the number of spare parts that will be demanded.

Another issue is that failure data is usually censored. There are some categorizations regarding censored data. In a *right censoring* mechanism only a lower bound on the lifetime for some individual items are available. In this case, if the product's lifetime is larger than the warranty time, the firm will only know that the lifetime is larger than the warranty time. *Left-censored* point values are known only to be less than a specific value. In Figure 2.1 the failure times of five different items of a product are demonstrated. At the end of the experiment, the lifetimes of items 1 and 2 were realized, but the exact lifetimes of item 3 and 4 operational at the end of the experiment remain unknown.

Censored data can also be of type I, type II or randomly censored. In type I censoring, the censoring levels are known, hence the number of censored observations is a random variable, while in type II censoring the number of censored observations is fixed in advance. In a randomly censored sample the censoring levels and the number of censored observations are random outcomes. In this thesis, the data exhibits type I right censoring, which often arises when a study is conducted over a specified period of time.

A number of researchers have developed methodologies to address censor-

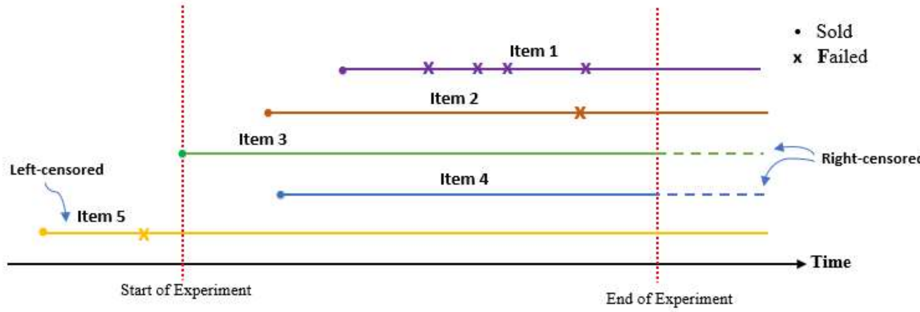


Figure 2.1: Censoring within the failure times data

ing [22, 29, 42, 6]. These methodologies can be either in non-parametric or parametric form. The non-parametric and graphical procedures introduced to estimate the distributional characteristics of the univariate lifetime data can be a simple relative-frequency table for censored data (life table methodology), or methods such as Kaplan-Meier (KM) Estimate.

The exact lifetime distribution theory for estimation of lifetimes is usually not available. Therefore, one must resort to approximations which are mainly based on Maximum-likelihood large-sample theory. Various parametric models are used in modeling the failure processes. Scholz [42] derives the maximum likelihood estimates of the Weibull regression model involving censored data. The author assumes a Gumbel distribution for the error terms, and provides unique maximum likelihood estimates.

Log-normal distribution is also used in modeling of lifetimes. Basak et al. [6] developed inferential methods from a sample of censored data of a three-parameter log-normal distribution. Kuş and Kaya [27] developed a maximum likelihood estimation procedure to estimate the parameters of log-logistic distributions for censored data.

Chapter 3

A MATHEMATICAL MODEL FOR THE TOTAL NUMBER OF FAILURES AND THE COST PROCESS

In this chapter we propose a mathematical model based on point processes to describe the total cumulative cost process and the total cumulative return process of failed items over time. In Section 3.1 we introduce a general continuous time model while in Section 3.2 we assume that the sales process is given by a non-homogeneous Poisson process. For a non-homogeneous Poisson sales process one can derive more detailed properties of the stochastic processes that represent the total number of returned items, and the total cost of repairing those items up to any time. In particular, under this assumption these stochastic processes belong to the class of *filtered Poisson processes* for which a lot of nice theoretical properties can be derived that will be useful in our statistical analysis.

In the proposed model we decompose the problem into a sales part and a failure part. Because the type of repair influences future failures, we discuss in this section different repair models available in the maintenance literature. To keep the model as simple as possible, and because it seems to be a realistic repair model for our statistical application, our main focus is on the *minimal repair* model.

3.1 A continuous time model for the total cost of repairing failed items and the total number of failed items

In this section we propose a continuous time stochastic model for the total number of failures over time of a particular product sold to customers and the total costs of repairs. To introduce the model let $(\Omega, \mathcal{H}, \mathbb{P})$ be a probability space hosting the point process $(\mathbf{T}_n)_{n\in\mathbb{N}}$ with $\mathbf{T}_n$, $n \in \mathbb{N}$ representing the random arrival time of the nth arriving customer who buys exactly one item of this product. It is assumed that this so-called *one-dimensional random point process* $(\mathbf{T}_n)_{n\in\mathbb{N}}$ is nonexplosive (see [11]) meaning the random variable $\mathbf{T}_n$ is finite with probability one for every n, $\mathbf{T}_n < \mathbf{T}_{n+1}$ for every n and

$$\mathbf{T}_\infty =\uparrow \lim_{n\uparrow\infty} \mathbf{T}_n = \infty \text{ a.s}$$

and denote by $\mathbf{S} = (\mathbf{S}(t))_{t\geq 0}$ the counting process of this point process $(\mathbf{T}_n)_{n\in\mathbb{N}}$ given by

$$\mathbf{S}(t) = \sum_{n=1}^{\infty} 1_{\{\mathbf{T}_n \leq t\}}. \tag{3.1.1}$$

Due to the interpretation of the random variable $\mathbf{T}_n$ the random variable $\mathbf{S}(t)$ counts the total number of arriving customers within the interval $[0, t]$. Since each customer buys exactly one product, the random variable $\mathbf{S}(t)$ also counts the cumulative sales up to time t. We also denote by Ψ the so-called *intensity measure* of the point process $(\mathbf{T}_n)_{n\in\mathbb{N}}$ (see [11] or section 1.4 of [28]) given by

$$\Psi(A) := \mathbb{E}(\sum_{n=1}^{\infty} 1_{\{\mathbf{T}_n\in A\}}) = \sum_{n=1}^{\infty} \mathbb{P}(\mathbf{T}_n \in A) \tag{3.1.2}$$

with A any Borel set on $\mathbb{R}_+$, and we introduce the right-continuous increasing function

$$\Psi(t) = \Psi([0, t]) = \sum_{n=1}^{\infty} F_n(t), t \geq 0 \tag{3.1.3}$$

with

$$F_n(t) = \mathbb{P}(\mathbf{T}_n \leq t). \tag{3.1.4}$$

Clearly $\Psi(t) = \mathbb{E}(\mathbf{S}(t))$ and to avoid pathological cases it is assumed in the remainder of this chapter that $\Psi(t) < \infty$ for every $t > 0$. This means that the expected amount of sales up to any time t is always finite. Also in the remainder of this chapter we call the item of the product bought by the nth arriving customer item n.

After item n is bought at the random time $\mathbf{T}_n$, it is used by the customer for a random amount of time, called the *usage time* of item n, which is denoted by the random variable $\mathbf{U}_n$. It is assumed that the sequence $(\mathbf{U}_n)_{n\in\mathbb{N}}$ is independent of the sequence $(\mathbf{T}_n)_{n\in\mathbb{N}}$ of selling times and that the random variables $\mathbf{U}_n$, $n \in \mathbb{N}$ are independent and identically distributed with cdf G satisfying $G(0) = 0$ having a finite first moment. For a lot of consumer goods which are not season-dependent the independence between selling times and usage time seems to be a reasonable assumption. It is also possible to include in the model that usage times depend on the time of sales but we will not discuss this extension in this thesis. Since we are interested in the total number of repairs of failed items and the corresponding costs of repairing these failed items up to any time t it is clear that the number of repaired items or the total repair costs up to any time depend on the products available in the market at any time t. Therefore we introduce the stochastic process $\mathbf{U} = \{\mathbf{U}(t) : t \geq 0\}$ with

$$\mathbf{U}(t) = \text{ number of items of a given product in use at time } t.$$

By its definition it is clear that

$$\mathbf{U}(t) = \sum\nolimits_{n=1}^{\infty} 1_{\{\mathbf{U}_n > t - \mathbf{T}_n\}} 1_{\{\mathbf{T}_n \leq t\}}. \tag{3.1.5}$$

Since the random variable $\mathbf{U}(t)$ is also known in the literature as the *installed base* at time t (see [26]), we refer to the stochastic process $\mathbf{U}$ as the *installed base stochastic process*. For this stochastic process it is easy to verify the following result.

Lemma 1 *For any non-explosive cumulative point process* $(\mathbf{T}_n)_{n \in \mathbf{N}}$ *it follows that*

$$\mathbb{E}(\mathbf{U}(t)) = \int_0^t (1 - G(t-u))\Psi(du), t \geq 0. \tag{3.1.6}$$

with $\Psi(t) = \mathbb{E}(\mathbf{S}(t))$*,* $t \geq 0$ *the intensity measure of the point process* $(\mathbf{T}_n)_{n \in \mathbb{N}}$.

Proof. By relation (3.1.5) and the independence of the sequences $(\mathbf{U}_n)_{n \in \mathbb{N}}$ and $((\mathbf{T}_n)_{n \in \mathbb{N}}$ we obtain using the monotone convergence theorem (see [14]) and the tower property for conditional expectations that

$$\begin{aligned}
\mathbb{E}(\mathbf{U}(t)) &= \mathbb{E}\left(\sum\nolimits_{n=1}^{\infty} 1_{\{\mathbf{X}_n > t - \mathbf{T}_n\}} 1_{\{\mathbf{T}_n \leq t\}}\right) \\
&= \sum\nolimits_{n=1}^{\infty} \mathbb{E}(1_{\{\mathbf{X}_n > t - \mathbf{T}_n\}} 1_{\{\mathbf{T}_n \leq t\}}) \\
&= \sum\nolimits_{n=1}^{\infty} \int_0^t (1 - G(t-u)) F_n(du) \\
&= \int_0^t (1 - G(t-u)) \Psi(du).
\end{aligned}$$

This shows the desired result. □

Usage times can be used in a lot of different situations. One situation is when a product is only used during the warranty period of length w and discarded by the customer after the expiration date of the warranty. This means that

$$\mathbf{U}_n \overset{a.s}{=} w.$$

This is the *warranty installed base* concept as defined in [26]. It is also possible that the customer uses the product much longer than the warranty period, thus in general the expected usage time of the product is much larger than the length of its warranty period. This is the *economic installed base* concept as defined in [26]. Most consumer household products belong to this class. If this scenario applies, the decision to stop using the product depends on the repair cost of failures occurring after the expiration date of the warranty. Another attitude that the consumers of a product might follow is to discard the product before its warranty ends. This usually happens for products with short life cycles. This is the *mixed installed base* concept as defined in [26].

Hence in general it is realistic to assume that the random usage time $\mathbf{U}_n$ of item n depends on the costs of repair in combination with the remaining value of the failed item. A repair of a consumer good is either done by a repair crew at the home address of the customer or by repairing the failed product at a repair facility. In both cases this generates repair costs for the manufacturer if a failure occurs during the warranty period, or for the customer after the expiration date of the warranty. In our proposed model we include these repair costs. After item n is sold at time $\mathbf{T}_n$ it generates during its usage time a sequence of failures and each failure causes a repair. It is assumed for simplicity that repairs do not take any time and a failed item after repair will always function again. An easy extension of this model would be to include the possibility that with a certain probability depending on the age of the defective item this item cannot be repaired and needs to be replaced. This extension can be analyzed using similar techniques but for simplicity we do not include this extension into this thesis.

To describe the counting process of failures and corresponding repairs of item n we introduce on the probability space $(\Omega, \mathcal{H}, \mathbb{P})$ a sequence of increasing random variables

$$0 < \overline{\mathbf{T}}_{1n} < \overline{\mathbf{T}}_{2n} < \overline{\mathbf{T}}_{3n} < ...$$

with $\mathbf{T}_n + \overline{\mathbf{T}}_{jn}$ representing the time of the jth failure of item n. The stochastic cumulative counting process $\mathbf{N}_n = \{\mathbf{N}_n(t) : t \geq 0\}$ describing the total number of repairs or failures of item n within the interval $[\mathbf{T}_n, \mathbf{T}_n + t)$ is then given by

$$\mathbf{N}_n(t) = \sum\nolimits_{j=1}^{\infty} 1_{\{\overline{\mathbf{T}}_{jn} \leq t\}}, t \geq 0. \tag{3.1.7}$$

Since the items are identical products, it seems reasonable to assume that for every $n \in \mathbb{N}$ the sequence $(\overline{\mathbf{T}}_{jn})_{j \in \mathbb{N}}$ of random variables are identically distributed and the sequences $(\overline{\mathbf{T}}_{jn})_{j \in \mathbb{N}}$, $n \in \mathbb{N}$ are independent. The assumption of independence might be restrictive but without this assumption it seems to be impossible to analyze the model. This implies that the counting processes $\mathbf{N}_n$, $n \in \mathbb{N}$ listed in relation (3.1.7) are independent and have the same probability law. As for the sales process in relation (3.1.2) we denote by Φ the intensity measure of the identically distributed point processes $(\overline{\mathbf{T}}_{jn})_{j\in\mathbb{N}}$, $n \in \mathbb{N}$ given by

$$\Phi(A) := \mathbb{E}\left(\sum\nolimits_{j=1}^{\infty} 1_{\{\overline{\mathbf{T}}_{jn}\in A\}}\right) = \sum\nolimits_{j=1}^{\infty} \mathbb{P}(\overline{\mathbf{T}}_{jn} \in A) \tag{3.1.8}$$

with A any Borel set on $\mathbb{R}_+$. As for the intensity measure of the sales process we also introduce the right-continuous function

$$\Phi(t) := \Phi([0,t]) = \mathbb{E}\left(\sum\nolimits_{j=1}^{\infty} 1_{\{\overline{\mathbf{T}}_{jn}\leq t\}}\right), t \geq 0 \tag{3.1.9}$$

The main remaining problem is to determine what the repair does to the state of the item since this determines the intensity measure Φ. In our statistical analysis we are dealing with household equipment that consist of a large number of components. Thus it seems reasonable to assume as an approximation that the *minimal repair* assumption holds. As observed in [15], the replacement of one failed component by an as-good-as-new component in a product consisting of a lot of components does not really change the failure rate of the product. Therefore, unless specified otherwise, we assume that the minimal repair assumption holds throughout this chapter. Of course, this assumption should be tested using our particular data sample of failures of individual items. This will be the topic of future research.

In this section we also discuss another type of repair process popular within the maintenance literature. It should be obvious that depending on the influence of the repair on the state of the product we obtain different intensity measures Φ, and most of these intensity measures can only be calculated using numerical evaluations. In the Appendix of [8] or [43] it is shown that the point process $\mathbf{N}_n$ of failures under the minimal repair assumption is given by a non-homogeneous Poisson process with arrival rate function the failure rate function of the cdf

$$F(t) = \mathbb{P}(\overline{\mathbf{T}}_{1n} \leq t) \tag{3.1.10}$$

with $\overline{\mathbf{T}}_{1n}$ the random operation time of item n until its first failure. It is assumed that this cumulative distribution function F satisfies $F(0) = 0$ and has a density f. Hence its failure rate function is given by

$$r(t) = \frac{f(t)}{1 - F(t)}$$

Since it is well-known that

$$1 - F(t) = e^{-\int_0^t r(s)ds} \tag{3.1.11}$$

this shows using the well known properties of a non-homogeneous Poisson process (see [13]) that the intensity measure of the minimal repair model defined in relation (3.1.9) is given by

$$\Phi(t) = \int_0^t r(u)du = -\ln(1 - F(t)), t \geq 0. \tag{3.1.12}$$

For more details on the minimal repair process, the reader is referred to [5].

In case the repair process brings the repaired item back to the *as good as new* state, it follows with $\overline{\mathbf{T}}_{0n} = 0$ that the random variables

$$\mathbf{X}_{jn} = \overline{\mathbf{T}}_{jn} - \overline{\mathbf{T}}_{j-1n}, j \in \mathbb{N}$$

are independent and identically distributed. This applies to items with only a few critical components. Since this is not likely to be the case for consumer goods, we only mention this type of repair process for completeness. Intermediate types of repair processes based on residual life time after a repair (so-called hazard rate repair models) can be found in [5]. The good as new condition implies that the intensity measure in relation (3.1.9) of the failure process is given by

$$\Phi(t) = \sum\nolimits_{j=1}^{\infty} \mathbb{P}(\overline{\mathbf{T}}_{jn} \leq t) = \sum\nolimits_{j=1}^{\infty} F^{j^*}(t), t \geq 0 \tag{3.1.13}$$

with F^{n^*} denoting the n-fold convolution of F. Observe it is well known for $F(0) = 0$ and F has a density f on $(0, \infty)$ that for every $n \in \mathbb{N}$

$$F^{n*}(t) = \int_0^t F^{(n-1)*}(t-u)f(u)du, t \geq 0$$

with

$$F^{0^*}(t) = 1, t \geq 0.$$

In this case the intensity measure Φ of the failure process is also called the *proper renewal function* associated with the cumulative distribution function F. (see [10], [23] and [3]). In this case only for very special cases a simple analytical formula is available. For a lot of details on the asymptotic behavior of the renewal function including close bounds the reader should consult [3] or [37] while for numerical approaches to evaluate numerically the renewal function the reader should consult [47] or [31].

If we are interested in the cost of repairs, we introduce for each item n bought at the random time $\mathbf{T}_n$, $n \in \mathbb{N}$ the sequence of random variables $(\mathbf{C}_{jn})_{j \in \mathbb{N}}$ with $\mathbf{C}_{jn}$ denoting the random cost of repairing the jth failure of item n. Since the items are identical products we assume for every $n \in \mathbb{N}$ that the sequence of random variables $(\mathbf{C}_{jn})_{j\in\mathbb{N}}$ are identically distributed and the sequences $(\mathbf{C}_{jn})_{j\in\mathbb{N}}$, $n \in \mathbb{N}$ are independent. Again the independence assumption is most restrictive but without this condition it seems difficult to analyze the model. Clearly the repair costs might also depend on the age of the item and so $\mathbf{C}_{jn}$ can be correlated with $\overline{\mathbf{T}}_{jn}$. Instead of counting the total number of repairs of item n within the interval $[\mathbf{T}_n, \mathbf{T}_n + t]$ we then need to introduce the stochastic process $\mathbf{C}_n = \{\mathbf{C}_n(t) : t \geq 0\}$ with $\mathbf{C}_n(t)$ denoting the total random cost within the interval $[\mathbf{T}_n, \mathbf{T}_n + t)$ of repairing item n . It is now obvious that the random variable $\mathbf{C}_n(t)$ is given by

$$\mathbf{C}_n(t) = \sum\nolimits_{j=1}^{\infty} \mathbf{C}_{jn} 1_{\{\overline{\mathbf{T}}_{jn} \leq t\}}, t \geq 0. \tag{3.1.14}$$

Since the repair costs are nonnegative, the cumulative costs processes $\mathbf{C}_n$, $n \in \mathbb{N}$ for each item n have increasing sample paths. By our assumption on the random cost $(\mathbf{C}_{jn})_{j\in\mathbb{N}}$ of a repair of item n and the point process $(\overline{\mathbf{T}}_{jn})_{j\in\mathbb{N}}$ of arrivals of defects of item n, it follows that the cumulative costs processes $\mathbf{C}_n, n \in \mathbb{N}$ have the same probability law and are independent. Taking into consideration the date of production of item n we can relax the assumption that the stochastic processes $\mathbf{C}_n, n \in \mathbb{N}$ have the same probability law. This extension is realistic since the technology of producing an item may change over time and this change of technology affects the probability law of the failure process. For simplicity we will not consider this extension as well.

Using relation (3.1.14) it follows by the monotone convergence theorem that

$$\mathbb{E}(\mathbf{C}_n(t)) = \sum\nolimits_{j=1}^{\infty} \mathbb{E}(\mathbf{C}_{jn} 1_{\{\overline{\mathbf{T}}_{jn} \leq t\}}). \tag{3.1.15}$$

To simplify this formula we need to make some additional assumptions. The simplest case to consider is to assume that the average cost of a repair at age y of any item n is given by $c(y)$ with $c : \mathbb{R}_+ \to \mathbb{R}_+$ some increasing function. This function can be estimated using cost data on repairs of items. Now it follows that

$$\mathbf{C}_n(t) = \sum\nolimits_{j=1}^{\infty} c(\overline{\mathbf{T}}_{jn}) 1_{\{\overline{\mathbf{T}}_{jn} \leq t\}} = \int_0^t c(u) d\mathbf{N}_n(u). \tag{3.1.16}$$

Using formula (3.1.15) this shows by the tower property of conditional expectations that

$$\mathbb{E}(\mathbf{C}_n(t)) = \sum\nolimits_{j=1}^{\infty} \int_0^t c(u) d\mathbb{P}(\overline{\mathbf{T}}_{jn} \leq u) = \int_0^t c(u) d\Phi(u) \tag{3.1.17}$$

with Φ the intensity measure of the failure process listed in relation (3.1.9). In particular under the minimal repair assumption we obtain using relation

(3.1.12) and (3.1.17) that

$$\mathbb{E}(\mathbf{C}_n(t)) = \int_0^t c(u)r(u)du. \tag{3.1.18}$$

By taking $c(y) = 1$ for every $y \geq 0$ we recover in relation (3.1.17) the total number $\mathbf{N}_n(t)$ of failures or repairs within the interval $[\mathbf{T}_n, \mathbf{T}_n + t]$. Another extension not to be considered in this thesis is that the random variable $\mathbf{C}_{jn}, j \in \mathbb{N}$ representing the cost of the jth repair is independent of the time of the jth failure. This might apply to items in which a failed component is always replaced by a completely new one and the repair costs are always given by the sum of the cost of that component and some fixed service cost. If this holds we obtain using relation (3.1.15) that

$$\mathbb{E}(\mathbf{C}_n(t)) = \sum\nolimits_{j=1}^{\infty} \beta_j \mathbb{P}(\overline{\mathbf{T}}_{jn} \leq t) \tag{3.1.19}$$

with $\beta_j := \mathbb{E}(\mathbf{C}_{jn})$. In general this seems to be difficult to analyze unless we specify the type of repair process. This will be the topic of future research.

Observe that the total number of repairs of a given item within a certain time interval can be seen as a counting process of a non-explosive point process starting at a different time $\mathbf{T}_n$, while the cumulative cost process of failures or repairs of a given item can be seen arising from a marked non-explosive point process with the markers given by the cost of a repair. We are now interested in the total number of failures up to time t and this stochastic process is denoted by the random process $\mathbf{R} = \{\mathbf{R}(t) : t \geq 0\}$. In the remainder of this thesis this process is called the *cumulative repair process*. The more general process we are also interested in is the total cost of repair process. We refer to this process as the *cumulative cost process* and denote it by $\mathbf{C} = \{\mathbf{C}(t) : t \geq 0\}$. To derive a representation of the cumulative repair process we observe that the repair process $\mathbf{N}_n$ of item n starts after it is bought at time $\mathbf{T}_n$, and the total number of recorded failures up to time t is given by the random variable $\mathbf{N}_n((t - \mathbf{T}_n) \wedge \mathbf{U}_n\})$ with $\mathbf{U}_n$ the usage time of item n and $\boldsymbol{\sigma} \wedge \boldsymbol{\tau} := \min\{\boldsymbol{\sigma}, \boldsymbol{\tau}\}$. Looking now at all the items of a particular product sold before time t and adding all the repairs it follows that the total number of repairs up to time $t \geq 0$ is given by

$$\mathbf{R}(t) = \sum\nolimits_{n=1}^{\infty} \mathbf{N}_n((t - \mathbf{T}_n) \wedge \mathbf{U}_n) 1_{\{\mathbf{T}_n \leq t\}}. \tag{3.1.20}$$

If we are interested in the stochastic process $\mathbf{R}_g = \{\mathbf{R}_g(t) : t \geq 0\}$ representing the total number of repairs of items under warranty up to any time t then it follows introducing

$$\mathbf{R}_g(t) = \text{total number of repairs under warranty up to time } t \tag{3.1.21}$$

that

$$\mathbf{R}_g(t) = \sum\nolimits_{n=1}^{\infty} \mathbf{N}_n((t-\mathbf{T}_n) \wedge w) 1_{\{\mathbf{T}_n \leq t\}} \tag{3.1.22}$$

with w denoting the length of the warranty period. This process shows the total cost of repairs for the manufacturer, as it is the manufacturer who pays for repairs that occur within the warranty period.

Clearly the stochastic processes $\mathbf{R}$ and $\mathbf{R}_g$ are non-negative integer valued stochastic processes. If we are interested in the cumulative cost process of all repairs we obtain in a similar way for every $t \geq 0$

$$\mathbf{C}(t) = \sum\nolimits_{n=1}^{\infty} \mathbf{C}_n((t-\mathbf{T}_n) \wedge \mathbf{U}_n) 1_{\{\mathbf{T}_n \leq t\}}. \tag{3.1.23}$$

with $\mathbf{C}_n$, $n \in \mathbf{N}$ the cumulative cost processes of each item n. Depending on whether the cost of repair is measured in integer or continuous values, the stochastic process $\mathbf{C}$ has state space $\mathbb{Z}_+$ or $\mathbb{R}_+$.

Although we are dealing with the cumulative cost process $\mathbf{C}$, the total cost of repair is not always covered by the manufacturer. After the expiration of the warranty, the customer has to pay these costs. Since the cost to be paid by the manufacturer only applies to items under warranty the total cost of repairs up to time t that is completely covered by the manufacturer is given by the stochastic process $\mathbf{C}_g = \{\mathbf{C}_g(t) : t \geq 0\}$ with

$$\mathbf{C}_g(t) = \sum\nolimits_{n=1}^{\infty} \mathbf{C}_n((t-\mathbf{T}_n) \wedge w) 1_{\{\mathbf{T}_n \leq t\}}. \tag{3.1.24}$$

with w representing the length of the warranty period.

In the next result we compute the expectation of the random variable $\mathbf{C}(t)$. Observe that the expected total costs to be paid by the manufacturer up to any time t, and the expected number of repairs up to any time t are a special case.

Lemma 2 *It follows for every $t \geq 0$ that*

$$\mathbb{E}(\mathbf{C}(t)) = \int_0^t \mu(t-u)\Psi(du) \tag{3.1.25}$$

with

$$\mu(s) = \mathbb{E}(\mathbf{C}_1(s \wedge \mathbf{U}_1)) = \mathbb{E}(\mathbf{C}_1(s) \wedge \mathbf{C}_1(\mathbf{U}_1)), s \geq 0. \tag{3.1.26}$$

denoting the expected total repair costs of item 1 within the interval $[\mathbf{T}_1, \mathbf{T}_1 + (s \wedge \mathbf{U}_1)]$ and Ψ the intensity measure of the sales process.

Proof. Since the random variable $\mathbf{C}_n(t \wedge \mathbf{U}_n)$ is independent of the random variable $\mathbf{T}_n$ for every $n \in \mathbb{N}$ and the random variables $\mathbf{C}_n(t \wedge \mathbf{U}_n), n \in \mathbb{N}$ are identically distributed, we obtain by relation (3.1.23) applying similar arguments as in Lemma 1 that

$$\begin{aligned}\mathbb{E}(\mathbf{C}(t)) &= \textstyle\sum_{n=1}^{\infty} \mathbb{E}(\mathbf{C}_n((t-\mathbf{T}_n) \wedge \mathbf{U}_n))1_{\{\mathbf{T}_n \leq t\}}) \\ &= \textstyle\sum_{n=1}^{\infty} \int_0^t \mu(t-u) F_n(du) \\ &= \textstyle\int_0^t \mu(t-u)\Psi(du)\end{aligned}$$

with Ψ the counting measure of the sales process listed in relation (3.1.3). Hence we have verified relation (3.1.25). The result in relation (3.1.26) follows using the non decreasing sample paths of the total cost process $\mathbf{C}_1$ of item 1 and we have shown the result. □

A special and important case of Lemma 2 listed in the next corollary is given by the expected number of cumulative repairs. In the statistical section of this thesis, we only consider the problem of estimating the expected number of failed items up to any time. This is because our data set does not include costs of repair but only sale and failure times.

Corollary 1 *It follows for every $t \geq 0$ that*

$$\mathbb{E}(\mathbf{R}(t)) = \int_0^t \mu(t-u)\Psi(du) \tag{3.1.27}$$

with the function $\mu : \mathbb{R}_+ \to \mathbb{R}_+$ given by

$$\mu(s) = \mathbb{E}(\mathbf{N}_1(s \wedge \mathbf{U}_1)). \tag{3.1.28}$$

In the above result we did not specify the type of repair applied to failed items. In the next result we give a simplified formula for the function μ under the minimal repair assumption.

Lemma 3 *If the repair process of items is of the minimal repair type and the cost process $\mathbf{C}_1$ is given by relation (3.1.16) then the cumulative cost process $\mathbf{C}_1$ has independent increments. Additionally, if the usage time $\mathbf{U}_1$ is a stopping time with respect to the filtration generated by the failure process $\mathbf{N}_1$ then for every $s \geq 0$*

$$\mathbb{E}(\mathbf{C}_1(s \wedge \mathbf{U}_1)) = \mathbb{E}\left(\int_0^{s \wedge \mathbf{U}_1} c(u)r(u)du\right) = \int_0^s c(u)r(u)\mathbb{P}(\mathbf{U}_1 > u)du \tag{3.1.29}$$

Proof. We know under the minimal repair assumption that the counting process $\mathbf{N}_1$ is a non-homogeneous Poisson process with arrival rate function $r(.)$ and this shows (see [14]) that the stochastic process $t \to \int_0^t c(y)d\mathbf{N}_n(y)$ has independent increments. Hence the process $\mathbf{M} = \{\mathbf{M}(t) : t \geq 0\}$ given by

$$\mathbf{M}(t) = \int_0^t c(u)d\mathbf{N}_1(u) - \int_0^t c(u)r(u)du \tag{3.1.30}$$

is a martingale. Since for every $s \geq 0$ the random variable $s \wedge \mathbf{U}_1$ is a bounded stopping time with respect to the filtration generated by the stochastic counting process, it follows by Doobs optimal stopping theorem for martingales (see [14]) that

$$\mathbb{E}(\mathbf{C}_1(s \wedge \mathbf{U}_1)) = \mathbb{E}\left(\int_0^{s\wedge \mathbf{U}_1} c(u)d\mathbf{N}_1(u)\right) = \mathbb{E}\left(\int_0^{s\wedge \mathbf{U}_1} c(u)r(u)du\right). \tag{3.1.31}$$

To simplify the above formula we observe

$$\int_0^{s\wedge \mathbf{U}_1} c(u)r(u)du = \int_0^s c(u)r(u)1_{\{\mathbf{U}_1 > u\}}du$$

and this implies by Fubinis theorem

$$\mathbb{E}\left(\int_0^{s\wedge \mathbf{U}_1} c(u)r(u)du\right) = \int_0^s c(u)r(u)\mathbb{P}(\mathbf{U}_1 > u\}du.$$

Hence we have verified the result. □

Again we mention as a corollary the following important special case of Lemma 3 if we only deal with the number of failures. It gives a formula for the expected number of defects of a particular sold item under the minimal repair assumption.

Corollary 2 *If the repair process is of the minimal repair type and the usage time $\mathbf{U}_1$ is a stopping time with respect to the filtration generated by the failure process $\mathbf{N}_1$ then for every $s \geq 0$*

$$\begin{aligned} \mathbb{E}(\mathbf{N}_1(s \wedge \mathbf{U}_1)) &= \mathbb{E}\left(\int_0^{s\wedge \mathbf{U}_1} r(u)du\right) \\ &= -\mathbb{E}(\ln(1 - F(s \wedge \mathbf{U}_1))) \end{aligned} \tag{3.1.32}$$

with F denoting the cumulative distribution function of the random time $\mathbf{T}_{11}$ until the first failure.

An immediate consequence of Lemma 2 and 3 is given by the next result.

Corollary 3 *If the repair process for each item is of the minimal repair type and is given by relation (3.1.16) and the usage time* $\mathbf{U}_1$ *is a stopping time with respect to the filtration generated by the failure counting process* $\mathbf{N}_1$ *then it follows for every* $t \geq 0$ *that*

$$\mathbb{E}(\mathbf{C}(t)) = \int_0^t \mu(t-u)\Psi(du)$$

with

$$\mu(s) = \int_0^s f(u)r(u)\mathbb{P}(\mathbf{U}_1 > u)du$$

and Ψ *the intensity measure of the sales process listed in relation (3.1.3).*

Proof. Apply Lemma 2 and 3. □

Again we list as a corollary the following special case of Corollary 3 related to the cumulative repair process. Actually this result will be of importance in our statistical section to obtain estimates of the expected cumulative repairs over time.

Corollary 4 *If the repair process is of the minimal repair type and the usage time* $\mathbf{U}_1$ *is a stopping time with respect to the filtration generated by the failure process* $\mathbf{N}_1$ *then for every* $t \geq 0$

$$\mathbb{E}(\mathbf{R}(t)) = \int_0^t \mu(t-u)\Psi(du), t \geq 0 \tag{3.1.33}$$

with

$$\mu(s) = -\mathbb{E}(\ln(1 - F(s \wedge \mathbf{U}_1)). \tag{3.1.34}$$

and F *denoting the cumulative distribution function of the random time* $\mathbf{T}_{11}$ *of item* 1 *until the first failure.*

Next, we shortly explain the relation between our proposed mathematical model and the statistical analysis that is presented in Chapters 4 and 5 of this thesis. The primary goal of this thesis is to obtain estimates of the expected number of returned defective items making use of the proposed model. Thus, it is sufficient to estimate the intensity measures of both the sales process and the failure process. In short we explain how these estimates are obtained. For more detailed information about the statistical techniques and used parametric forms one should consult Chapter 4 (estimating sales) and Chapter 5 (estimating failure).

As discussed in Chapter 4, the sales in each year follow a similar pattern. The sales are lower in the beginning and end of the year, and higher during the middle of the year. To capture this, we approximate each year's sales by the often-used parametric function $t \mapsto at^b e^{-ct}$ approximating the derivative $\Psi^{(1)}(t)$ of the intensity measure Ψ. Originally (see [12]) this parametric function is well known in the estimation of the total number of sales during the total life cycle of a product. In particular, time t is measured in days and we fit this parametric function to each year sales. This means we set for any parameters $a, b, c \geq 0$

$$\Psi(t) = a \int_0^t u^b e^{-cu} du, 0 \leq t \leq 365 \tag{3.1.35}$$

and for every $k \in \mathbb{N}$

$$\Psi(t) = \Psi(365k) + a \int_0^{t (\operatorname{mod} 365)} u^b e^{-cu} du, 365k \leq t \leq 365(k+1) \tag{3.1.36}$$

Observe for every $t \geq 0$ the notation $t (\operatorname{mod} 365)$ means

$$t (\operatorname{mod} 365) := t - \left\lfloor \frac{t}{365} \right\rfloor 365$$

with $\left\lfloor \frac{t}{365} \right\rfloor$ the largest non-negative integer smaller or equal to $\frac{t}{365}$. Using now our sales data over each year and employing the above parametric representation, we estimate the parameters a, b, c over each year separately. We also consider some alternative parametric functions in that chapter.

In Chapter 5 we estimate the intensity measure of the counting process of failures. Under the reasonable assumption of minimal repair, the estimation of the intensity measure Φ reduces to the estimation of the cumulative distribution function of the first time to failure. As already mentioned, this minimal repair assumption should be tested on our data set and this is part of our future research. Given now that this assumption is reasonable we use as an example in our statistical section the often-used parametric class of Weibull cumulative distribution functions to fit to the sample failure data. This is done by using the *maximum likelihood principle for censored data* which is discussed in Chapter 5. Note that the class of Weibull cumulative distribution functions is given by

$$F(x) = 1 - e^{-(x\beta^{-1})^{\alpha}}, \alpha \geq 0, \beta \geq 0 \tag{3.1.37}$$

This shows that

$$\ln(1 - F(x)) = (x\beta^{-1})^{\alpha}$$

and so by Corollary 4 it follows for this parametric class and the used parametric class for the sales process that

$$\mathbb{E}(\mathbf{R}(t)) = \beta^{-\alpha} \int_0^t \mathbb{E}((t-u)^\alpha \wedge \mathbf{U}_1^\alpha)\Psi(du) \tag{3.1.38}$$

with Ψ listed in relation (3.1.2) and (3.1.36). In our case study the policy of the manufacturer is to repair only items under warranty, and so in this case the random variable $\mathbf{U}_1$ is equal to the length w of the warranty period. Hence we obtain from relation (3.1.38) for $t < w$ that

$$\mathbb{E}(\mathbf{R}(t)) = \beta^{-\alpha} w^\alpha \Psi(t) \tag{3.1.39}$$

while for $t > w$

$$\begin{aligned} \mathbb{E}(\mathbf{R}(t)) &= \beta^{-\alpha} \textstyle\int_0^t ((t-u)^\alpha \wedge w^\alpha)\Psi(du) \\ &= \beta^{-\alpha} w^\alpha \Psi(t-w) + \beta^{-\alpha} \textstyle\int_{t-w}^t (t-u)^\alpha \Psi(du). \end{aligned} \tag{3.1.40}$$

As can be seen from relation (3.1.22), an exact interpretation of the above formula is the number of returned defective items under warranty given by $\mathbb{E}(\mathbf{R}_g(t))$.

To come up with a reasonable model for the usage time, denote by $\theta(y)$ the salvage value of an item of age y. Clearly the salvage value function θ is decreasing. A customer might now use the following reasonable decision rule: discard the item at the random time $\mathbf{U}_1$ with

$$\mathbf{U}_1 = \inf\{t > w : \mathbf{C}_1(t) - \mathbf{C}_1(w) \geq \theta(y)\} \tag{3.1.41}$$

with w representing the length of the warranty period. Note that under the minimal repair model

$$\mathbf{C}_1(t) - \mathbf{C}_1(w) = \int_w^t f(u) d\mathbf{N}_1(u)$$

denotes the cost the customer has to pay for repairs due to failures occurring after the warranty period. We will not continue with this approach and leave the details of deriving all kinds of properties for this chosen usage time under the minimal repair to future research.

However, in some practical cases due to the policy of the manufacturer the usage time can be independent of the failure process. An example is given by the following. The manufacturer might only repair items which fall under the warranty and to allocate resources a manufacturer likes to know an estimate of the total number of defective items which need to be repaired by him. If this applies, the manufacturer is now interested in the stochastic process $\mathbf{R}$ with a usage time given by $\mathbf{U} \overset{a.s}{=} w$ with w denoting the length of the warranty period. In this case the formula in relation (3.1.26) reduces to

$$\mu(s) = \mathbb{E}(\mathbf{N}_1(s \wedge w)), s \geq 0 \tag{3.1.42}$$

and it equals the total number of repairs on item n to be done by the manufacturer within the interval $[\mathbf{T}_n \mathbf{T}_n + s]$. Under the minimal repair assumption this function equals

$$\mu(s) = \int_0^{s \wedge w} r(u)du = -\ln(1 - F(s \wedge w)) \tag{3.1.43}$$

with F denoting the cumulative distribution function of the random first time to failure. Moreover,

$$\mu(\infty) = \lim_{s \uparrow \infty} \mu(s) = \mathbb{E}(\mathbf{N}_1(w)) = -\ln(1 - F(w)) \tag{3.1.44}$$

denotes the expected total number of times a repair occurs during the warranty period and

$$\int_0^w c(u)r(u)du$$

the expected repair costs for each particular item .

Although we are dealing with the cumulative cost process C, the total cost of repair is not always covered by the manufacturer. After the expiration of the warranty time the customer has to pay these costs. Because the cost of the manufacturer only applies to items under warranty, it follows by relation (3.1.24) that the expected total cost of repairs up to time t for items under warranty is given by

$$\mathbb{E}(\mathbf{C}_g(t)) = \int_0^t \mu(t-u)\Psi(du) \tag{3.1.45}$$

with (take $\mathbf{U}_1 = w$)

$$\mu(s) = \mathbb{E}(\mathbf{C}_1(s \wedge w)), s \geq 0 \tag{3.1.46}$$

As already observed, under the minimal repair assumption and using relation (5.2.15) it follows that

$$\mu(s) = \mathbb{E}(\mathbf{C}_1(s \wedge w)) = \int_0^{s \wedge w} f(u)r(u)du.$$

In some cases the usage time $\mathbf{U}_1$ is independent of the failure counting process $\mathbf{N}_1$ of item 1 (remember all items have identically distributed failure processess). If this independence assumption holds we can simplify the function μ for any type of failure process. Clearly in this case we obtain

$$
\begin{aligned}
\mu(s) &= \mathbb{E}(\mathbf{N}_1(s \wedge \mathbf{U}_1)) \\
&= \mathbb{E}(\mathbf{N}_1(s \wedge \mathbf{U}_1)1_{\{\mathbf{U}_1 \leq s\}}) + \mathbb{E}(\mathbf{N}_1(s \wedge \mathbf{U}_1)1_{\{\mathbf{U}_1 > s\}}) \\
&= \mathbb{E}(\mathbf{N}_1(\mathbf{U}_1)1_{\{\mathbf{U}_1 \leq s\}}) + \mathbb{E}(\mathbf{N}_1(s)1_{\{\mathbf{U}_1 > s\}}) \\
&= \int_0^s \Phi(u)G(du) + \Phi(s)(1 - G(s))
\end{aligned}
\tag{3.1.47}
$$

with Φ the intensity measure of the failure process listed in relation (3.1.9). If the usage time depends on the failure process and hence is a stopping time with respect to the filtration generated by the failure process, we might proceed for general type of repair processes as follows. Since the stochastic failure process $\mathbf{N}$ is a non-explosive counting process we need to determine the so-called *compensator* of that point processes. (see Chapter 3 of [4]). As shown in Definition 1.8 on page 53 of [4] most counting processes have as a compensator the stochastic process $\mathbf{A} = \{\mathbf{A}(t) : t \geq 0\}$ given by

$$
\mathbf{A}(t) = \int_0^t \mathbf{a}(s)ds \tag{3.1.48}
$$

with $(\mathbf{a}(t))_{t \geq 0}$ some non-negative progressively measurable stochastic process. The simplest example is a compensator for which the stochastic process $(\mathbf{a}(t))_{t \geq 0}$ is given by a deterministic function and this corresponds to the minimal repair model being a Poisson counting process. Also for the as-good-as-new condition type of repair process the stochastic process $(\mathbf{a}(t))_{t \geq 0}$ is given by (see [4] and [23])

$$
\mathbf{a}(s) = r(\boldsymbol{\Delta}(s))
$$

with $\boldsymbol{\Delta} = \{\boldsymbol{\Delta}(s) : s \geq 0\}$ the so-called age process associated with the failure times and r the failure rate function of the cdf F of the time to first failure listed in relation (3.1.10). Also by Theorem 10 of [4] the stochastic process $\mathbf{A}$ measurable with respect of the filtration generated by the counting process $\mathbf{N}$ is a compensator if and only the stochastic process $\mathbf{M} = \{\mathbf{M}(t) : t \geq 0\}$ given by

$$
\mathbf{M}(t) = \mathbf{N}(t) - \mathbf{A}(t) \tag{3.1.49}
$$

is a martingale (see Theorem 11 of [4]). Since the counting process satisfies clearly $\mathbf{N}(0) = 0$ and by relation (3.1.48) we obtain $\mathbf{A}(0) = 0$ we obtain $\mathbf{M}(0) = 0$. Applying now Doobs stopping theorem for martingales (see [14]) to this martingale M it follows for any stopping time $\boldsymbol{\tau}$ with respect to the filtration generated by the point process that

$$\mathbb{E}(\mathbf{M}(t \wedge \boldsymbol{\tau})) = \mathbb{E}(\mathbf{M}(0)) = 0$$

This shows by relation (3.1.49) that for any stopping time $\boldsymbol{\tau}$

$$\mathbb{E}(\mathbf{N}(t \wedge \boldsymbol{\tau})) = \mathbb{E}(\mathbf{A}(t \wedge \boldsymbol{\tau})) = \mathbb{E}\left(\int_0^{t \wedge \boldsymbol{\tau}} \mathbf{a}(s) ds\right). \tag{3.1.50}$$

Hence if the failure process of each item is a point process with a compensator process $\mathbf{A}$ given by relation (3.1.48) and the usage time $\mathbf{U}$ is a stopping time then we obtain

$$\mu(t) = \mathbb{E}\left(\int_0^{t \wedge \mathbf{U}} \mathbf{a}(s) ds\right)$$

and as already determined

$$\mathbb{E}(\mathbf{R}(t)) = \int_0^t \mu(t-u)\Psi(du). \tag{3.1.51}$$

This concludes our general discussion of the cumulative cost process $\mathbf{C}$ of defective products and the cumulative repair process $\mathbf{R}$ without using the probability law of both processess. If we want to derive more detailed properties of these stochastic processes like the variance or covariance function we need to impose conditions on the cumulative sales process. This is the topic of the next section.

3.2 On non-homogeneous Poisson cumulative sales processes.

In this section we restrict ourselves to a cumulative sales process represented by a non-homogeneous Poisson process with arrival intensity function ψ and a compound counting process of failure costs. Under this framework the cumulative cost process $\mathbf{C}$ or its special case the cumulative repair process $\mathbf{R}$ is a so-called filtered Poisson process (see [38]). Although a more general class of cumulative sales processes to consider is the class of renewal processes equipped with a time transformation and so the total cost process is a so-called filtered renewal processes (see [1]) we only discuss in detail cumulative sales processes given by a non-homogeneous Poisson process. A very interesting survey on more sophisticated techniques then used here is given in Chapter 6 on Poisson random measures of [14]. Some of the results listed here are special cases of general results discussed in that chapter. Also a very good survey on Poisson processes and filtered Poisson processes is given in Chapter 4 of [38]. As already observed at the end of the previous section it seems that an expression for the cumulative expected cost $\mathbb{E}(\mathbf{C}(t)), t \geq 0$ is the only general result for the cumulative cost process C one can derive without imposing additional assumptions on the probability law of the cumulative sales process.

To derive more detailed results like the cumulative distribution function of the random variable $\mathbf{C}(t)$ we need to assume that the cumulative sales process S is a non-homogeneous Poisson process with arrival intensity function ψ. For such a process it is well known that the expected cumulative sales up to time t is given by (see relation (3.1.3)

$$\Psi(t) := \int_0^t \psi(u)du. \tag{3.2.1}$$

Observe this is the intensity measure of the counting process S. Hence by Lemma 2 we obtain for a sales process modelled by a non-homogeneous Poisson process that

$$\mathbb{E}(\mathbf{C}(t)) = \int_0^t \psi(u)\mathbb{E}(\mathbf{C}_1((t-u) \wedge \mathbf{U}_1)du \tag{3.2.2}$$

In a lot of statistical applications a parametric model is fitted to cumulative sales data and so we obtain an estimation of the expected cumulative sales up to time t for every $t \geq 0$. If this estimated expected cumulative sales function has a derivative function $\psi(.)$ we may assume as a *first moment* approximation that the total sales process is represented by a non-homogeneous Poisson process with arrival intensity function ψ. Using this approach we obtain a first moment approximation of the cumulative costs process. Note this is related to the widely used practice in engineering to determine the cdf of a random variable by a so-called two-moment fit within a selected class of cumulative distributions functions. Well known examples of these classes are Gamma and Log-normal distributions (see Appendix B of [47]). Hence next to first moments we are also be interested in second and higher moments and in general deriving an expression for the cumulative distribution function of the random variable $\mathbf{C}(t)$. To derive an expression for the probability Laplace-Stieltjes tranform of this random $\mathbf{R}(t)$ under the assumption that the sales process is a non-homogeneous Poisson process we need the following result.

Lemma 4 *If the cumulative sales process* $\mathrm{S} = \{\mathbf{S}(t) : t \geq 0\}$ *is a non-homogeneous Poisson process with Borel arrival intensity function* ψ *independent of the independent and identically distributed cumulative cost processes*

$\mathbf{C}_n, n \in \mathbb{N}$ *listed in relation (3.1.14) then for every* $k \in \mathbb{Z}_+$ *and* $s \geq 0$

$$\mathbb{E}(e^{-s\mathbf{C}(t)} \mid \mathbf{S}(t) = k) = \mathbb{E}\left(e^{-s\mathbf{C}_1((t-\mathbf{Y}_1)\wedge \mathbf{U}_1)}\right)^k$$

with the random variable $\mathbf{Y}_1$ *concentrated on* $[0,t]$ *and independent of the stochastic process* $\{\mathbf{C}_1(s \wedge \mathbf{U}_1) : s \geq 0\}$ *having cdf*

$$F(y) = \frac{\Psi(y)}{\Psi(t)}, 0 \leq y \leq t. \tag{3.2.3}$$

with Ψ *listed in relation (3.2.1).*

Proof. By the properties of a non-homogeneous Poisson process (see [41] or [13]) it follows for every $s \geq 0$ that

$$\mathbb{E}(e^{-s\mathbf{C}(t)} \mid \mathbf{S}(t) = k) = \mathbb{E}\left(e^{-s\sum_{n=1}^{k}\mathbf{C}_n((t-\mathbf{Y}_{n:k})\wedge \mathbf{U}_n)}\right) \tag{3.2.4}$$

with $(\mathbf{Y}_{1:k}, ..., \mathbf{Y}_{k:k})$ the joint order statistics of a sequence of independent and identically distributed random variables $\mathbf{Y}_n, n = 1, ..., k$ on $[0,t]$ having the continuous cumulative distribution function

$$F(y) = \mathbb{P}(\mathbf{Y} \leq y) = \frac{\Psi(y)}{\Psi(t)}, 0 \leq y \leq t. \tag{3.2.5}$$

Introducing the set Π_k of all permutations π on $\{1, ..., k\}$ and for every $\pi \in \Pi_k$ the event

$$E_\pi = \{(\mathbf{Y}_1, ..., \mathbf{Y}_k) : \mathbf{Y}_{\pi(1)} \leq .... \leq \mathbf{Y}_{\pi(k)}\}$$

it follows using the random vector $(\mathbf{Y}_1, ..., \mathbf{Y}_k)$ has a continuous joint cumulative distribution function that

$$\begin{aligned}
&\mathbb{E}\left(e^{-s\sum_{n=1}^{k}\mathbf{C}_n((t-\mathbf{Y}_{n:k})\wedge \mathbf{U}_n)}\right) \\
= \;& \textstyle\sum_{\pi\in\Pi_k} \mathbb{E}\left(e^{-s\sum_{n=1}^{k}\mathbf{C}_n((t-\mathbf{Y}_{\pi(n)})\wedge \mathbf{U}_n)} 1_{E_\pi}\right) \\
= \;& \textstyle\sum_{\pi\in\Pi_k} \int\cdots\int_{[0,t]^k} \mathbb{E}\left(e^{-s\sum_{n=1}^{k}\mathbf{C}_n((t-y_{\pi(n)})\wedge \mathbf{U}_n)}\right) 1_{\overline{E}_\pi}(y_1, ..., y_k) dF(y_1, ..., y_k)
\end{aligned} \tag{3.2.6}$$

with

$$\overline{E}_\pi = \{(y_1, ..., y_k) : y_{\pi(1)} \leq y_{\pi(2)} \leq .... \leq y_{\pi(k)}\} \tag{3.2.7}$$

and F the joint cumulative distribution function of the random vector $(\mathbf{Y}_1, ..., \mathbf{Y}_k)$ given by

$$F(y_1, ..., y_k) = \Psi(t)^{-n}\Pi_{n=1}^{k}\Psi(y_n).$$

Since the cumulative cost processes $\mathbf{C}_n(. \wedge \mathbf{U}_n), n \in \mathbb{N}$ are by assumption independent and identically distributed and clearly for every $\pi \in \Pi_k$

$$\sum_{n=1}^{k} \mathbf{C}_{\pi(n)}((t - y_{\pi(n)}) \wedge \mathbf{U}_{\pi(n)}) = \sum_{n=1}^{k} \mathbf{C}_n((t - y_n) \wedge \mathbf{U}_n)$$

it follows for every $y \in \overline{E}_\pi$ that

$$\begin{aligned}
\mathbb{E}\left(e^{-s\sum_{n=1}^{k}\mathbf{C}_n((t-y_{\pi(n)})\wedge\mathbf{U}_n)}\right) &= \Pi_{n=1}^{k}\mathbb{E}\left(e^{-s\mathbf{C}_n((t-y_{\pi(n)})\wedge\mathbf{U}_n)}\right) \\
&= \Pi_{n=1}^{k}\mathbb{E}\left(e^{-s\mathbf{C}_{\pi(n)}((t-y_{\pi(n)})\wedge\mathbf{U}_{\pi(n)})}\right) \\
&= \mathbb{E}\left(e^{-s\sum_{n=1}^{k}\mathbf{C}_{\pi(n)}((t-y_{\pi(n)})\wedge\mathbf{U}_{\pi(n)})}\right) \\
&= \mathbb{E}\left(e^{-s\sum_{n=1}^{k}\mathbf{C}_n((t-y_n)\wedge\mathbf{U}_n)}\right) \\
&= \mathbb{E}\left(e^{-s\mathbf{C}_1((t-y_1)\wedge\mathbf{U}_1)}\right)^k.
\end{aligned}$$

This shows applying relation (3.2.6)

$$\begin{aligned}
&\mathbb{E}\left(e^{-s\sum_{n=1}^{k}\mathbf{C}_n((t-\mathbf{Y}_{n:k})\wedge\mathbf{U}_n)}\right) \\
=\ &\sum_{\pi\in\Pi_k}\int\cdots\int_{[0,t]^k}\mathbb{E}\left(e^{-s\mathbf{C}_1((t-y_1))\wedge\mathbf{U}_1)}\right)^k 1_{\overline{E}_\pi}(y_1,...,y_k)dF(y_1,...,y_k) \\
=\ &\int_0^t\mathbb{E}\left(e^{-s\mathbf{C}_1((t-y_1))\wedge\mathbf{U}_1)}\right)^k dF(y_1) \\
=\ &\mathbb{E}\left(e^{-s\mathbf{C}_1((t-\mathbf{Y}_1)\wedge\mathbf{U}_1)}\right)^k
\end{aligned}$$

and using relation (3.2.4) we obtain the desired result. □

A special case of Lemma 4 applied to the cumulative repair process $\mathbf{R}$ is given by

$$\mathbb{E}(e^{-s\mathbf{R}(t)} \mid \mathbf{S}(t) = k) = \mathbb{E}(e^{-s\mathbf{N}_1((t-\mathbf{Y}_1)\wedge\mathbf{U}_1)})^k \tag{3.2.8}$$

with $\mathbf{N}_1$ the failure counting process of item 1.

Before discussing the next result we need to introduce the following well known definition (see [45]).

Definition 1 *Let* X *be a non-negative random variable on the probability space* $(\Omega, \mathcal{H}, \mathbb{P})$. *The function* $\pi_{\mathbf{X}} : [0, \infty) \to [0, 1]$ *defined by*

$$\pi_{\mathbf{X}}(s) = \mathbb{E}(e^{-s\mathbf{X}}), s \geq 0 \tag{3.2.9}$$

is called the probability Laplace-Stieltjes transform (pLSt) of the random variable $\mathbf{X}$. *If the random variable* $\mathbf{X}$ *is also integer valued the function* $P_{\mathbf{X}} : \mathbb{D} \to \mathbb{C}$ *with* $\mathbb{D} = \{z \in \mathbb{C} : \mid z \mid \leq 1\} \subseteq \mathbb{C}$ *the unit disk in the complex plane* $\mathbb{C}$ *defined by*

$$P_{\mathbf{X}}(z) = \mathbb{E}(z^{\mathbf{X}}), z \in \mathbb{D} \tag{3.2.10}$$

is called the probability generating function (pgf) of the random variable $\mathbf{X}$.

It is well known (see [51]) that the probability Laplace-Stieltjes transform of the random variable $\mathbf{X}$ uniquely determines the underlying right continuous cumulative distribution function of the same random variable. We also introduce for every $t \geq 0$ the functions $Q_t : \mathbb{R}_+ \to \mathbb{R}$ given by

$$Q_t(s) := \int_0^t \psi(y)\mathbb{E}(e^{-s\mathbf{C}_1((t-y)\wedge \mathbf{U}_1)})dy \tag{3.2.11}$$

It is obvious that

$$Q_t(0) = \int_0^t \psi(y)dy = \Psi(t)$$

with Ψ the intensity measure of the sales process. One can now show the following important result about the probability Laplace-Stieltjes transform of the random variable $\mathbf{C}(t)$ for any $t \geq 0$.

Theorem 1 *If the cumulative sales process* $\mathbf{S} = \{\mathbf{S}(t) : t \geq 0\}$ *is a non-homogeneous Poisson process with Borel arrival intensity function* ψ *and this cumulative sales process is independent of the independent and identically distributed cumulative cost processes* $\mathbf{C}_n, n \in \mathbb{N}$ *then the probability Laplace-Stieltjes transform of the random variable* $\mathbf{C}(t)$ *is given by*

$$\pi_{\mathbf{C}(t)}(s) = \mathbb{E}(e^{-s\mathbf{C}(t)}) = e^{-(Q_t(0)-Q_t(s))}, s \geq 0 \tag{3.2.12}$$

with

$$Q_t(s) := \int_0^t \psi(y)\mathbb{E}(e^{-s\mathbf{C}_1((t-y)\wedge \mathbf{U}_1)})dy, s \geq 0. \tag{3.2.13}$$

Proof. Since the cumulative sales process $\mathbf{S}$ is a non-homogeneous Poisson process with Borel intensity function ψ we obtain

$$\begin{aligned}\mathbb{E}(e^{-s\mathbf{C}(t)}) &= \textstyle\sum_{k=0}^{\infty} \mathbb{E}\left(e^{-s\mathbf{C}(t)} \mid \mathbf{S}(t) = k\right)\mathbb{P}(\mathbf{S}(t) = k) \\ &= e^{-\Psi(t)}\textstyle\sum_{k=0}^{\infty} \mathbb{E}\left(e^{-s\mathbf{C}(t)} \mid \mathbf{S}(t) = k\right)\frac{\Psi(t)^k}{k!}.\end{aligned}$$

Applying Lemma 4 it follows

$$\begin{aligned}\mathbb{E}(e^{-s\mathbf{C}(t)}) &= e^{-\Psi(t)}\sum_{k=0}^{\infty}\frac{\left(\mathbb{E}\left(e^{-s\mathbf{C}_1((t-\mathbf{Y}_1)\wedge\mathbf{U}_1)}\right)\Psi(t)\right)^k}{k!} \\ &= e^{-\Psi(t)\left(1-\mathbb{E}\left(e^{-s\mathbf{C}_1((t-\mathbf{Y}_1)\wedge\mathbf{U}_1)}\right)\right)}.\end{aligned} \tag{3.2.14}$$

Since the stochastic process $C_1(t\wedge\mathbf{U}_1)$ is independent of the arrival process and hence independent of the random variable $\mathbf{Y}_1$ having cdf F listed in relation (3.2.3) it follows that

$$\begin{aligned}\Psi(t)\mathbb{E}(e^{-s(\mathbf{C}_1((t-\mathbf{Y}_1)\wedge\mathbf{U}_1)}) &= \Psi(t)\int_0^t\frac{\psi(y)}{\Psi(t)}\mathbb{E}(e^{-s(\mathbf{C}_1((t-y)\wedge\mathbf{U}_1)})dy \\ &= \int_0^t\psi(y)\mathbb{E}(e^{-s(\mathbf{C}_1((t-y)\wedge\mathbf{U}_1)})dy \\ &= Q_t(s)\end{aligned} \tag{3.2.15}$$

Substituting relation (3.2.15) and $Q_t(0)=\Psi(t)$ into relation (3.2.14) yields the desired result. □

In the statistical section, because we do not have cost data, we are interested in the estimation of the parameters of the cumulative stochastic repair process $\mathbf{R}$. Hence, in the next corollary we mention the probability Laplace-Stieltjes transform of the random variable $\mathbf{R}(t)$. Clearly this is a special case of Theorem 1.

Corollary 5 *If the cumulative sales process* $\mathbf{S}=\{\mathbf{S}(t):t\geq 0\}$ *is a non homogeneous Poisson process with Borel arrival intensity function* ψ *and this cumulative sales process is independent of the independent and identically distributed cumulative cost processes* $\mathbf{C}_n, n\in\mathbb{N}$ *then the probability Laplace-Stieltjes transform of the random variable* $\mathbf{R}(t)$ *is given by*

$$\pi_{\mathbf{R}(t)}(s)=\mathbb{E}(e^{-s\mathbf{R}(t)})=e^{-(Q_t(0)-Q_t(s))}, s\geq 0. \tag{3.2.16}$$

with

$$Q_t(s):=\int_0^t\psi(y)\mathbb{E}(e^{-s\mathbf{N}_1((t-y)\wedge\mathbf{U}_1)})dy. \tag{3.2.17}$$

Since the cumulative repair process $\mathbf{R}$ is an integer-valued process then by exactly the same arguments we obtain for every $z\in\mathbb{D}$

$$P_{\mathbf{R}(t)}(z)=\mathbb{E}(z^{\mathbf{R}(t)})=e^{-(Q_t(1)-Q_t(z))} \tag{3.2.18}$$

with

$$Q_t(z) := \int_0^t \psi(y)\mathbb{E}(z^{\mathbf{N}_1((t-y)\wedge \mathbf{U}_1)})dy = \int_0^t \psi(y)P_{\mathbf{N}_1((t-y)\wedge \mathbf{U}_1)}(z)dy. \tag{3.2.19}$$

with P denoting the probability generating function.

Up to now we did not specify in our model the type of repair process. Clearly this is needed to simplify the expression for Q_t both in relations (3.2.13) and (3.2.17). In this chapter we mainly consider the special case that $\mathbf{U}_1$ is independent of the costs process $\mathbf{C}_1$ and its special case $\mathbf{U}_1 \overset{a.s}{=} w$ with w denoting the length of the warranty period. Observe that the last case covers the total repair costs to be paid be the manufacturer due to warranty obligations. The much more practical case that the random variable $\mathbf{U}_1$ is a stopping time with respect to the cost process (see relation (3.1.41) for an example of such a stopping time) will also be discussed for the minimal repair model. However in this case we can only derive some relations which might be of use in future research.

If the usage time $\mathbf{U}_1$ of an item is independent of the cost of repair process of that item we obtain

$$\begin{aligned} \mathbb{E}\left(e^{-s\mathbf{C}_1(t\wedge \mathbf{U}_1)}\right) &= \mathbb{E}(e^{-s\mathbf{C}_1(t\wedge \mathbf{U}_1)}1_{\{\mathbf{U}_1>t\}}) + \mathbb{E}(e^{-s\mathbf{C}_1(t\wedge \mathbf{U}_1)}1_{\{\mathbf{U}_1\leq t\}}) \\ &= \mathbb{E}(e^{-s\mathbf{C}_1(t)}1_{\{\mathbf{U}_1>t\}}) + \mathbb{E}(e^{-s\mathbf{C}_1(\mathbf{U}_1)}1_{\{\mathbf{U}_1\leq t\}}) \\ &= \mathbb{E}(e^{-s\mathbf{C}_1(t)})(1-G(t)) + \textstyle\int_0^t \mathbb{E}(e^{-s\mathbf{C}_1(u)})G(du) \end{aligned} \tag{3.2.20}$$

with G the cdf of the usage time $\mathbf{U}_1$. For the special case of $\mathbf{U}_1 \overset{a.s}{=} w$ with w the length of the warranty period of the product we obtain the simplified formula

$$\mathbb{E}\left(e^{-s\mathbf{C}_1(t\wedge \mathbf{U}_1)}\right) = \mathbb{E}\left(e^{-s\mathbf{C}_1(t\wedge w)}\right). \tag{3.2.21}$$

By a similar proof we obtain for the cumulative failure process and the random usage time $\mathbf{U}_1$ independent of the failure process that

$$\mathbb{E}(e^{-s\mathbf{N}_1(t\wedge \mathbf{U}_1)}) = \mathbb{E}(e^{-s\mathbf{N}_1(t)})(1-G(t)) + \int_0^t \mathbb{E}(e^{-s\mathbf{N}_1(u)})G(du). \tag{3.2.22}$$

For $\mathbf{U}_1 \overset{a.s}{=} w$ it follows that

$$\mathbb{E}(e^{-s\mathbf{N}_1(t\wedge \mathbf{U}_1)}) = \mathbb{E}(e^{-s\mathbf{N}_1(t\wedge w)}). \tag{3.2.23}$$

Looking at the formulas in relations (3.2.13) and (3.2.20) for the cumulative cost process and relations (3.2.17) and (3.2.22) for the cumulative repair process we need to derive for the evaluation of the function Q_t a more detailed expression for the probability Laplace -Stieltjes transform of the individual item cumulative cost or failure process. If we assume that repairs are minimal repairs not changing the age of the item and we consider the item-dependent cumulative cost process given by

$$\mathbf{C}_1(t) = \sum\nolimits_{j=1}^{\infty} c(\mathbf{T}_{j1}) 1_{\{\mathbf{T}_{j1} \leq t\}} = \int_0^t c(y) d\mathbf{N}_1(y) \tag{3.2.24}$$

we can simplify these formulas. Under these assumptions the next result holds. Note that the proof of this result is actually easier than the proof of Lemma 4 and for completeness it is listed.

Lemma 5 *If the repair process of items is a minimal repair process and the cumulative cost process is given by relation (3.2.24) then it follows that the stochastic process C_1 has independent increments and*

$$\pi_{\mathbf{C}_1(t)}(s) = \mathbb{E}(e^{-s\mathbf{C}_1(t)}) = e^{-(\overline{Q}_t(0) - \overline{Q}_t(s))}, s \geq 0 \tag{3.2.25}$$

with

$$\overline{Q}_t(s) := \int_0^t r(y) e^{-sc(y)} dy \tag{3.2.26}$$

and r denoting the failure rate function of the cumulative distribution function F of the random time to the first failure.

Proof. Since the repair process is a minimal repair process we know that the counting process $\mathbf{N}_1$ is a non-homogeneous Poisson process having independent increments. It is now obvious using the definition of the cost process in relation (3.2.24) that the cost process $\mathbf{C}_1$ has independent increments. Also we know that the arrival rate function of the counting process $\mathbf{N}_1$ of failures is given by r with r denoting the failure rate function of the cumulative distribution function of the random time to the first failure. Hence it follows by the tower property that

$$\begin{aligned} & \mathbb{E}(e^{-s\mathbf{C}_1(t)}) \\ = \ & \mathbb{P}(\mathbf{N}_1 = 0) + \sum\nolimits_{n=1}^{\infty} \mathbb{E}(e^{-s\mathbf{C}_1(t)} \mid \mathbf{N}_1(t) = n) \mathbb{P}(\mathbf{N}_1(t) = n) \\ = \ & e^{-\Phi(t)} \left(1 + \sum\nolimits_{n=1}^{\infty} \mathbb{E}(e^{-s\mathbf{C}_1(t)} \mid \mathbf{N}_1(t) = n) \tfrac{\Phi(t)^n}{n!}\right) \end{aligned} \tag{3.2.27}$$

with $\Phi(t) = \int_0^r r(u) du$ the intensity measure of the counting process $\mathbf{N}_1$. By the well known properties of a non-homogeneous Poisson process (see [13]) with arrival rate function r we obtain that

$$\mathbb{E}(e^{-s\mathbf{C}_1(t)} \mid \mathbf{N}_1(t)=k) = \mathbb{E}\left(e^{-s\sum_{n=1}^{k} c(\mathbf{Z}_{n:k})}\right)$$

with $(\mathbf{Z}_{1:k},...\mathbf{Z}_{k:k})$ the joint order statistics of a sequence of independent and identivally distributed random variables $\mathbf{Z}_n, n=1,...,k$ on $[0,t]$ with cumulative distribution function

$$\mathbb{P}(\mathbf{Z}_1 \leq z) = \frac{\Phi(z)}{\Phi(t)}, 0 \leq z \leq t.$$

This implies using

$$\sum\nolimits_{k=1}^{n} c(\mathbf{Z}_{k:n}) = \sum\nolimits_{k=1}^{n} c(\mathbf{Z}_k)$$

that by the independence of the random variables $\mathbf{Z}_k$, k=1,....,n and hence the independence of the random variables $f(\mathbf{Z}_k), k=1,....,n$

$$\begin{aligned}
\mathbb{E}\left(e^{-s\sum_{k=1}^{n} c(\mathbf{Z}_{k:n})}\right) &= \mathbb{E}\left(e^{-s\sum_{k=1}^{n} c(\mathbf{Z}_k)}\right) \\
&= \mathbb{E}\left(\Pi_{k=1}^{n} e^{-sc(\mathbf{Z}_k)}\right) \\
&= \mathbb{E}(e^{-sc(\mathbf{Y}_1)})^k
\end{aligned}$$

This shows by relation (3.2.27)

$$\mathbb{E}(e^{-s\mathbf{C}_1(t)}) = e^{-\Phi(t)(1-\mathbb{E}(e^{-sc(\mathbf{Y}_1)}))}$$

and since

$$\Phi(t)\mathbb{E}(e^{-sc(\mathbf{Y}_1)})) = \Phi(t)\int_0^t \frac{r(y)}{\Phi(t)} e^{-sc(y)dy} = \int_0^t r(y)e^{-sc(y)}dy = \overline{Q}_t(s)$$

the result follows. □

Applying Lemma 5 and relation (3.2.20) it follows for $\mathbf{U}_1$ independent of the individual item cumulative cost process $\mathbf{C}_1$ that

$$\mathbb{E}(e^{-s\mathbf{C}_1(t\wedge\mathbf{U}_1)}) = e^{-(\overline{Q}_t(0)-\overline{Q}_t(s))}(1-G(t)) + \int_0^t e^{-(\overline{Q}_t(0)-\overline{Q}_u(s))}G(du) \quad \textbf{(3.2.28)}$$

with $\overline{Q}_u(s)$ listed in relation (3.2.26). For the special case $\mathbf{U}_1 \overset{a.s}{=} w$ we obtain

$$\mathbb{E}(e^{-s\mathbf{C}_1(t\wedge\mathbf{U}_1)}) = e^{-(\overline{Q}_t(0)-\overline{Q}_{t\wedge w}(s))}.$$

Using relation (3.2.13) we can therefore numerically evaluate $Q_t(s)$. Hence we are able to numerical evaluate the probability Laplace-Stieltjes transform of the random variable $\mathbf{C}(t)$ for different values of s and we now might use inversion techniques for Laplace-Stieltjes transforms (see Appendix F of [48]) to calculate numerically the cumulative distribution function of the random variable $\mathbf{C}(t)$. This might be a topic of future research in case data on cost of repairs are available and it is tested on our particular data set that the assumption of minimal repair is a reasonable assumption. An example of such an approach applied to the integer valued random variable $\mathbf{R}(t)$ is discussed at the end of this chapter.

In Lemma 5 we show that the individual item cost process has independent increments. To evaluate for such a cost process the cumulative distribution function of an increment we show the following generalization of Lemma 5.

Lemma 6 *If the repair process is a minimal repair process and the cumulative cost process is given by relation (3.2.24) then for every* $t > u > 0$

$$\mathbb{E}(e^{-s(\mathbf{C}_1(t)-\mathbf{C}_1(u))}) = e^{-((\overline{Q}_t(0)-\overline{Q}_u(0))-(\overline{Q}_t(s)-\overline{Q}_u(0)))} \tag{3.2.29}$$

with

$$\overline{Q}_t(s) = \int_0^t r(y)e^{-sc(y)}dy. \tag{3.2.30}$$

Proof. Since by Lemma 5 the cost process has independent increments we obtain that

$$\mathbb{E}(e^{-s\mathbf{C}_1(t)}) = \mathbb{E}(e^{-s(\mathbf{C}_1(u)+\mathbf{C}_1(t)-\mathbf{C}_1(u))}) = \mathbb{E}(e^{-s\mathbf{C}_1(u)})\mathbb{E}(e^{-s(\mathbf{C}_1(t)-\mathbf{C}_1(u))})$$

Applying now again Lemma 5 yields the desired result. □

Since the cost process $\mathbf{C}_1$ has independent increments it follows using standard techniques (see [2]) and lemma 6 that the stochastic process $\mathbf{M} = \{\mathbf{M}_t : t \geq 0\}$ given by

$$\mathbf{M}_t = e^{-s\mathbf{C}_1(t)}e^{\overline{Q}_t(0)-\overline{Q}_t(s)}$$

is a martingale. This observation might be useful for the case that the usage time is a stopping time with respect to the failure process. This might also be a topic of future research.

Again we mention as a corollary (take $c(y) = 1$ for every $y \geq 0$) the following important special case of Lemma 5 in which we only deal with the number of returned and defective items. It gives a formula for the probability-Laplace Stieltjes transform of the number of defects of a particular sold item for a minimal repair process. Although we already are familiar with this result it is listed for completeness.

Corollary 6 *If the repair process of items is a minimal repair process then it follows that its stochastic counting process* N_1 *is a non-homogeneous Poisson process and*

$$\mathbb{E}(e^{-s\mathbf{N}_1(t)}) = e^{-\Phi(t)(1-e^{-s})} \tag{3.2.31}$$

with Φ *the intensity measure of the counting process* $\mathbf{N}_1$.

Again we can now derive a more detailed expression for the probability Laplace-Stieltjes transform of the random variable $\mathbf{R}(t)$ using Corollary 5 and relation (3.2.22) for $\mathbf{U}_1$ independent of the failure process. In particular applying relation (3.2.22) we obtain for a minimal repair process that

$$\mathbb{E}(e^{-s\mathbf{N}_1(t\wedge\mathbf{U}_1)}) = e^{-\Phi(t)(1-e^{-s})}(1 - G(t) + \int_0^t e^{-\Phi(u)(1-e^{-s})}G(du) \tag{3.2.32}$$

with Φ the intensity measure of the failure counting process $\mathbf{N}_1$. For the special case $\mathbf{U}_1 \overset{a.s}{=} w$ it follows that

$$\mathbb{E}(e^{-s\mathbf{N}_1(t\wedge\mathbf{U}_1)}) = e^{-\Phi(t\wedge w)(1-e^{-s})}. \tag{3.2.33}$$

It is also possible using standard techniques and the special form of the probability Laplace-Stieltjes transform of the random variable $\mathbf{C}(t)$ given in Theorem 1 to derive a recurrent relation for the nth moment of the cumulative cost process at time t relating it to the kth moment, $1 \leq k \leq n$, of the cumulative cost process of each individual item.

Lemma 7 *If the functions* $c_n : [0,\infty) \to [0,\infty], n \in \mathbb{Z}_+$ *are given by*

$$c_0(t) := 1 \textit{ and } c_n(t) := \mathbb{E}(\mathbf{C}^n(t)), n \in \mathbb{N}$$

and $\bar{c}_n : [0,\infty) \to [0,\infty], n \in \mathbb{N}$ *by*

$$\bar{c}_n(t) := \int_0^t \psi(y)\mathbb{E}(\mathbf{C}_1^n((t-y)\wedge\mathbf{U}_1))dy$$

then for every $n \in \mathbb{N}$ *and* $t \geq 0$

$$c_n(t) = \sum\nolimits_{k=0}^{n-1} \bar{c}_{n-k}(t)c_k(t). \tag{3.2.34}$$

Proof. We know from Theorem 1 that for every $s > 0$

$$\pi_{\mathbf{C}(t)}(s) = \mathbb{E}(e^{-s\mathbf{C}(t)}) = e^{Q_t(s)-Q_t(0)}$$

with $Q_t(s) = \int_0^t \psi(y)\mathbb{E}(e^{-s\mathbf{C}_1((t-y)\wedge \mathbf{U}_1)})dy$. This shows that

$$\pi^{(1)}_{\mathbf{C}(t)}(s) = Q^{(1)}_t(s)e^{Q_t(s)-Q_t(0)} = Q^{(1)}_t(s)\pi_{\mathbf{C}(t)}(s). \tag{3.2.35}$$

Applying Lemma 18 in the Appendix A and Leibniz rule for the product of functions to relation (3.2.35) it follows for every $n \in \mathbb{N}$ and $s > 0$ that

$$\pi^{(n)}_{\mathbf{C}(t)}(s) = \sum_{k=0}^{n-1} \binom{n-1}{k} Q^{(n-k)}_t(s)\pi^{(k)}_{\mathbf{C}(t)}(s).$$

Letting $s \downarrow 0$ this yields

$$\pi^{(n)}_{\mathbf{C}(t)}(0^+) = \sum_{k=0}^{n-1} \binom{n-1}{k} Q^{(n-k)}_t(0^+)\pi^{(k)}_{\mathbf{C}(t)}(0^+)$$

and so

$$(-1)^n\pi^{(n)}_{\mathbf{C}(t)}(0^+) = \sum_{k=0}^{n-1} \binom{n-1}{k} (-1)^{n-k}Q^{(n-k)}_t(0^+)(-1)^k\pi^{(k)}_{\mathbf{C}(t)}(0^+).$$

Again by Lemma 18 and relation (3.2.11) we know that

$$(-1)^n Q^{(n)}_t(0^+) = \int_0^t \psi(y)\mathbb{E}(\mathbf{C}^n_1((t-y)\wedge \mathbf{U}_1))dy$$

and

$$(-1)^k\pi^{(k)}_{\mathbf{C}(t)}(0^+) = \mathbb{E}(\mathbf{C}^k_1(t))$$

showing the desired result. □

For $n = 1$ it follows

$$\pi^{(0)}_{\mathbf{C}(t)}(0^+) = \pi_{\mathbf{C}(t)}(0) = 1$$

and so we obtain by Lemma 7

$$\mathbb{E}(\mathbf{C}(t)) = \int_0^t \psi(y)\mathbb{E}(\mathbf{C}_1((t-y)\wedge \mathbf{U}_1))dy$$

rediscovering the result in relation (3.2.2).

To compute the nth moment of the cumulative cost process we need to compute beforehand the functions $\bar{c}_k(t)$ for every $1 \leq k \leq n-1$. In case the usage time $\mathbf{U}_1$ is independent of the cost process $\mathbf{C}_1$ it follows for every $y \geq 0$ that

$$\begin{aligned}
\mathbb{E}(\mathbf{C}_1^k(y \wedge \mathbf{U}_1)) &= \mathbb{E}(\mathbf{C}_1^k(y \wedge \mathbf{U}_1)1_{\{\mathbf{U}_1 \leq y\}}) + \mathbb{E}(\mathbf{C}_1^k(y \wedge \mathbf{U}_1)1_{\{\mathbf{U}_1 > y\}}) \\
&= \mathbb{E}(\mathbf{C}_1^k(\mathbf{U}_1)1_{\{\mathbf{U}_1 \leq y\}}) + \mathbb{E}(\mathbf{C}_1^k(y)1_{\{\mathbf{U}_1 > y\}}) \qquad (3.2.36) \\
&= \int_0^y \mathbb{E}(\mathbf{C}_1^k(u))dG(u) + \mathbb{E}(\mathbf{C}_1^k(y))(1 - G(y))
\end{aligned}$$

To compute the kth moments one may use the following result.

Lemma 8 *Let the repair process be a minimal repair process and the cumulative cost process be given by relation (3.2.24). If the functions* $c_n : [0, \infty) \mapsto [0, \infty)$*,* $n \in \mathbb{Z}_+$ *are given by*

$$c_0(t) := 1 \textit{ and } c_n(t) := \mathbb{E}(\mathbf{C}_1^n(t))$$

and $\overline{c}_n : [0, \infty) \mapsto [0, \infty), n \in \mathbb{N}$ *by*

$$\overline{c}_n(t) := \int_0^t r(y)c^n(y))dy$$

then for every $n \in \mathbb{N}$ *and* $t \geq 0$

$$c_n(t) = \sum_{k=0}^{n-1} \overline{c}_{n-k}(t)c_k(t). \qquad (3.2.37)$$

Proof. By Lemma 5 we know that $\pi^{(1)}_{\mathbf{C}_1(t)}(s) = \overline{Q}_t^{(1)}(s)\ \pi_{\mathbf{C}_1(t)}(s)$ with $\overline{Q}_t(s) = \int_0^t r(y)e^{-sc(y)}dy$. Applying now the same arguments as the proof of Lemma 7 the desired result follows. □

Hence under the minimal repair assumption and using the cost model of relation (3.2.24) it seems to be in principle possible to compute using numerical techniques these unknown constants. This approach will not be pursued in this thesis. If we are only interested in the nth moment of the cumulative repair process $\mathbf{R}$ up to time t we obtain the following special case of Lemma 7.

Corollary 7 *If the functions* $c_n : [0, \infty) \to [0, \infty]$ *are given by*

$$c_0(t) := 1 \textit{ and } c_n(t) := \mathbb{E}(\mathbf{R}^n(t)), n \in \mathbb{N}$$

and $\overline{c}_n : [0, \infty) \to [0, \infty], n \in \mathbb{N}$ *by*

$$\overline{c}_n(t) := \int_0^t \psi(y)\mathbb{E}(\mathbf{N}_1^n((t - y) \wedge \mathbf{U}_1))dy$$

then for every $n \in \mathbb{N}$ and $t \geq 0$

$$c_n(t) = \sum\nolimits_{k=0}^{n-1} \bar{c}_{n-k}(t) c_k(t). \tag{3.2.38}$$

In a statistical analysis it is important to derive a formula for the variance to measure the variability of the random variable. An immediate consequence of Lemma 7 for the second moment $c_2(t)$ is the following expression for the variance of the random variable $\mathbf{C}(t)$.

Lemma 9 *It follows for any $t \geq 0$ that*

$$Var(\mathbf{C}(t)) = \int_0^t \psi(y) \mathbb{E}(\mathbf{C}_1^2((t-y) \wedge \mathbf{U}_1)) dy. \tag{3.2.39}$$

Proof. By Lemma 7 we obtain for $n = 2$ that

$$\mathbb{E}(\mathbf{C}^2(t)) = c_2(t) = \bar{c}_2(t) c_0(t) + \bar{c}_1(t) c_1(t) = \bar{c}_2(t) + \bar{c}_1(t) c_1(t). \tag{3.2.40}$$

Again by Lemma 7 $c_1(t) = \bar{c}_1(t)$ and so we obtain

$$\mathbb{E}(\mathbf{C}^2(t)) = c_2(t) = \bar{c}_2(t) + c_1^2(t). \tag{3.2.41}$$

This shows

$$Var(\mathbf{C}(t)) = c_2(t) - c_1^2(t) = \bar{c}_2(t) = \int_0^t \psi(y) \mathbb{E}(\mathbf{C}_1^2((t-y) \wedge \mathbf{U}_1) dy \tag{3.2.42}$$

and we have verified relation (3.2.39). □

To evaluate the variance of $\mathbf{C}(t)$ given in relation (3.2.39) we need to know $\mathbb{E}(\mathbf{C}_1^2(t \wedge \mathbf{U}_1))$ for every $t \geq 0$. To compute this for a minimal repair process and a cumulative costs process satisfying relation (3.2.24) we first mention the following result.

Lemma 10 *It follows for any minimal repair process and a cumulative costs process $\mathbf{C}_1$ satisfying relation (3.2.24) that*

$$Var(\mathbf{C}_1(t)) = \int_0^t r(y) c^2(y) dy. \tag{3.2.43}$$

Proof. Apply the same proof as in Lemma 9 and use Lemma 8. □

In particular under the minimal repair assumption we obtain using Lemma 10 that

$$\mathbb{E}(\mathbf{C}_1^2(t)) = \int_0^t r(y) c^2(y) dy + \left(\int_0^t r(y) c(y) dy \right)^2. \tag{3.2.44}$$

Assuming that the random usage time $\mathbf{U}_1$ is independent of the individual item cumulative cost process $\mathbf{C}_1$ it follows by relation (3.2.36) for $k = 2$ that

$$\mathbb{E}(\mathbf{C}_1^2(t \wedge \mathbf{U}_1)) = \int_0^t \mathbb{E}(\mathbf{C}_1^2(u))dG(u) + \mathbb{E}(\mathbf{C}_1^2(t))(1 - G(t)).$$

with $\mathbb{E}(\mathbf{C}_1^2(t)))$ listed in relation (3.2.44). Applying now Lemma 9 we can write down an expression for the variance of $\mathbf{C}_1(t)$ which can be numerically evaluated. Clearly computing the variance is important since it gives a measure of the variability of the random variable $\mathbf{R}(t)$.

Before discussing the challenging task to compute or approximate $\mathbb{E}(\mathbf{C}_1(t \wedge \mathbf{U}_1))$ for any stopping time $\mathbf{U}_1$ we observe that under the minimal repair assumption the cost process $\mathbf{C}_1$ given by relation (3.2.24) has independent increments. It is well-known that for any stochastic process $\mathbf{X}$ with independent increments and satisfying $\mathbb{E}(\mathbf{X}(t)) = 0$ that the stochastic process $\mathbf{M} = \{\mathbf{M}_t : t \geq 0\}$ given by

$$\mathbf{M}(t) = \mathbf{X}^2(t) - \mathbb{E}\mathbf{X}^2(t)$$

is a martingale (see [11]). Applying this result to the cumulative cost process $\mathbf{C}_1$ we observe by Lemma 10 that the stochastic process

$$\begin{aligned} \mathbf{M}(t) &= \left(\mathbf{C}_1(t) - \textstyle\int_0^t r(u)c(u)du\right)^2 - Var(\mathbf{C}_1(t)) \\ &= \left(\mathbf{C}_1(t) - \textstyle\int_0^t r(u)c(u)du\right)^2 - \textstyle\int_0^t r(u)c^2(u)du \end{aligned}$$

is a martingale. Applying Doobs stopping theorem to this martingale we obtain for a usage time $\mathbf{U}_1$ being a stopping time with respect to the cost process that

$$\mathbb{E}\mathbf{M}(t \wedge \mathbf{U}_1) = 0$$

Hence we obtain a relation which might be useful in approximating or computing $\mathbb{E}(\mathbf{C}_1^2(t \wedge \mathbf{U}_1))$. This is a topic of future research.

If we are only interested in the cumulative repair process we obtain the following special case of Lemma 9.

Corollary 8 *It follows for any $t \geq 0$ that*

$$Var(\mathbf{R}(t)) = \int_0^t \psi(y)\mathbb{E}(\mathbf{N}_1^2((t-y) \wedge \mathbf{U}_1))dy. \tag{3.2.45}$$

To compute for the cumulative repair process the integral

$$\int_0^t \psi(y)\mathbb{E}(\mathbf{N}_1^2((t-y)\wedge \mathbf{U}_1))dy$$

we need to know the second moment of the counting process of repairs. A special case of Corollary 8 is given by the next result. Actually this result is of importance in the statistical section. It yields a measure of the variability of the random variable $\mathbf{R}(t)$.

Corollary 9 *If the repair process of items is a minimal repair process and the usage time* $\mathbf{U}_1$ *equals* $\mathbf{U}_1 \overset{a.s}{=} w$ *with* w *denoting the length of the warranty period then for any* $t \geq 0$

$$\begin{aligned} & Var(\mathbf{R}(t)) \\ = \; & -\textstyle\int_0^t \psi(t-y)\ln(1-F(y\wedge w))dy - \int_0^t \psi(t-y)(\ln(1-F(y\wedge w)))^2 dy \end{aligned} \tag{3.2.46}$$

with F *denoting the cumulatiive distribution function of the random time to the first failure.*

Proof. If the repair process is of the so-called minimal repair type we know by Lemma 6 that the counting process $\mathbf{N}_1$ is a non-homogeneous Poisson process with arrival intensity function given by the failure rate function

$$r(t) = \frac{f(t)}{1-F(t)}$$

of the cdf F of the random time to the first failure. This implies for $\mathbf{U}_1 = w$ a.s that

$$\mathbb{E}(\mathbf{N}_1^2(s\wedge \mathbf{U}_1)) = \mathbb{E}(\mathbf{N}_1^2(s\wedge w)) = Var(\mathbf{N}_1^2(s\wedge w)) + \mathbb{E}(\mathbf{N}_1(s\wedge w))^2.$$

Since for a Poisson distributed random variable we know that

$$Var(\mathbf{N}_1(s\wedge w)) = \mathbb{E}(\mathbf{N}_1(s\wedge w))$$

this shows that

$$\mathbb{E}(\mathbf{N}_1^2(s\wedge \mathbf{U}_1)) = \mathbb{E}(\mathbf{N}_1(s\wedge w)) + \mathbb{E}(\mathbf{N}_1(s\wedge w))^2$$

and hence we obtain

$$\mathbb{E}(\mathbf{N}_1^2(s\wedge \mathbf{U}_1)) = \Phi(s\wedge w) + \Phi^2(s\wedge w)$$

with

$$\Phi(s) := \int_0^s r(s)ds.$$

the intensity measure of the cumulative repair process $\mathbf{N}_1$. This shows for $\mathbf{U}_1 = w$ that by Corollary 8

$$Var(\mathbf{R}(t)) = \int_0^t \psi(t-y)\Phi(y \wedge w)dy + \int_0^t \psi(t-y)\Phi^2(y \wedge w)dy. \qquad (3.2.47)$$

Applying relation (3.1.12) yields the desired result. □

In general we need numerical integration techniques (see [46] or [16]) to compute the integrals in relation (3.2.46).

In case we are only interested in the cumulative stochastic process $\mathbf{R}$ of returned and defective items it is also possible to derive a recursive equation for

$$p_k(t) := \mathbb{P}(\mathbf{R}(t) = k), k \in \mathbb{Z}_+ \qquad (3.2.48)$$

A way to derive these recursive equations we first observe for every $t \geq 0$ that (see Section 1.2.2 of [48])

$$p_k(t) = \frac{P^{(k)}_{\mathbf{R}(t)}(0)}{k!}, k \in \mathbb{Z}_+ \qquad (3.2.49)$$

with $P^{(k)}_{\mathbf{R}(t)}(0)$ denoting the kth derivative of the pgf $P_{\mathbf{R}(t)}(z)$ evaluated in $z = 0$. Using this result and relation (3.2.18) it is relatively easy to show the following recurrent relation for the pdf of the random variable $\mathbf{R}(t)$.This result is actually know as Adelsons recursion formula (see [47]). Before discussing this result we introduce the functions $\varphi_k : \mathbb{R}_+ \to \mathbb{R}_+$, $k \in \mathbb{Z}_+$ given by

$$\varphi_k(t) := \int_0^t \psi(y)\mathbb{P}(\mathbf{N}_1((t-y) \wedge \mathbf{U}_1) = k)dy \qquad (3.2.50)$$

Lemma 11 *It follows for every $t \geq 0$ that*

$$p_0(t) = e^{-\Psi(t)} \qquad (3.2.51)$$

while for $n \in \mathbb{N}$

$$np_n(t) = \sum\nolimits_{k=0}^{n-1}(n-k)p_k(t)\varphi_{n-k}(t) \qquad (3.2.52)$$

Proof. Clearly

$$p_0(t) = \mathbb{P}(\mathbf{S}(t) = 0) = e^{-\Lambda(t)}.$$

By relation (3.2.18) we know

$$P_{\mathbf{R}(t)}(z) = e^{Q_t(z)-Q_t(1)}$$

This shows

$$P^{(1)}_{\mathbf{R}(t)}(z) = Q^{(1)}_t(z)P_{\mathbf{R}(t)}(z).$$

Hence by Leibniz rule for the product of two differentiable functions we obtain for every $n \in N$ that

$$P^{(n)}_{\mathbf{R}(t)}(z) = \sum_{k=0}^{n-1} \binom{n-1}{k} Q^{(n-k)}_t(z)P^{(k)}_{\mathbf{R}(t)}(z)$$

This shows applying relation (3.2.49) that

$$\begin{aligned} p_n(t) &= \frac{P^{(n)}_{\mathbf{R}(t)}(0)}{n!} \\ &= \frac{1}{n!}\sum_{k=0}^{n-1}\binom{n-1}{k}Q^{(n-k)}_t(0)P^{(k)}_{\mathbf{R}(t)}(0) \\ &= \frac{1}{n}\sum_{k=0}^{n-1}(n-k)\frac{Q^{(n-k)}_t(0)}{(n-k)!}\frac{P^{(k)}_{\mathbf{R}(t)}(0)}{k!} \\ &= \frac{1}{n}\sum_{k=0}^{n-1}(n-k)\frac{Q^{(n-k)}_t(0)}{(n-k)!}p_k(t). \end{aligned} \tag{3.2.53}$$

Since it is easy to verify for every k that

$$Q^{(k)}_t(z) = \int_0^t \psi(y)P^{(k)}_{\mathbf{N}_1((t-y)\wedge\mathbf{U}_1))}(z)dy$$

it follows that

$$\frac{Q^{(k)}_t(0)}{k!} = -\int_0^t \psi(y)\mathbb{P}(\mathbf{N}_1((t-y)\wedge\mathbf{U}_1) = k)dy.$$

Applying now relation (3.2.53) we obtain the desired result. □

Using the above lemma after having estimated both the arrival intensity function of the non-homogeneous Poisson cumulative sales process (see [30]) and the probability law of the counting process of repairs we can compute recursively the cdf of the random variable $\mathbf{R}(t)$. It is well known that this recursive procedure (see [48]) is in general stable. To compute the integral

$$\int_0^t \psi(y)\mathbb{P}(\mathbf{N}_1((t-y)\wedge\mathbf{U}_1) = k)dy$$

we need to know the cdf of the counting process of repairs.

If the repair process is of the so-called minimal repair type (this repair type applies to a lot of consumer goods consisting of a lot of different components) we know that the random counting process $\mathbf{N}_1$ is a non=homogeneous Poisson process with arrival intensity function given by the failure rate function $r(t) = \frac{f(t)}{1-F(t)}$ of the cdf F of the random time to first failure. This shows that

$$\mathbb{P}(\mathbf{N}_1(s) = k) = e^{-\Lambda_1(s)} \frac{\Phi(s)^k}{k!}, j \in \mathbb{Z}_+$$

If we consider the simplest case $\mathbf{U}_1 \stackrel{a.s}{=} w$ this shows that

$$\begin{aligned}\int_0^t \psi(y)\mathbb{P}(\mathbf{N}_1((t-y) \wedge \mathbf{U}_1) = k)dy &= \int_0^t \psi(t-y)\mathbb{P}(\mathbf{N}_1(y \wedge w) = k)dy \\ &= \int_0^t \psi(t-y)e^{-\Phi(y\wedge w)} \frac{\Phi(y\wedge w)^k}{k!}dy.\end{aligned}$$

We now introduce the next definition.

Definition 2 *For any* $\mathbf{s} = (s_1, s_2) \in \mathbb{R}_+^2$ *the function* $\pi_{\mathbf{X}} : \mathbb{R}_+^2 \to [0,1]$ *given by*

$$\pi_{\mathbf{X}}(\mathbf{s}) := \mathbb{E}(e^{-s_1\mathbf{X}_1 - s_2\mathbf{X}_2}) = \mathbb{E}(e^{-\mathbf{s}^\intercal \mathbf{X}}). \tag{3.2.54}$$

is called the bivariate probability Laplace-Stieltjes transform (bpLSt) of the non-negative random vector $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2)$. *If additionally the random vector* $(\mathbf{X}_1, \mathbf{X}_2)$ *is integer valued the function* $P_{\mathbf{X}} : \mathbb{D}^2 \to \mathbb{C}$ *given by*

$$P_{\mathbf{X}}(z_1, z_2) := \mathbb{E}\left(z_1^{\mathbf{X}_1} z_2^{\mathbf{X}_2}\right). \tag{3.2.55}$$

is called the bivariate probability generating function of the random vector $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2)$

In general we are also interested in the dependence of the cumulative cost process $\mathbf{C}$ over time given by for example a covariance function and or any finite dimensional cdf of this stochastic process. To derive some of these expressions we first need to introduce useful notation. If $\mathbf{C}_n(t \wedge \mathbf{U}_n), t \geq 0$ is the cumulative cost of repair process for item n introduce the following function

$$\overline{\mathbf{C}}_n(t \wedge \mathbf{U}_n) = \begin{cases} \mathbf{C}_n(t \wedge \mathbf{U}_n) & \text{if } t \geq 0 \\ 0 & \text{if } t < 0 \end{cases} \tag{3.2.56}$$

For $t_2 \geq t_1$ it follows by the above convention that

$$\mathbf{C}(t_1) = \sum_{n=1}^{\infty} \overline{\mathbf{C}}_n((t_1 - \mathbf{T}_n) \wedge \mathbf{U}_n) 1_{\{\mathbf{T}_n \leq t_2\}} \tag{3.2.57}$$

One can now derive using a similar proof as done in Lemma 4 the following generalization of this lemma.

Lemma 12 *If the cumulative sales process* $\mathbf{S} = \{\mathbf{S}(t) : t \geq 0\}$ *is a non-homogeneous Poisson process with Borel arrival intensity function* ψ *independent of the independent and identically distributed cumulative cost processes* $\mathbf{C}_n, n \in \mathbb{N}$ *listed in relation (3.1.14) then for every* $t_2 \geq t_1$ *and* $k \in \mathbb{Z}_+$ *it follows that*

$$\mathbb{E}(e^{-s_1\mathbf{C}(t_1)-s_2\mathbf{C}(t_2)} \mid \mathbf{S}(t_2) = k) = \mathbb{E}(e^{-s_1\overline{\mathbf{C}}_1((t_1-\mathbf{Y}_1)\wedge\mathbf{U}_1)-s_2\overline{\mathbf{C}}_1((t_2-\mathbf{Y}_1)\wedge\mathbf{U}_1)})^k$$

with the random variable $\mathbf{Y}_1$ *concentrated on* $[0, t_2]$ *having cdf*

$$F(y) = \frac{\Psi(y)}{\Psi(t_2)}, 0 \leq y \leq t_2. \tag{3.2.58}$$

and being independent of the stochastic process $\{\mathbf{C}_1(s \wedge \mathbf{U}_1) : s \geq 0\}$

Proof. By relation (3.2.57) we know for $t_2 \geq t_1$ that

$$\mathbf{C}(t_1) = \sum\nolimits_{n=1}^{\infty} \overline{\mathbf{C}}_n((t_1 - \mathbf{T}_n) \wedge \mathbf{U}_n) 1_{\{\mathbf{T}_n \leq t_2\}}.$$

Introducing for the selected $s_1, s_2 \geq 0$ the random variable

$$\mathbf{Q}_n(t_1, t_2) = -s_1\overline{\mathbf{C}}_n(t_1 \wedge \mathbf{U}_n) - s_2\overline{\mathbf{C}}_n(t_2 \wedge \mathbf{U}_n)$$

it follows that

$$\mathbb{E}(e^{-s_1\mathbf{C}(t_1)-s_2\mathbf{C}(t_2)} \mid \mathbf{S}(t_2) = k) = \mathbb{E}\left(e^{\sum_{n=1}^{k} \mathbf{Q}_n(t_1-\mathbf{Y}_{n;k}, t_2-\mathbf{Y}_{n;k})}\right)$$

with $(\mathbf{Y}_{1:k}, ..., \mathbf{Y}_{k:k})$ the joint order statistics of a sequence of independent and identically distributed random variables $\mathbf{Y}_n, n = 1, ..., k$ on $[0, t]$ concentrated on $[0, t_2]$ and having the continuous cumulative distribution function

$$F(y) = \mathbb{P}(\mathbf{Y} \leq y) = \frac{\Psi(y)}{\Psi(t_2)}, 0 \leq y \leq t_2 \tag{3.2.59}$$

Since the random variables $\mathbf{Q}_n$ are independent of the arrival process we obtain by the same proof as used in Lemma 4 that

$$\mathbb{E}\left(e^{\sum_{n=1}^{k} \mathbf{Q}_n(t_1-\mathbf{Y}_{n;k}, t_2-\mathbf{Y}_{n;k})}\right) = \mathbb{E}\left(e^{\mathbf{Q}_1(t_1-\mathbf{Y}_1, t_2-\mathbf{Y}_1)}\right)^k.$$

This shows

$$\begin{aligned}\mathbb{E}(e^{-s_1\mathbf{C}(t_1)-s_2\mathbf{C}(t_2)} \mid \mathbf{S}(t_2)=k) &= \mathbb{E}\left(e^{\mathbf{Q}_1(t_1-\mathbf{Y}_1,t_2-\mathbf{Y}_1)}\right)^k \\ &= \mathbb{E}(e^{-s_1\overline{\mathbf{C}}_1((t_1-\mathbf{Y}_1)\wedge\mathbf{U}_1)-s_2\overline{\mathbf{C}}_1((t_2-\mathbf{Y}_1)\wedge\mathbf{U}_1)})^k\end{aligned}$$

and we have verified the desired result □

Before discussing the next result we introduce for $t_2 \geq t_1$ and $\mathbf{t} = (t_1, t_2)$ and $\mathbf{s} = (s_1, s_2) \in \mathbb{R}^2_+$ the function $Q_{\mathbf{t}} : \mathbb{R}^2_+ \to \mathbb{R}_+$ given by

$$Q_{\mathbf{t}}(\mathbf{s}) = \int_0^{t_2} \psi(y)\mathbb{E}(e^{-s_1\overline{\mathbf{C}}_1((t_1-y)\wedge\mathbf{U}_1)-s_2\overline{\mathbf{C}}_1((t_2-y)\wedge\mathbf{U}_1)})dy. \tag{3.2.60}$$

Clearly $Q_{\mathbf{t}}(\mathbf{0}) = \Psi(t_2)$. It is now possible to show the following result for the bivariate probability Laplace-Stieltjes transform of the random vector $(\mathbf{C}(t_1), \mathbf{C}(t_2))$ using a similar approach as done in Theorem 1.

Theorem 2 *If the cumulative sales process* $\mathrm{S} = \{\mathbf{S}(t) : t \geq 0\}$ *is a non-homogeneous Poisson process with Borel intensity function* ψ *and this cumulative sales process is independent of the independent and identically distributed cumulative costs processes* $\mathbf{C}_n, n \in \mathbb{N}$ *then for every* $t_2 \geq t_1$ *and* $\mathbf{t} = (t_1, t_2)$ *and* $\mathbf{s} = (s_1, s_2) \in \mathbb{R}^2_+$

$$\mathbb{E}(e^{-s_1\mathbf{C}(t_1)-s_2\mathbf{C}(t_2)}) = e^{-(Q_{\mathbf{t}}(\mathbf{0})-Q_{\mathbf{t}}(\mathbf{s}))}$$

with $Q_{\mathbf{t}}(\mathbf{s})$ *given by relation (3.2.60).*

Looking in detail at the proof of Theorem 2 one can actually conclude the following. Introducing for any finite sequence $0 \leq t_1 \leq t_2 \leq t_3 \leq ...t_n$ and $\mathbf{t} = (t_1, ..., t_n)$ and $\mathbf{s} = (s_1, ...s_n) \in \mathbb{R}^n_+$ the multivariate probability Laplace-Stieltjes transform $\pi_{\mathbf{C}(\mathbf{t})} : \mathbb{R}^n_+ \to [0,1]$ given by

$$\pi_{\mathbf{C}(\mathbf{t})}(\mathbf{s}) := \mathbb{E}(e^{-\mathbf{s}^{\intercal}\mathbf{C}(\mathbf{t})}) \tag{3.2.61}$$

with $\mathbf{C}(\mathbf{t}) = (\mathbf{C}(t_1), ...\mathbf{C}(t_n))$ and introducing the function $Q_{\mathbf{t}}(\mathbf{s}) : \mathbb{R}^n_+ \mapsto [0,1]$ given by

$$Q_{\mathbf{t}}(\mathbf{s}) = \int_0^{t_n} \psi(y)\mathbb{E}\left(e^{-\sum_{\imath=1}^n s_\imath\overline{\mathbf{C}}_1((t_\imath-y)\wedge\mathbf{U}_1)}\right)dy \tag{3.2.62}$$

one can show by the same approach as done in Theorem 2 for the bivariate case the following theorem.

Theorem 3 *If the cumulative sales process* $\mathrm{S} = \{\mathbf{S}(t) : t \geq 0\}$ *is a non-homogeneous Poisson process with Borel intensity function* ψ *and this cumulative sales process is independent of the independent and identi cally distributed* *cumulative costs processes* $\mathbf{C}_n, n \in \mathbb{N}$ *then for every vector* $\mathbf{t} = (t_1, ...t_n)$ *satisfying* $t_1 \leq .... \leq t_n$ *and* $\mathbf{s} \in \mathbb{R}_+^n$

$$\mathbb{E}(e^{-\mathbf{s}^\intercal \mathbf{C}(\mathbf{t})}) = e^{-(Q_\mathbf{t}(\mathbf{0}) - Q_\mathbf{t}(\mathbf{s}))}. \tag{3.2.63}$$

Using Theorem 2 and for $\mathbf{t} = (t_1, t_2)$ with $t_1 \leq t_2$

$$\frac{\partial^2 \pi_{\mathbf{R}(t)}}{\partial s_1 \partial s_2}(\mathbf{0}) = -\,\mathbb{E}(\mathbf{C}(t_1)\mathbf{C}(t_2)) \tag{3.2.64}$$

we obtain the following expression for the covariance function.

Lemma 13 *It follows for* $t_2 \geq t_1$ *that*

$$Cov(\mathbf{C}(t_1), \mathbf{C}(t_2)) = \int_0^{t_1} \psi(y)\mathbb{E}(\mathbf{C}_1((t_1 - y) \wedge \mathbf{U}_1)\mathbf{C}_1((t_2 - y) \wedge \mathbf{U}_1))dy.$$

Proof. Apply Theorem 2 and use relation (3.2.64). □

This concludes the discussion of the model in this thesis. It is still to be investigated whether the stochastic process $\mathbf{R}$ has more useful properties under the minimal repair assumption. In the second part of this thesis we discuss the statistical techniques in the estimation of the parametric models describing the failure and the sales process.

Chapter 4

ESTIMATING THE INTENSITY MEASURE OF THE SALES PROCESS

The prevailing approach in predicting the future sales of a product has been time series forecasting. In this study we propose to use a parametric function to estimate the sales quantities, which is a new approach independent of the model proposed for returned items. In this chapter we present the simplest possible parametric models to estimate the intensity measure of the sales process of our two sample products. First, a simple polynomial regression model is analyzed. Then, another method derived from [12] is tested and the results of the two methods are compared. In addition, we also conduct a fit with weekly data which smooths fluctuations in daily sales quantities.

4.1 Sales data

We begin by presenting the characteristics of our data. We have the item-level data on the sales and failure dates of two products (Product A and Product B) of a household durable goods manufacturer. To be precise, each "product" here refers to a product group, including multiple models and SKUs.

The data covers the five consecutive years 2013,...,2017. Each row in the database represents an individual item of the product and the number of data in each years data set is around one million items. Columns identify the bill number, the item's production code, the year and month of production, serial number and the date of installation (Figure 4.1). The installation date denotes the day on which the warranty starts. We assume that the installation date also represents the sales date of the product. In practice, there might be a couple of days' difference between the sale of the product and its installation at the home of a customer, but this should have only a negligible impact on our results.

	A	B	C	D	E	F
1	BillNo	ProductionCode	ProductionYear	ProductionMonth	SerieNo	WarrantyStart
2	253843286	70076	2012	January	100037	1/1/2013
3	258204975	70274	2011	March	100058	1/1/2013
4	253840466	70195	2012	January	100065	1/1/2013
5	253839061	70358	2012	January	100071	1/1/2013

Figure 4.1: Sales data in raw form

In Figures 4.2 and 4.3, the daily sales quantities in five years for the two products are depicted. For both products, we observe the same repeating annual pattern. Sales start low around January, and keep increasing towards the summer months. After reaching a peak around September, sales drop again in the subsequent months. This peak around September is caused by regular promotional activity. We do not observe a significant upward or downward trend over years. Since the sales refer to a product group containing multiple models and SKUs (such as all refrigerator SKUs of the firm), we do not observe a standard life cycle shape starting from 2013. Instead, the dominant factor in the sales is the seasonality that repeats itself annually.

We also observe near-zero sales values on Sundays. Recall that our sales data is actually installation data, and on Sundays the installation crews do not operate in most cities.

Since the plots in different years follow the same pattern, and there is no trend over the years, one may assume that this pattern will be repeated in the future. Therefore, we use the average of the five years' daily sales in the estimation of the parameters of the sales model. These averages are depicted in Figures 4.4 and 4.5.

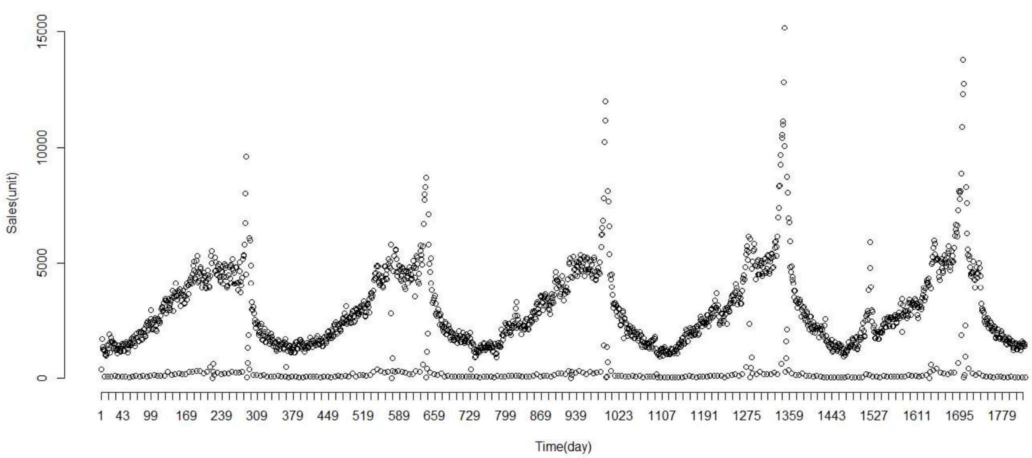


Figure 4.2: Daily sales data of product A (2013-2017)

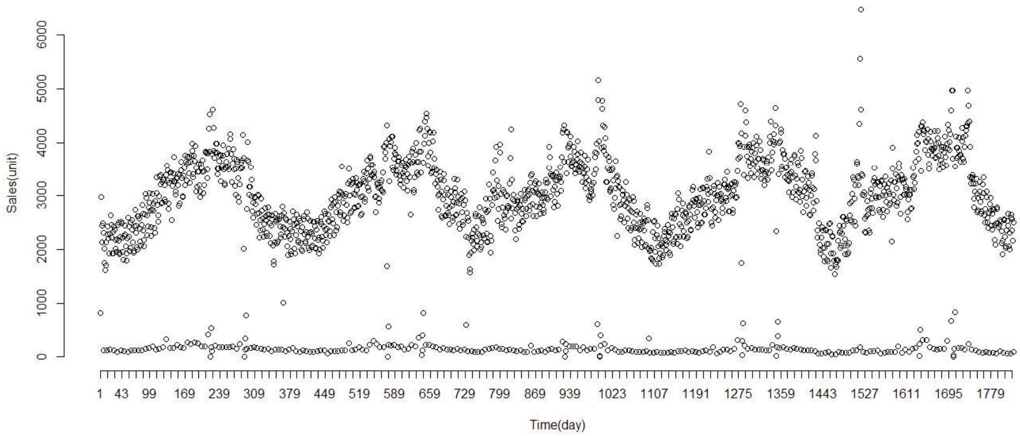


Figure 4.3: Daily sales data of product B (2013-2017)

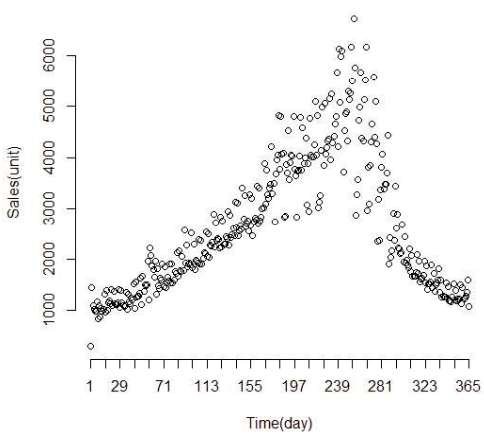


Figure 4.4: Average daily sales data of product A

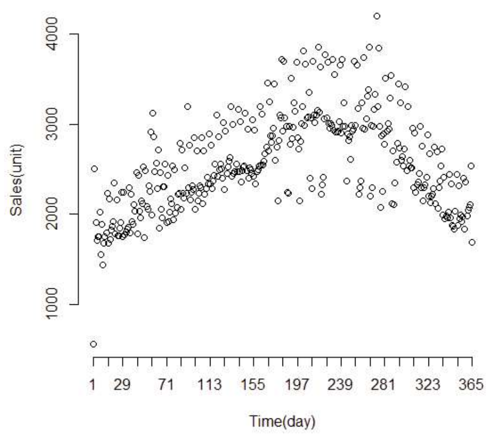


Figure 4.5: Average daily sales data of product B

4.2 Methodology for estimating sales

Traditionally, sales forecasting is based on simple extrapolation of historical data. Different from this approach, we assume that the sales process is represented by some stochastic process with an unknown probability law. Since our data is available in daily granularity we use a discrete time stochastic process model. The accumulative non-negative sales process is given by $\mathbf{S} = \{\mathbf{S}(t) : t \in \mathbb{N}\}$ and $\Delta\mathbf{S}(t) = \mathbf{S}(t) - \mathbf{S}(t-1), t \in \mathbb{N}$, with $\mathbf{S}(0) = 0$ denoting the total sales in period t. To keep the estimation of the sales process as simple as possible we assume that the random variables $\Delta \mathbf{S}(t), t \in \mathbb{N}$ are independent with finite unknown first moment $m(t) := \mathbb{E}(\Delta \mathbf{S}(t))$. This means that

$$\Delta \mathbf{S}(t) = m(t) + \boldsymbol{\epsilon}_t \tag{4.2.1}$$

with $\boldsymbol{\epsilon}_t$, $t \in \mathbb{N}$ a sequence of independent and identically distributed random variables having zero mean. In an alternative model we also assume

$$\Delta \mathbf{S}(t) = m(t)\boldsymbol{\epsilon}_t \tag{4.2.2}$$

and $\boldsymbol{\epsilon}_t$, $t \in \mathbb{N}$ a sequence of independent and non-negative identically distributed random variables with mean 1. The model in relation (4.2.1) is called the *additive* sales model while the model in relation (4.2.2) is called the *multiplicative* sales model. Clearly the multiplicative sales model can be transformed into the additive sales model by using the $\ln$-transformation. This means that the model for the logarithm of the data set is given by an additive sales model. If T denotes the number of observations, we propose in parametric statistics some class of parametrized functions $f_\theta : \{1, 2, ..., T\} \to \mathbb{R}, \theta \in \Theta \subseteq \mathbb{R}^k$ with $\Theta \subseteq \mathbb{R}^k$ the so-called *parameter set* which serve as an estimation for the mean of the sales at time t. This means we assume that the true functional form of the random variable $\Delta \mathbf{S}(t)$ is given by

$$\Delta \mathbf{S}(t) = f_\theta(t) + \boldsymbol{\epsilon}_t$$

or

$$\Delta \mathbf{S}(t) = f_\theta(t)\boldsymbol{\epsilon}_t.$$

and using our data we need to estimate the unknown θ. Hence using some penalty function and our data vector $\Delta \mathbf{S} := (\Delta \mathbf{S}(1), ..., \Delta \mathbf{S}(T))$ we determine the parameter $\theta \in \Theta$ which has the smallest penalty. In this section we always use mean squared error as the penalty function. This means with T denoting the number of observations and introducing

$$\| v \|_2 := \left(\sum_{t=1}^{T} | v(t) |^2 \right)^{\frac{1}{2}}$$

for any finite sequence $v : \{1, 2..., T\} \to \mathbb{R}$ that for the model in relation (4.2.1) we need to solve the optimization problem

$$\inf_{\theta \in \Theta} \{ \| \Delta \mathbf{S} - f_\theta \|_2^2 \}. \tag{P}$$

For the multiplicative sales model in relation (4.2.2) our optimization problem reduces to

$$\inf_{\theta\in\Theta}\{\| \ln(\Delta\mathbf{S}) - \ln(f_\theta) \|_2^2\} \qquad (Q)$$

with $\ln(\Delta\mathbf{S}) = (\ln(\Delta\mathbf{S}(1)), ..., \ln(\Delta\mathbf{S}(T))$.

4.2.1 The additive model

The first parametric model we consider is given by the additive sales model

$$\Delta\mathrm{S}(t) = f_\theta(t) + \boldsymbol{\epsilon}_t, t = 1, ..., T$$

with the class of parametric functions f_θ, $\theta \in \Theta$ given by

$$f_\theta(t) = \theta_0 + \theta_1 t + ... + \theta_k t^k, k \le T+1$$

and $\Theta = \mathbb{R}^{k+1}$. This is the so-called *polynomial regression model* [7]. This means we assume that our true functional model is given

$$\Delta\mathrm{S}(t) = \theta_0 + \theta_1 t + ... + \theta_k t^k + \boldsymbol{\epsilon}_t, t = 1, ..., T. \qquad (4.2.3)$$

In matrix notation relation (4.2.3) reduces to

$$\Delta\mathrm{S} = X\boldsymbol{\theta} + \boldsymbol{\epsilon} \qquad (4.2.4)$$

with $\boldsymbol{\theta}^\intercal = (\theta_0, ..., \theta_k)$, $\Delta\mathrm{S}^\intercal = (\Delta\mathrm{S}(1), ..., \Delta\mathrm{S}(T))$, $\boldsymbol{\epsilon}^\intercal = (\boldsymbol{\epsilon}_1, ..., \boldsymbol{\epsilon}_T)$ and

$$X = \begin{pmatrix} 1 & 1 & 1 & ... & 1 \\ 1 & 2 & 2^2 & ... & 2^k \\ ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... \\ 1 & T-1 & (T-1)^2 & ... & (T-1)^k \\ 1 & T & T^2 & ... & T^k \end{pmatrix}$$

Using relation (4.2.4) it follows that

$$\| \Delta\mathrm{S} - f_\theta \|_2^2 = (\Delta\mathrm{S} - X\boldsymbol{\theta})^\intercal(\Delta\mathrm{S} - X\boldsymbol{\theta}) = \Delta\mathrm{S}^\intercal\Delta\mathrm{S} - 2\Delta\mathrm{S}^\intercal X\boldsymbol{\theta} + \boldsymbol{\theta}^\top X^\intercal X\boldsymbol{\theta}$$

and so our optimization problem (P) is given by

$$\inf\{\Delta\mathbf{S}^{\intercal}\Delta\mathbf{S} - 2\Delta\mathbf{S}^{\intercal}X\boldsymbol{\theta} + \boldsymbol{\theta}^{\top}X^{\intercal}X\boldsymbol{\theta} : \boldsymbol{\theta} \in \mathbb{R}^{k+1}\}.$$

The same optimization problem shows up in linear regression and it well known ([32]) that the optimal $\boldsymbol{\theta}_{opt}$ satisfies the so-called *normal equations*

$$X^{\intercal}X\boldsymbol{\theta}_{opt} = X^{\intercal}\Delta\mathrm{S}. \tag{4.2.5}$$

Hence we can use a linear regression package to compute the optimal $\boldsymbol{\theta}_{opt}$. Although the matrix $X^{\intercal}X$ is related to the so-called Hankel matrix [39] it seems to be difficult to give a closed form expression in relation (4.2.5) for this matrix. Observe that the matrix $X^{\top}X = (\overline{x}_{ij})$ is a symmetric $(k+1) \times (k+1)$ matrix with

$$\overline{x}_{ij} = \sum\nolimits_{t=1}^{T} t^{i+j-2}, 1 \le i \le k+1, 1 \le j \le k+1$$

In particular, for $k = 2$ we obtain a 3×3 matrix and $\sum_{t=1}^{T} t^{i+j-2}$ can be computed analytically. In the next section we only fit for $k = 1,2,3,4$ this class of parametric functions to our data set. By the same approach it is also easy to fit the class of parametric functions

$$f_{\theta}(t) = \theta_0 + \theta_1 g_1(t) + ... + \theta_k g_k(t)$$

with $g_j(t)$ some given functions. The set $g_j, j = 1, ...k$ are known as *basis functions* (Chapter 3 of [7]). In this case we assume that our true functional model is given by

$$\Delta\mathrm{S}(t) = \theta_0 + \theta_1 g_1(t) + ... + \theta_k g_k(t) + \boldsymbol{\epsilon}_t, t = 1, ...T. \tag{4.2.6}$$

Observe relation (4.2.3) reduces in matrix notation to

$$\Delta\mathrm{S} = X\boldsymbol{\theta} + \boldsymbol{\epsilon} \tag{4.2.7}$$

with $\boldsymbol{\theta}^\intercal = (\theta_0, ..., \theta_k)$, $\Delta\mathbf{S}^\intercal = (\Delta\mathbf{S}(1), ..., \Delta\mathbf{S}(T))$, $\boldsymbol{\epsilon}^\intercal = (\boldsymbol{\epsilon}_1, ..., \boldsymbol{\epsilon}_T)$ and

$$X = \begin{pmatrix} 1 & g_1(1) & g_2(1) & ... & g_k(1) \\ 1 & g_1(2) & g_2(2) & ... & g_k(2) \\ ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... \\ 1 & g_1(T-1) & g_2(T-1) & ... & g_k(T-1) \\ 1 & g_1(T) & g_2(T) & ... & g_k(T). \end{pmatrix}$$

and we proceed the analysis as before.

4.2.2 The multiplicative model

The other parametric model we consider is the multiplicative sales model

$$\Delta\mathbf{S}(t) = f_\theta(t)\boldsymbol{\epsilon}_t, t = 1, ..., T$$

with the class of parametric functions given by

$$f_\theta(t) = e^{\theta_0 + \theta_1 t + ...\theta_{k-1} t^{k-1}} t^{\theta_k}, \boldsymbol{\theta}^\intercal = (\theta_0, ..., \theta_k) \in \mathbb{R}^{k+1}.$$

For $k = 2$ this model is the same as the model proposed in [12]. Observe it is equivalent to assume that the true functional model for the logarithm transformation of the data set is given by

$$\ln(\Delta\mathbf{S}(t)) = \theta_0 + \theta_1 t + ...\theta_{k-1} t^{k-1} + \theta_k \ln(t) + \ln(\boldsymbol{\epsilon}_t), t = 1, ..., T. \quad (4.2.8)$$

In matrix notation relation (4.2.8) reduces to

$$\ln(\Delta\mathbf{S}) = X\boldsymbol{\theta} + \boldsymbol{\epsilon} \quad (4.2.9)$$

with $\boldsymbol{\theta}^\intercal = (\theta_0, ..., \theta_k)$, $\ln(\Delta \mathbf{S})^\intercal = (\ln(\Delta \mathbf{S}(1)), ..., \ln(\Delta \mathbf{S}(T))$, $\boldsymbol{\epsilon}^\intercal = (\boldsymbol{\epsilon}_1, ..., \boldsymbol{\epsilon}_T)$ and

$$X = \begin{pmatrix} 1 & 1 & 1 & ... & 0 \\ 1 & 2 & 2^2 & ... & \ln(2) \\ ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... \\ 1 & T-1 & (T-1)^2 & ... & \ln(T-1) \\ 1 & T & T^2 & ... & \ln(T). \end{pmatrix}$$

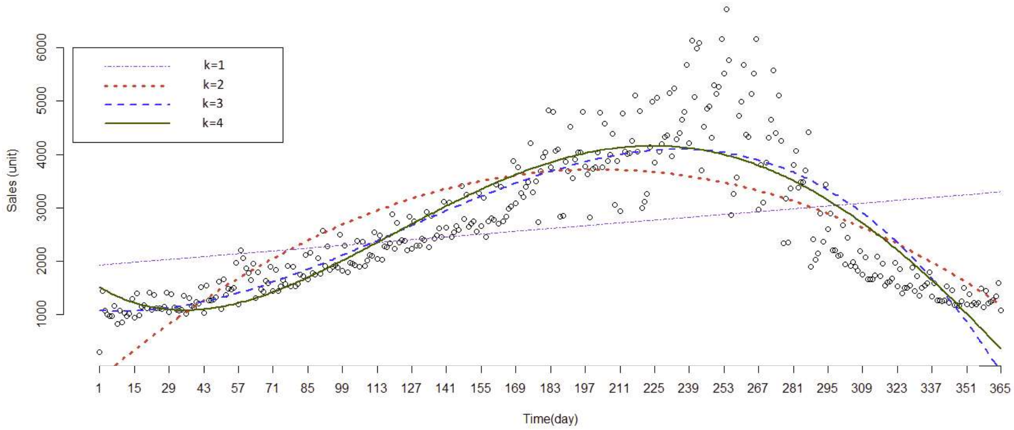
In Appendix B we discuss part of the theory of the linear model. In the next two sections we fit the additive and multiplicative models to our daily sales data. Parameters are estimated and by providing some goodness-of-fit statistics, we determine the best model to fit to our sales data.

4.3 Fitting the additive model to daily sales

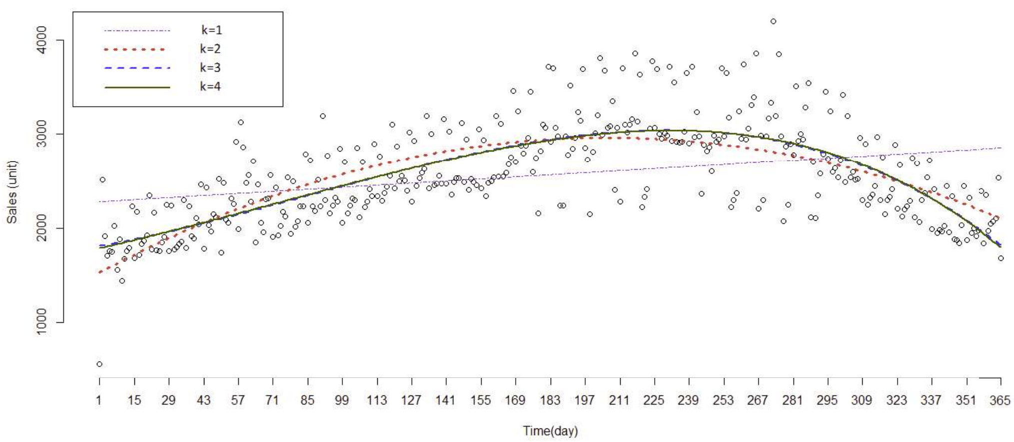
In this subsection we compute the least squares estimates for the unknown parameters of the polynomial regression model given in Equation (4.2.3). In Figures 4.6a and 4.6b, the polynomial regression model for $k = 1, 2, 3, 4$ is fitted to the daily sales data. As can be seen from the figures, and supported by the coefficient R^2 of multiple correlation ([32]) and the adjusted coefficient R^2 (adj) shown in Table 4.1, the additive sales model with $k = 4$ is the best fit to the daily sales data for both products. Note we do not take k larger than 4 since for larger k we face the problem of overfitting ([7]). Since there is no noticeable difference for both products in the coefficient of multiple correlation and the adjusted coefficient with $k = 3$ or 4 we prefer due to reducing the problem of overfitting to select the additive sales model with $k = 3$ as the best choice to fit our data.

Table 4.1: Goodness-of-fit measurements for the additive model

	Product A		Product B	
Degree	R^2	R^2 (adj)	R^2	R^2 (adj)
k=1	0.0903	0.0878	0.0976	0.0951
k=2	0.6155	0.6133	0.5030	0.5002
k=3	0.7517	0.7446	0.5445	0.5407
k=4	0.7647	0.7621	0.5447	0.5397



(a) Product A



(b) Product B

Figure 4.6: The additive models fitted over average daily sales data for product A (a) and product B (b)

4.4 Fitting the multiplicative model to daily sales

Our multiplicative model is an extension of the model discussed in [12]. The author proposes two models to estimate the parameters of life cycle shape. These functions are drawn in Figures 4.7 and 4.8 respectively.

The first model is simply the additive combination of two functions describing the sales data in terms of time

$$\mathbb{E}(\Delta \mathbf{S}_t) = a + bt_1^2 + \frac{c}{t_2} \tag{4.4.1}$$

where $\Delta \mathrm{S}(t)$ stands for sales, t_1 for the time of growth and t_2 for the time of

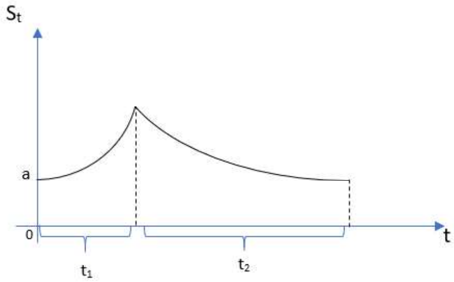


Figure 4.7: Brockhoff's additive model

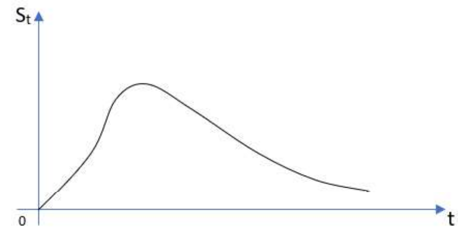


Figure 4.8: Brockhoff's multiplicative model

decline, and a, b and c are the parameters. The second proposed model in this paper is the multiplicative model for $k = 2$ given by

$$\mathbb{E}(\Delta \mathbf{S}_t) = at^b e^{-ct} \tag{4.4.2}$$

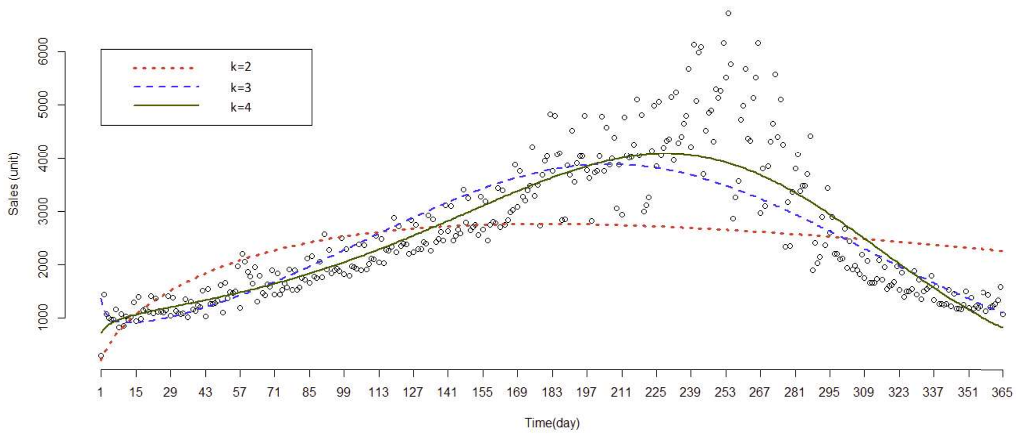
where $\Delta \mathrm{S}(t)$ stands for random sales at time t. In particular our relation (4.2.8) reduces for $k = 2$ to relation (4.4.2) with $a = e^{\theta_0}$, $b = \theta_2$ and $c = -\theta_1$.

The fitted multiplicative models for $k = 2, 3, 4$ are depicted in Figure 4.9a and 4.9b. As expected the multiplicative model for $k = 4$ gives the best fit to the data. In Table 4.2 we also compute for both products the coefficient R^2 of multiple correlation and the adjusted coefficient R^2 (adj) applied to the logarithm transformation of the data set. From Tables 4.1 and 4.2, we observe that assuming a linear relation for the logarithm transformation of the daily sales explains the data much better than assuming a linear relation for the daily sales data. Hence we believe the best model for daily sales for both product types is the multiplicative sales model with $k = 4$.

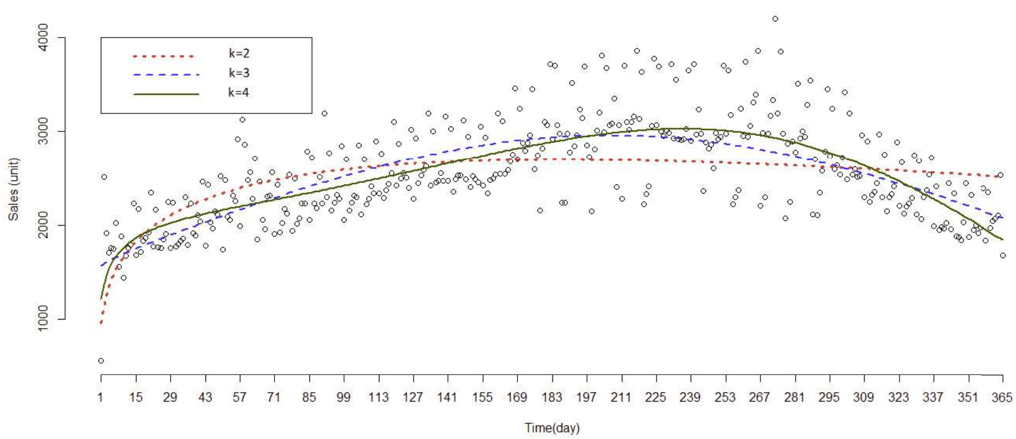
Table 4.2: Goodness-of-fit measurements for the multiplicative model

	Product A		Product B	
Degree	R^2	R^2 (adj)	R^2	R^2 (adj)
k=2	0.4027	0.3994	0.3776	0.3741
k=3	0.8033	0.8017	0.5353	0.5314
k=4	0.8524	0.8508	0.5800	0.5753

The daily data exhibits large fluctuation. Therefore, we expect the R-squared measure to improve if we fit the models using weekly data, which smooths the daily fluctuations. This is what we do in the next subsection.



(a)

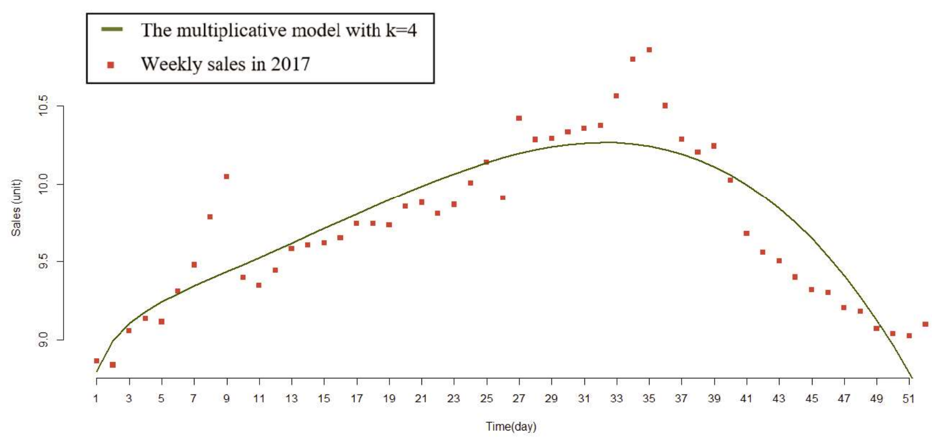


(b)

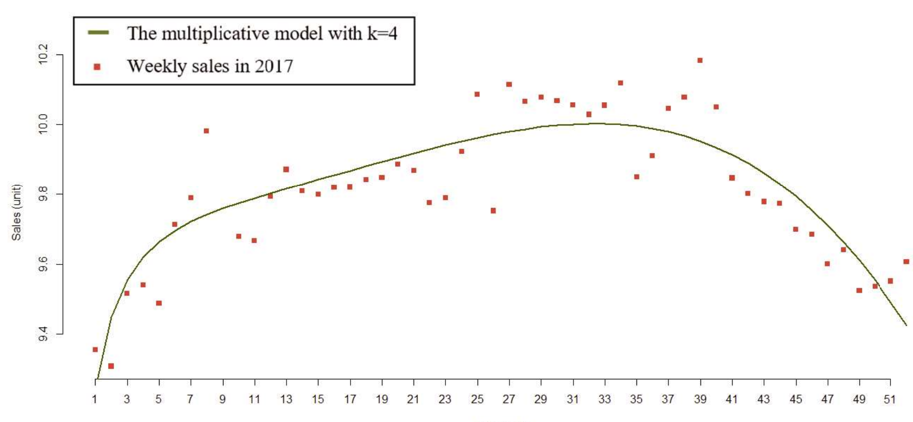
Figure 4.9: The multiplicative models fitted over average daily sales data of product A (a) and product B (b)

4.5 Fitting the multiplicative model using a smoothed version of the data

In this subsection we aggregate the data and compute the weekly sales data. By doing so we expect to eliminate the fluctuations due to daily sales variations. For this purpose we take the data of year 2017 since the pattern of sales is almost the same in each year. To construct the weekly sales we simply sum up the sales amount of every seven days starting from Sunday 1/1/2017. We use the multiplicative model with $k = 4$ as we observed this model to be successful with daily data. Figures 4.10a and 4.10b present the weekly average (over five years) sales data together with the multiplicative model fit. As it also be seen from the figure, the model seems to fit well and the R-squared values are 0.80 and 0.69 for products A and B respectively.



(a)



(b)

Figure 4.10: The multiplicative model fitted over weekly sales of product A (a) and product B (b) in 2017

Next, we obtain the daily sales estimations from the weekly sales estimations by calculating daily weights. Let X_i denote the total sales in week i, and $X_{i,j}$ denote the sales on day j of week i, where $j \in 1, 2, ..., 7$ and 7 correspnds to the days of the week. We calculate the daily weight for day j as

$$\bar{Y}_j = \frac{\sum_{i=1}^{52} X_{i,j}}{\sum_{i=1}^{52} X_i}, \qquad j = 1, ..., 7. \tag{4.5.1}$$

This weight estimates the percentage of weekly sales happening on day j of any week. To obtain daily sales estimations, one multiplies the relevant daily weight with the weekly sales estimation for a particular week. The weights that we obtain are provided in Table 4.3 for different days of the week.

Table 4.3: Weights for each day of the week

Day-of-week	Tue	Wed	Thu	Fri	Sat	Sun	Mon
Weight (Product A)	14.08	16.74	15.39	15.50	16.23	4.50	17.57
Weight (Product B)	14.27	16.73	15.31	15.04	16.19	3.82	17.50

Figure 4.11 illustrate the daily sales estimations obtained with this approach together with the real sales data for the two products. The R-squared values are 0.27 and 0.21 for products A and B respectively which shows that the model does not fit very well to the data. This is because we are using equal daily weights for different weeks of the year while there might be different in daily weights from one week to another.

Figure 4.12 is the residuals vs. order plot which is used to verify the assumption that the error terms are independent. Residuals are almost distributed around zero and there is no trend or seasonality within them so they are independent. But what we see in the plot is that there are some points spread randomly but not around 0. These error terms belong to sales on Sunday which as we discussed before provide a separate cluster of points (Look at Figure 4.11).

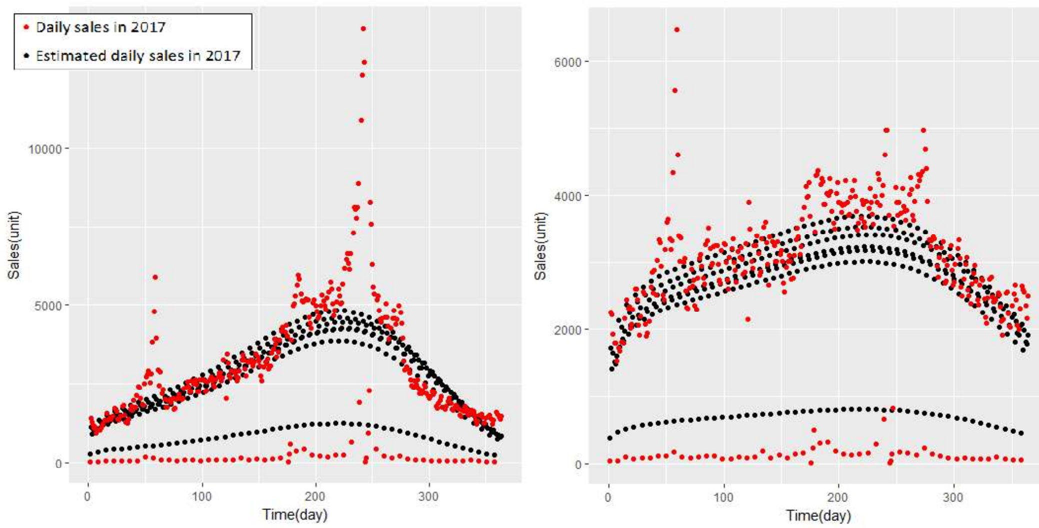


Figure 4.11: The multiplicative model fitted over daily sales of product A (a) and product B (b) in 2017. Left: Product A, right: product B.

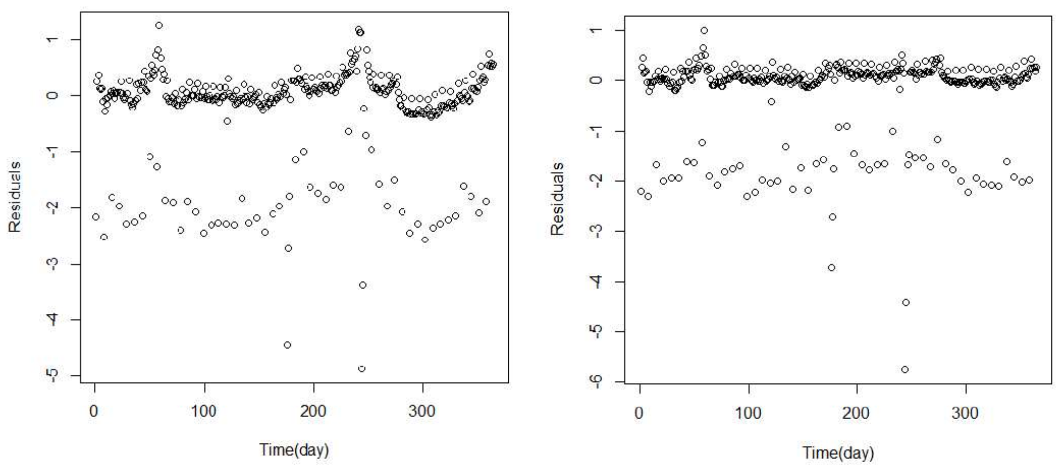


Figure 4.12: Residuals vs. time (Left: product A , right: product B)

4.6 Model selection

So far in this chapter, after testing the additive and multiplicative models, we observe that the multiplicative model fits better to our daily sales data. One can refer to Section 4.2.2 for the parametric function of the multiplicative model with degree k. Later, we performed the fit over weekly data, assuming that a smoothed version of the data would give a better fit. The weekly estimated values are then converted into daily estimations using estimated daily weight factors. Use of weekly values improved the R-squared performance of product B whereas it slightly deteriorated the performance for product A (Look at table 4.2). Thus, the best performing estimation approach can be product-dependent. In order to provide estimations for the daily sales, we used the daily weights provided in 4.3. The R-squared values we obtained were rather low. Which means using weekly data to provide estimations for the daily sales is not successful. Therefore, we propose using daily values for estimation of daily sales and using the weekly data in case you the weekly estimations are needed.

Subsequently, the model that we use for the estimation of the weekly sales of product A and B becomes

$$\mathbb{E}_A(\Delta\mathbf{S}(t)) = e^{8.83-0.044t+0.0038t^2-0.00007t^3}t^{0.34} \tag{4.6.1}$$

$$\mathbb{E}_B(\Delta\mathbf{S}(t)) = e^{9.3+0.057t+0.0022t^2-0.00003t^3}t^{0.38} \tag{4.6.2}$$

One can refer to relations (3.1.27) in Corollary 1 and (3.1.35) to understand the relation between the analysis in this chapter and the proposed model discussed in Chapter 3. Note that because time is measured in weeks, some of the coefficients in these equations need to be very small values. Otherwise, the function cannot capture the required curvature of the sales process.

Chapter 5

ESTIMATING THE INTENSITY MEASURE OF THE FAILURE PROCESS

In this chapter we apply statistical techniques to our data set to estimate the intensity measure of the failure process. We do so by estimating the cumulative distribution function of the random time to the *first failure*. This is sufficient because we assume that the minimal repair assumption holds (see Chapter 3). The issue we face is that we do not observe at least one failure for all sold items; we have failure time data only for those items that failed and returned to the firm within their warranty coverage. Hence we are dealing with a censored data set. In Section 5.1, we discuss the main characteristics of this set. In Section 5.2, we consider the statistical method to estimate the intensity measure of our failure process.

In particular, we discuss the well-known Maximum Likelihood Principle applied to censored data and to an arbitrary parametric class of cumulative distribution functions. We specify this approach to the parametric class of Weibull cumulative distribution functions, and present an algorithm to calculate the so-called maximum likelihood estimators for this class. The main reason for using this specific parametric class is its popularity within the maintenance literature. Finally in Section 5.3 we apply this approach to our data set and discuss the quality of the fit we obtain.

5.1 Failure data

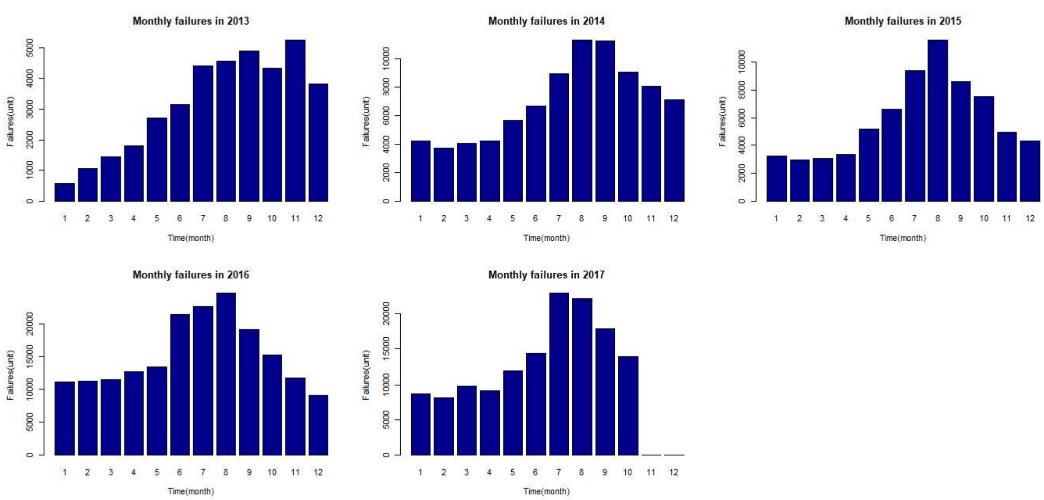
In this section we provide some graphical and statistical information about the failure data of the two products (Product A and Product B) that we study. Each individual item of these two products is tracked with a specific identifier. We can match the sales and failure data of items using that identifier. For each item sold, the firm provided us the warranty start date (which corresponds to the item installation date) and dates in which the item is brought (if any) to the firm for repairs within the three-year warranty period. The data is in daily detail, covering the period from January 2013 to October 2017. Note that we use the “repair time” in the dataset as a proxy for the real “failure time” because we do not know when exactly the failure occurred. In practice, the consumer may bring the item to repair after some delay, but this is not likely to be a long time. Figure 5.1 depicts the first six columns of our database.

Figure 5.1: Failure data in raw form

	A	B	C	D	E	F
1	ProductCode	ProductionYear	ProductionMonth	SerieNo	RepairTime	WarrantyStart
2	60001	1980	1	121927	6/14/2013	4/23/2013
3	60001	1980	1	121927	5/11/2013	4/23/2013
4	60001	1994	2	201210	5/24/2013	4/25/2013
5	60001	2005	10	118416	2/4/2015	5/21/2013

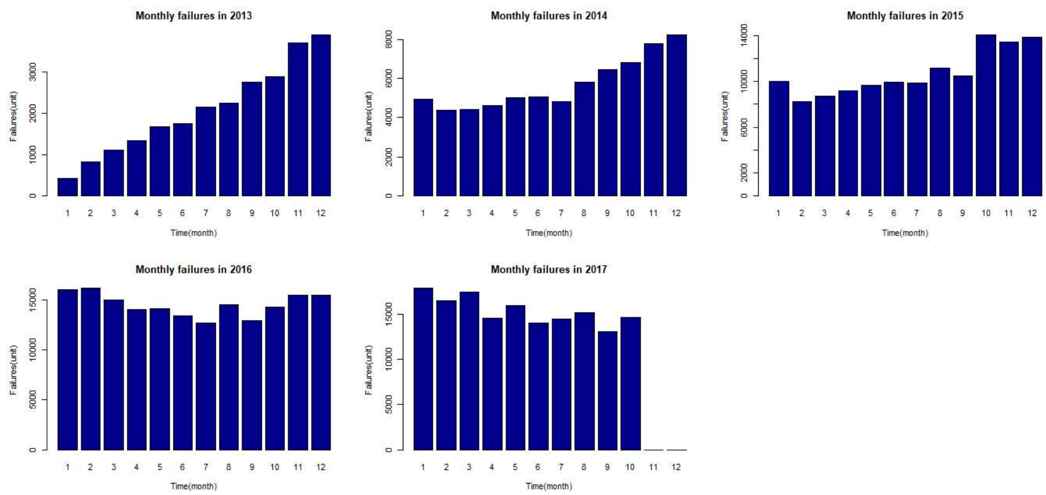
Using this data, we first check if there is a pattern or trend in the total failure data. Figures 5.2 and 5.3 present the monthly number of failures for Products A and B respectively. We observe the number of failures to be higher is summer months for product A. We could not find such a pattern for product B. For instance, in 2013 the number of product B failures increased over months, while in 2015 there was no such trend. Note, however, that one cannot understand the failure rate (or, failure intensity) of the product from these figures as this also depends on the number of products in use in each month.

Figure 5.2: Number of failures by month (product A)



In our following estimation study, we use only the data of items sold (installed) in year 2013. We refer to the time between the installation of the item and its first failure as the item’s *lifetime*. As discussed earlier, we know the

Figure 5.3: Number of failures by month (product B



lifetime of items that failed at least once within the warranty period (uncensored data), but not of items that did not fail during the warranty periods (censored data). As summarized in Table 5.1, only a small percentage of items experience a failure within the warranty period.

Table 5.1: Failure data statistics

	Product A (%)	Product B (%)
Uncensored	160,770 (17.38)	115,527 (12.7)
Censored	763,906 (82.62)	794,378 (87.3)
Total	924,676	909,905

Each item can fail multiple times during its warranty period. Table 5.2 shows the frequency of failures. For example, we observe 484 items of product B to fail five times during the warranty period.

Table 5.2: Frequency of failures during the warranty period

Number of failures	1	2	3	4	5	> 5	total	total failures
Product A	128,787	24,481	5,404	1,337	344	117	160,770	201,911
Product B	92,327	16,747	4,304	1,337	484	328	115,624	148,613

Figures 5.4 and 5.5 depict the frequency distribution of lifetimes of products A and B. We observe a significant percentage of failures to occur within the first 50 days for both products. This is consistent with the reported observations in maintenance literature [52]. Products with production defects are likely to be discovered within this initial usage period. Beyond that, product A failures exhibit seasonality such that the item is more likely to fail around one year, two

year etc. after installation. This may be due to a joint seasonality in installation time and time of failure, such as refrigerators installed and failed mostly in summer months. On the other hand, we don't observe such seasonality with product B. With that product, after the initial 50 day period, the frequency of failures increase slowly over time.

Figure 5.4: Lifetimes of uncensored Product A items sold in 2013

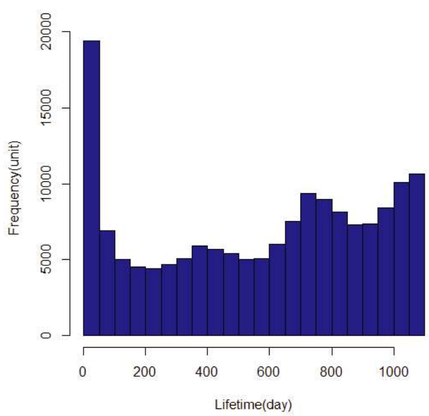
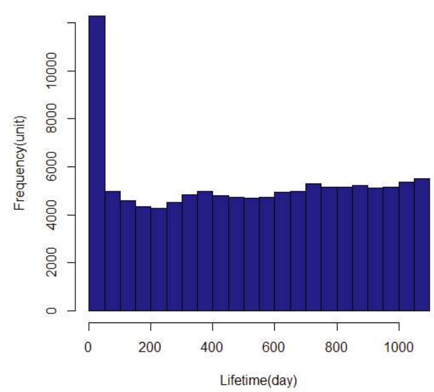


Figure 5.5: Lifetimes of uncensored Product B items sold in 2013



We also plotted the hazard rate function of the uncensored and censored lifetimes all together (Figures 5.6 and 5.7). The formulas of obtaining this functions values are described in the Section 5.2. As it can be seen in the figures, the hazard rate function of both products has the same pattern as their frequency plot. The hazard rate functions start with a high peak and end with a slight upward shape. This means the hazard rate function has a bathtub shape.

Figure 5.6: Hazard rate function of lifetimes of items sold in 2013 (Product A)

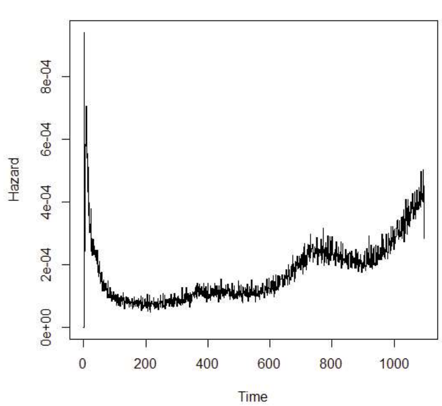
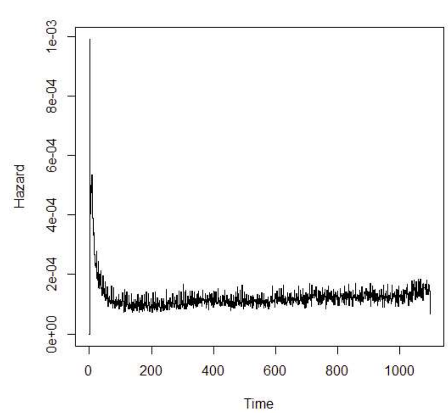


Figure 5.7: Hazard rate function of lifetimes of items sold in 2013 (Product B)



Although Weibull hazard does not form a bathtub shape, considering its popularity in the lifetime data analysis we will use this distribution in our estimations of lifetimes.

5.2 Methodology

In this section, we first provide some definitions on the basic concepts used in lifetime estimation process. Then, we discuss statistical procedures to estimate censored random variables (MLE method). In the next section, we introduce the Weibull distribution, a MLE-tailored method for the Weibull-distributed censored data, and a special purpose algorithm to obtain the estimated optimal parameters of the distribution.

5.2.1 Basic concepts

We first list some well-known definitions.

Definition 3 *The so-called failure rate or hazard function* $\lambda : \mathbb{R}_+ \to \mathbb{R}_+$ *of a non-negative random variable* $\mathbf{X}$ *with unbounded support is given by*

$$\lambda(x) := \frac{f(x)}{1 - F(x)}, x > 0 \tag{5.2.1}$$

with f *denoting the density of the cumulative distribution function* F *of the random variable* $\mathbf{X}$*. The so-called survival function is given by*

$$S(x) := 1 - F(x). \tag{5.2.2}$$

By the definition of the survival function we know that

$$\frac{d \ln S}{dx}(x) = -\lambda(x)$$

and this shows by the main theorem of integration that

$$\ln(S(x)) - \ln(S(0)) = -\Lambda(x)$$

with

$$\Lambda(x) = \int_0^x \lambda(u)du \tag{5.2.3}$$

the so-called *cumulative hazard function*. Since we always assume for non-negative random variables that $F(0) = 0$ and hence $S(0) = 1$ it follows that

$$S(x) = 1 - F(x) = e^{-\Lambda(x)}. \tag{5.2.4}$$

Although not standard, one can also define the failure rate function for real valued random variables $\mathbf{X}$ having a cdf on $\mathbb{R}$. In this case we also obtain by standard differentiation that

$$S(x) = e^{-\Lambda(x)}, x \in \mathbb{R} \tag{5.2.5}$$

with

$$\Lambda(x) = \int_{-\infty}^{x} \lambda(u)du.$$

Since $F(\infty) := \lim_{x\uparrow\infty} F(x) = 1$ it follows that $\lim_{x\uparrow\infty} S(x) = 0$ and so by relation (5.2.5)

$$\Lambda(\infty) = \lim_{x\uparrow\infty} \Lambda(x) = \infty. \tag{5.2.6}$$

In order to portray a random sample one can graph the empirical survival function (ESF). If there are no censoring within the data, ESF is defined as:

$$\hat{S}(t) = \frac{Number\ of\ observations \geq t}{n}, t \geq 0 \tag{5.2.7}$$

In case the data we are dealing contains censored values, we have to modify the survival function as follow:

$$\hat{S}(t) = \Pi_{j:t_j<t} \frac{n_j - d_j}{n_j} \tag{5.2.8}$$

Where t_j is the failure time of the $j^t h$ item, $j = 1, ..., n$, d_j is the number of failures that occur at t_j and n_j is the number of uncensored items with lifetimes t_j and greater then t_j.

We next introduce the definitions of location, scale and shape parameters.

Definition 4 *We introduce the definitions of location, scale and shape parameters. The notation $\mathbf{X} \stackrel{d}{=} \mathbf{Y}$ means that the random variables $\mathbf{X}$ and $\mathbf{Y}$ have the same cumulative distribution function.*

1. *A parameter $\gamma \in \mathbb{R}$ is called a location parameter of the random variable $\mathbf{X}$ if*

$$\mathbf{X} \stackrel{d}{=} \gamma + \mathbf{Y}$$

with $\mathbf{Y}$ having a cdf which is independent of γ.

2. *A parameter $\beta > 0$ is called a scale parameter if*

$$\mathbf{X} \stackrel{d}{=} \beta\mathbf{Y}$$

with $\mathbf{Y}$ having a cdf which is independent of β.

3. *The parameters $\gamma \in \mathbb{R}$ and $\beta > 0$ are called the location and scale parameters of the random variable $\mathbf{X}$, respectively if*

$$\mathbf{X} \overset{d}{=} \gamma + \beta \mathbf{Y}$$

with the cdf of $\mathbf{Y}$ independent of γ and β

4. *The parameter $\alpha > 0$ is called a shape parameter of the nonnegative random variable $\mathbf{X}$ if*

$$\mathbf{X} \overset{d}{=} \mathbf{Y}^{\alpha^{-1}}$$

with $\mathbf{Y}$ a non-negative random variable with cdf independent of α

5. *The parameters $\gamma \in \mathbb{R}$ and $\beta > 0$ and $\alpha > 0$ are called respectively the location, scale and shape parameters of the random variable $\mathbf{X}$ if*

$$\mathbf{X} \overset{d}{=} \gamma + \beta \mathbf{Y}^{\alpha^{-1}}$$

with $\mathbf{Y}$ a non-negative random variable with cdf independent of γ, β and α.

A well known example of a cdf in the domain of maintenance belonging to the above class mentioned in part 5 of Definition 4 and depending on the three parameters γ, α and β is given by the Weibull cdf. In this case the random variable $\mathbf{Y}$ has an exponential cdf with parameter 1 given by

$$\mathbb{P}(\mathbf{Y} \leq x) = 1 - e^{-x}. \tag{5.2.9}$$

If we consider the most general parametric model having a location parameter $\gamma \in \mathbb{R}$, a scale parameter $\beta > 0$ and a shape parameter $\alpha > 0$ then one can achieve the density function and the cumulative hazard rate function as in the following.

Lemma 14 *Let $\mathbf{Y}$ be a non-negative random variable with cdf F independent of the parameters $\gamma \in \mathbb{R}$ and $\beta, \alpha > 0$ with density f, failure rate function λ and cumulative hazard rate function Λ and set $\mathbf{X} \overset{d}{=} \gamma + \beta \mathbf{Y}^{\alpha^{-1}}$. Then the following results hold.*

1. *For every $x \geq \gamma$*

$$\mathbb{P}_\theta(\mathbf{X} \leq x) = F(\boldsymbol{\beta}^{-\alpha}(x-\gamma)^\alpha) \tag{5.2.10}$$

with density

$$f_\theta(x) = \alpha\beta^{-\alpha}(x-\gamma)^{\alpha-1} f(\boldsymbol{\beta}^{-\alpha}(x-\gamma)^\alpha) 1_{[\gamma,\infty)}(x). \tag{5.2.11}$$

2. *The cumulative hazard rate function is given by*

$$\Lambda_\theta(x) = \Lambda(\beta^{-\alpha}(x-\gamma)^\alpha). \tag{5.2.12}$$

Proof. To show the first part we observe using $\mathbf{X} \stackrel{d}{=} \gamma + \beta \mathbf{Y}^{\alpha^{-1}}$ that for every $x \geq \gamma$

$$\begin{aligned} F_\theta(x) &= \mathbb{P}_\theta(\mathbf{X} \leq x) \\ &= \mathbb{P}_\theta(\gamma + \beta \mathbf{Y}^{\alpha^{-1}} \leq x) \\ &= \mathbb{P}(\mathbf{Y} \leq (\boldsymbol{\beta}^{-1}(x-\gamma))^\alpha) \\ &= F(\boldsymbol{\beta}^{-\alpha}(x-\gamma)^\alpha). \end{aligned} \tag{5.2.13}$$

This implies taking the derivative of the cdf F_θ that its density is given by

$$f_\theta(x) = \alpha\beta^{-\alpha}(x-\gamma)^{\alpha-1} f(\boldsymbol{\beta}^{-\alpha}(x-\gamma)^\alpha) 1_{[\gamma,\infty)}(x). \tag{5.2.14}$$

To show the second part of the lemma we observe by relation (5.2.4) and (5.2.13) that for every $x \geq \gamma$

$$\begin{aligned} \Lambda_\theta(x) &= -\ln(1 - F_\theta(x)) \\ &= -\ln(1 - F((\boldsymbol{\beta}^{-1}(x-\gamma))^\alpha) \\ &= \Lambda((\boldsymbol{\beta}^{-1}(x-\gamma))^\alpha) \end{aligned}$$

and we have shown the result. □

In case we take $\gamma = 0$ we obtain as a special case

$$\mathbb{P}_\theta(\mathbf{X} \leq x) = F(\boldsymbol{\beta}^{-\alpha} x^\alpha) \tag{5.2.15}$$

and

$$\Lambda_\theta(x) = \Lambda(\beta^{-\alpha} x^\alpha). \tag{5.2.16}$$

5.2.2 Maximum Likelihood Estimation for censored data

In this section we introduce the maximum likelihood principle to estimate the underlying cdf of censored data. If we have a sample of censored type-1 distributed random variables then this sample is the realization $(t_1, ..., t_m)$ of the random vector $(\mathbf{T}_1, ..., \mathbf{T}_m)$ with

$$\mathbf{T}_n = \mathbf{X}_n 1_{\{\mathbf{X}_n \leq C_n\}} + \Delta 1_{\{\mathbf{X}_n > C_n\}}. \tag{5.2.17}$$

If each of these independent and identically distributed random variables $\mathbf{X}_n$ satisfy $\mathbf{X}_n \sim F_\theta, \theta \in \Theta$ with density $f_\theta(x), \theta \in \Theta$, then the likelihood function of the given sample $(t_1, ..., t_m)$ is given by the function

$$L(\theta; t_1, ..., t_m) = \Pi_{n=1}^m f_\theta(t_n)^{1_{\{t_n \neq \Delta\}}(t_n)} S_\theta(C_n)^{1_{\{t_n = \Delta\}}(t_n)} \tag{5.2.18}$$

with $S_\theta(t) := \mathbb{P}_\theta(\mathbf{X}_n > t)$ and

$$1_A(t) = \begin{cases} 1 & \text{if } t \in A \\ 0 & \text{if } t \notin A. \end{cases}$$

Introducing the log-likelihood function $LL : \Theta \mapsto \mathbb{R}_+$ given by

$$LL(\theta) := \ln(L(\theta; t_1, ..., t_m) \tag{5.2.19}$$

it follows by relation (5.2.18) that

$$LL(\theta) = \sum_{n=1}^m 1_{\{t_n \neq \Delta\}}(t_n) \ln(f_\theta(t_n)) + 1_{\{t_n = \Delta\}}(t_n) \ln S_\theta(C_n). \tag{5.2.20}$$

If $C = \{n = 1, ..., m : t_n = \Delta\}$ denotes the set of censored data then an equivalent representation is given by

$$LL(\theta) = \sum_{n \notin C} \ln(f_\theta(t_n)) + \sum_{n \in C} \ln(S_\theta(C_n)). \tag{5.2.21}$$

In case $C_n = C$ for every n and and $\mid C \mid$denote the cardinality or total number of elements of the set C this reduces to

$$LL(\theta) = \sum_{n \notin C} \ln(f_\theta(t_n)) + (m - \mid C \mid) \ln S_\theta(C). \tag{5.2.22}$$

If the random variables $\mathbf{X}_n$ are discrete values and strictly positive then we obtain the log-likelihood function

$$LL(\theta) = \sum_{n \notin C} \ln(\mathbb{P}_\theta(\mathbf{X}_n = t_n)) + \sum_{n \in C} \ln(S_\theta(C_n)). \tag{5.2.23}$$

To rewrite the log-likelihood function in relation (5.2.21) in terms of failure rate functions, we observe by the definition of the failure rate function and relation (5.2.4) that

$$\begin{aligned} \ln(f_\theta(x)) &= \ln(\lambda_\theta(x)(1-F_\theta(x)) \\ &= \ln(\lambda_\theta(x)) + \ln(1-F_\theta(x)) \\ &= \ln(\lambda_\theta(x)) - \Lambda_\theta(x). \end{aligned} \tag{5.2.24}$$

This implies using relation (5.2.20) and (5.2.4) that

$$LL(\theta) = \sum\nolimits_{n\notin C} \ln \lambda_\theta(t_n) - \sum\nolimits_{n\notin C} \Lambda_\theta(t_n)) - \sum\nolimits_{n\in C} \Lambda_\theta(C_n). \tag{5.2.25}$$

Applying the maximum likelihood principle in statistics, we need to solve the optimization problem

$$v(P) = \sup\nolimits_{\theta\in\Theta} LL(\theta). \tag{P}$$

One of the main issues to prove is whether the above optimization problem has a unique optimal solution, and under which conditions we can find an algorithm which identifies this optimal solution. A sufficient condition to guarantee this is to show that the log-likelihood function is strictly concave in the parameter θ. If the parameter set Θ only consists of scale and shape parameters then the log-likelihood function can be simplified. This is shown in the next result.

Lemma 15 *If* $\mathbf{X} \stackrel{d}{=} \beta\mathbf{Y}^{\alpha^{-1}}$ *with* $\beta > 0$ *a scale parameter and* $\alpha > 0$ *a shape parameter and the non-negative random variable* $\mathbf{Y}$ *has a cdf* F *independent of the parameters* $\alpha, \beta > 0$ *with density* f *and cumulative hazard rate function* Λ *then the log-likelihood function of the given sample* $(t_1, ..., t_m)$ *is given by*

$$LL(\alpha,\beta) = \begin{cases} (m- \mid C \mid)(\ln(\alpha) + \ln(\beta^{-\alpha})) + (\alpha-1)\sum_{n\notin C}\ln(t_n) \\ + \sum_{n\notin C} \ln f(\beta^{-\alpha} t_n^\alpha) - \sum_{n\in C} \Lambda(\beta^{-\alpha} C_n^\alpha). \end{cases} \tag{5.2.26}$$

Proof. By relation (5.2.11) it follows with $\gamma = 0$ that

$$\ln f_\theta(t_n) = \ln(\alpha) + \ln(\beta^{-\alpha}) + (\alpha-1)\ln(t_n) + \ln f(\beta^{-\alpha} t_n^\alpha).$$

Moreover, by relation (5.2.4) and (5.2.12) we obtain

$$\ln(S_\theta(C_n)) = -\Lambda_\theta(C_n) = -\Lambda(\beta^{-\alpha} C_n^\alpha).$$

Applying relation (5.2.21) it follows

$$LL(\alpha,\beta)=\begin{cases}(m-\mid C\mid)(\ln(\alpha)+\ln(\beta^{-\alpha}))+(\alpha-1)\sum_{n\notin C}\ln(t_n)\\ +\sum_{n\notin C}\ln f(\beta^{-\alpha}t_n^{\alpha})-\sum_{n\in C}\Lambda(\beta^{-\alpha}C_n^{\alpha})\end{cases}$$

and we have shown the result. □

To analyse the structure of the optimization problem associated with Lemma 15 we observe the following. It follows using the definition of optimization problem P that by Lemma 15

$$\upsilon(P)=\sup_{\alpha>0}\varphi(\alpha)$$

with

$$\begin{aligned}\varphi(\alpha) &= \sup_{\beta>0} LL(\alpha,\beta)\\ &= (m-\mid C\mid)\ln(\alpha)+(\alpha-1)\sum_{n\notin C}\ln(t_n)+\sup_{\beta>0}h_\alpha(\beta^{-\alpha})\end{aligned}$$

and

$$h_\alpha(y)=(m-\mid C\mid)\ln(y)+\sum_{n\notin C}\ln f(yt_n^{\alpha})-\sum_{n\in C}\Lambda(yC_n^{\alpha}). \tag{5.2.27}$$

Since for every $\alpha>0$ it is easy to see that solving

$$\sup_{y>0}h_\alpha(y) \tag{5.2.28}$$

having optimal solution y_{opt} is the same as solving $\sup_{\beta>0}h_\alpha(\beta^{-\alpha})$ having as an optimal solution

$$\beta_{opt}(\alpha)=y_{opt}^{-\frac{1}{\alpha}}. \tag{5.2.29}$$

We need to come up with sufficient conditions on the density f to show the concavity of the function h_α. This is discussed in the next result.

Lemma 16 *If the density function f on $(0,\infty)$ is logconcave then the function h_α is strictly concave on $\mathbb{R}_+$. For $\lim_{y\downarrow 0}yf(y)=0$ it follows that*

$$\lim_{y\downarrow 0}h_\alpha(y)=-\infty$$

and for $\lim_{y\uparrow\infty}yf(y)=0$ we obtain

$$\lim_{y\uparrow\infty}h_\alpha(y)=-\infty.$$

Proof. Since by assumption the density f is log-concave on $\mathbb{R}_+$ it follows using Theorem $5.8B$ on page 74 of [24] that the survival function $y \to 1 - F(y)$ is also log-concave on $\mathbb{R}_+$. Since

$$1 - F(y) = e^{-\Lambda(y)}$$

this shows that the function $y \to -\Lambda(y)$ is concave on $\mathbb{R}_+$. Using now the definition of h_α in relation (5.2.27) and $y \to \ln(y)$ is strictly concave we obtain that the function h_α defined in relation (5.2.27) is strictly concave on $\mathbb{R}_+$. To show the limit relations we only prove it for $y \uparrow \infty$. The result for $y \downarrow 0$ can be proved similarly. It follows that

$$\ln(y) + \ln f(yt_n^\alpha) = \ln(yf(yt_n^\alpha)$$

and this implies using $\lim_{y\uparrow\infty} yf(y) = 0$ that

$$\lim_{y\uparrow\infty} \ln(y) + \ln(f(yt_n^\alpha) = -\infty.$$

This shows using relation (5.2.27) the desired result. □

By Lemma 16 we obtain for f log-concave that the optimization problem (5.2.28) for each $\alpha > 0$ has a unique finite optimal solution y_{opt} satisfying $h'_\alpha(y_{opt}) = 0$. We observe that

$$h'_\alpha(y) = (m- \mid C \mid)y^{-1} + \sum_{n\notin C} t_n^\alpha \frac{d\ln f}{dy}(yt_n^\alpha) - \sum_{n\in C} C_n^\alpha \lambda(yC_n^\alpha)$$

and this shows

$$h'_\alpha(y) = 0 \Longleftrightarrow m- \mid C \mid = -\sum_{n\notin C} yt_n^\alpha \frac{d\ln f}{dy}(yt_n^\alpha) + \sum_{n\in C} yC_n^\alpha \lambda(yC_n^\alpha). \quad (5.2.30)$$

One may now wonder under which parametric model the one dimensional optimization problem (5.2.28) has an analytical optimal solution. In the next example we discuss the case of Weibull distributed random variables.

5.2.3 Weibull distribution

Weibull distribution is one of the most accepted distributions in the reliability engineering [29] and this is mainly due to its potential in taking on the characteristics of other distributions depending on the values of its parameters.

Definition 5 *A random variable* $\mathbf{X}$ *has a Weibull cumulative distribution function with location parameter* $\gamma > 0$*, scale parameter* $\beta > 0$ *and shape parameter* $\alpha > 0$ *if*

$$\mathbf{X} \stackrel{d}{=} \gamma + \beta \mathbf{Y}^{\alpha^{-1}} \tag{5.2.31}$$

with the non-negative random variable $\mathbf{Y}$ *having an exponential cdf with parameter* 1*. We denote this by* $\mathbf{X} \sim Weibull(\alpha, \beta, \gamma)$.

For a Weibull cdf we know that the cdf of the random variable $\mathbf{Y}$ is given by $F(y) = 1 - e^{-y}$ with density $f(y) = e^{-y} 1_{[0,\infty)}(y)$. Also its cumulative hazard rate function is given by

$$\Lambda(y) = y. \tag{5.2.32}$$

Introducing for $\alpha > 0$ the Gamma function

$$\Gamma(\alpha) := \int_0^\infty y^{\alpha-1} e^{-y} dy \tag{5.2.33}$$

it is well known that

$$\Gamma(\alpha + 1) = \alpha \Gamma(\alpha) \tag{5.2.34}$$

and

$$\Gamma\left(\frac{1}{2}\right) = \sqrt[2]{\pi}. \tag{5.2.35}$$

This shows for $\mathbf{X} \sim Weibull(\alpha, \beta, \gamma)$ and using relation (5.2.34)

$$\begin{aligned} \mathbb{E}(\mathbf{X}) &= \mathbb{E}(\gamma + \beta \mathbf{Y}^{\alpha^{-1}}) \\ &= \gamma + \beta \mathbb{E}(\mathbf{Y}^{\alpha^{-1}}) \\ &= \gamma + \beta \textstyle\int_0^\infty y^{\alpha^{-1}} e^{-y} dy \\ &= \gamma + \beta \frac{\beta}{\alpha} \Gamma\left(\alpha^{-1}\right) \end{aligned} \tag{5.2.36}$$

Finally we also obtain

$$\begin{aligned} Var(\mathbf{X}) &= Var((\gamma + \beta \mathbf{Y}^{\alpha^{-1}}) \\ &= \beta^2 Var(\mathbf{Y}^{\alpha^{-1}}) \\ &= \beta^2 (\mathbb{E}(\mathbf{Y}^{2\alpha^{-1}}) - \mathbb{E}(\mathbf{Y}^{\alpha^{-1}})^2) \\ &= \beta^2 \left(\Gamma\left(\tfrac{2+\alpha}{\alpha}\right) - \Gamma\left(\tfrac{1+\alpha}{\alpha}\right)^2\right) \end{aligned}$$

5.2.4 MLE optimization for Weibull distribution

We now analyze the log-likelihood optimization problem (P) for the censored data with Weibull distribution as the underlying lifetime distribution with location parameter $\gamma = 0$. By Definition 5 and Lemma 14 we obtain for every $x \geq \gamma$ that the Weibull cdf with location parameter $\gamma > 0$, scale parameter $\beta > 0$ and shape parameter $\alpha > 0$ (take $\theta = (\alpha, \beta, \gamma)$) is given by

$$\begin{aligned} F_\theta(x) &= \mathbb{P}_\theta(\mathbf{X} \leq x) \\ &= 1 - e^{-(\beta^{-1}(x-\gamma))^\alpha} \end{aligned} \tag{5.2.37}$$

with density

$$f_\theta(x) = \alpha\beta^{-1}(\beta^{-1}(x-\gamma))^{\alpha-1}e^{-(\beta^{-1}(x-\gamma))^\alpha}1_{[\gamma,\infty)}(x)$$

Also by Lemma 14 and relation (5.2.32) we obtain

$$\Lambda_\theta(x) := (\beta^{-1}(x-\gamma))^\alpha 1_{[\gamma,\infty)}(x)$$

Since for $\gamma = 0$ we obtain that

$$\ln f(\beta^{-\alpha}t_n^\alpha) = \ln(e^{-(\beta^{-\alpha}t_n^\alpha)}) = -\beta^{-\alpha}t_n^\alpha$$

and

$$\Lambda(\beta^{-\alpha}C_n^\alpha) = \beta^{-\alpha}C_n^\alpha$$

We obtain by Lemma 15 that

$$LL(\alpha,\beta) = \begin{cases} (m- \mid C \mid)[\ln(\alpha) - \alpha\ln(\beta)] + (\alpha-1)\sum_{n\notin C}\ln(t_n) \\ -\beta^{-\alpha}\left(\sum_{n\notin C} {t_n}^\alpha + \sum_{n\in C} {C_n}^\alpha\right) \end{cases} \tag{5.2.38}$$

and we need to solve the optimization problem

$$v(P) = \sup_{\alpha>0,\beta>0} LL(\alpha,\beta). \tag{5.2.39}$$

Before analyzing optimization problem (5.2.39) we rewrite the log-likelihood function and related it to a so-called norm. Let $z \in \mathbb{R}^m$ be any vector and introduce the function $\parallel z \parallel_\alpha : \mathbb{R}_+ \to \mathbb{R}_+$ given by

$$\parallel z \parallel_\alpha = \left(\sum_{n=1}^m \mid z_n \mid^\alpha\right)^{\alpha^{-1}}. \tag{5.2.40}$$

For every $\alpha > 1$ the function $z \to \| z \|_\alpha$ represents the so-called L_α-norm of the vector z. [40]. In this notation using $t_n \geq 0$ and $C_n \geq 0$ we obtain

$$\sum\nolimits_{n \in S} {t_n}^\alpha + \sum\nolimits_{n \notin S} {C_n}^\alpha = \| \bar{t} \|_\alpha^\alpha$$

with the $m-$dimensional vector $\bar{t} = (\bar{t}_1, ..., \bar{t}_m)$ given by

$$\bar{t}_n = \begin{cases} t_n & n \in S \\ \\ C_n & n \notin S. \end{cases} \tag{5.2.41}$$

This shows by relation (5.2.38) that an equivalent representation of the log-likelihood function is given by

$$LL(\alpha, \beta) = (m - \mid C \mid)[\ln(\alpha) - \alpha \ln(\beta)] + (\alpha - 1) \sum\nolimits_{n \notin C} \ln(t_n) - \beta^{-\alpha} \| \bar{t} \|_\alpha^\alpha \tag{5.2.42}$$

To analyze the optimization problem (5.2.39) we first observe the following. Introducing for every shape parameter $\alpha > 0$, the optimization problem

$$\varphi(\alpha) = \sup\nolimits_{\beta > 0} LL(\alpha, \beta) \tag{5.2.43}$$

the optimization problem (5.2.39) is the same as

$$\upsilon(P) = \sup\nolimits_{\alpha > 0} \varphi(\alpha).$$

To show that for each $\alpha > 0$ the optimization problem (5.2.43) has an optimal solution we first observe

$$-\alpha \ln(\beta) = \ln(\beta^{-\alpha}) \text{ and } \lim\nolimits_{\beta \downarrow 0} \frac{\ln(\beta^{-\alpha})}{\beta^{-\alpha}} = 0$$

This implies using relation (5.2.42) that

$$\begin{aligned} & \lim\nolimits_{\beta \downarrow 0} LL(\alpha, \beta) - \mid S \mid \ln(\alpha) - (\alpha - 1) \sum\nolimits_{n \notin C} \ln(t_n) \\ = & \lim\nolimits_{\beta \downarrow 0} \mid S \mid \ln(\beta^{-\alpha}) - \beta^{-\alpha} \| \bar{t} \|_\alpha^\alpha \\ = & -\infty \end{aligned} \tag{5.2.44}$$

and

$$\lim\nolimits_{\beta \uparrow \infty} LL(\alpha, \beta) = -\infty. \tag{5.2.45}$$

By the continuity of the function $\beta \mapsto LL(\alpha, \beta)$ and relations (5.2.44) and (5.2.45) it follows therefore by Weierstrass theorem that for any $\alpha > 0$ optimization problem (5.2.43) has a finite positive optimal solution $\beta(\alpha)$. Actually as shown in the next lemma it is possible to write down an analytical formula for $\beta(\alpha)$ and show that the function φ is strictly concave.

Lemma 17 *It follows for every $\alpha > 0$ that the optimal solution of optimization problem (5.2.43) is unique and given by*

$$\beta(\alpha) = \frac{\| \bar{t} \|_\alpha}{| S |^{\alpha-1}}. \tag{5.2.46}$$

Moreover, the function $\varphi : \mathbb{R}_+ \to \mathbb{R}$ being the optimal objective value of optimization problem (5.2.43) is strictly concave on $\mathbb{R}_+$ and has the form

$$\varphi(\alpha) = \begin{cases} (m- \mid C \mid)[\ln(\alpha) - \alpha \ln(\| \bar{t} \|_\alpha)] \\ \\ +(\alpha - 1) \sum_{n \notin C} \ln(t_n) + \mid S \mid (\ln(\mid S \mid) - 1) \end{cases} \tag{5.2.47}$$

Proof. Let $\alpha > 0$ be given and introduce the function $g_\alpha : \mathbb{R}_+ \to \mathbb{R}$ defined by

$$g_\alpha(\beta) = \beta^{-\alpha}.$$

It is now well known that

$$g'_\alpha(\beta) = -\alpha\beta^{-(1+\alpha)}.$$

Hence we obtain using relation (5.2.42) that

$$\frac{\partial LL}{\partial \beta}(\alpha, \beta) = -\alpha\beta^{-1} \mid S \mid + \alpha\beta^{-(1+\alpha)} \| \bar{t} \|_\alpha^\alpha . \tag{5.2.48}$$

Since we already verified for any $\alpha > 0$ that the optimization problem (5.2.43) has a positive optimal solution $\beta(\alpha)$ it must follow for any $\alpha > 0$ and using $\alpha\beta^{-1} > 0$ for any $\alpha > 0$ and $\beta > 0$ that

$$\begin{aligned} \tfrac{\partial LL}{\partial \beta}(\alpha, \beta(\alpha)) = 0 \quad &\Leftrightarrow \quad \beta(\alpha)^{-\alpha} \| \bar{t} \|_\alpha^\alpha = m- \mid C \mid \\ &\Leftrightarrow \quad \beta(\alpha)^\alpha = \tfrac{\|\bar{t}\|_\alpha^\alpha}{m-|C|} \\ &\Leftrightarrow \quad \beta(\alpha) = \tfrac{\|\bar{t}\|_\alpha}{(m-|C|)^{\alpha-1}}. \end{aligned}$$

This shows for any $\alpha > 0$ that there is only one unique maximizer given by $\beta(\alpha)$ and substituting this into the objective function we obtain the expression for $\varphi(\alpha) = LL(\alpha, \beta(\alpha))$. In [40] it is shown in Theorem B on page 196 that the function $\alpha \to \alpha \ln(\| \bar{t} \|_\alpha)$ is convex. This shows since $\alpha \to \ln(\alpha)$ is a strictly concave function and using relation (5.2.47) that the function φ is strictly concave. $\square$

To finish the MLE estimation procedure for the Weibull case we now have to solve the one dimensional strictly concave maximization problem

$$\max_{\alpha>0} \varphi(\alpha) \tag{5.2.49}$$

with

$$\varphi(\alpha) =\mid S \mid [\ln(\alpha) - \alpha \ln(\| \bar{t} \|_\alpha)] + (\alpha - 1) \sum_{n \notin C} \ln(t_n) + \mid S \mid (\ln(\mid S \mid) - 1)$$

a strictly concave function. It is now sufficient to solve the problem

$$\max_{\alpha>0} \overline{\varphi}(\alpha)$$

with

$$\overline{\varphi}(\alpha) = (m- \mid C \mid)[\ln(\alpha) - \alpha \ln(\| \bar{t} \|_\alpha)] + (\alpha - 1) \sum_{n \notin C} \ln(t_n)$$

a strictly concave function . It follows (see [40]) that for a strictly concave function φ its derivative φ' is strictly decreasing and so the above problem can be solved numerically by finding the unique zero point of the derivative φ' of the function φ using bisection. This also shows that for the Weibull case the MLE optimization problem has a unique optimal solution. The derivative φ' of the function φ is given by

$$\varphi'(\alpha) = (m- \mid C \mid) \left(\alpha^{-1} - \frac{\sum_{n \notin C} \ln(t_n) {t_n}^\alpha + \sum_{n \in C} \ln(C_n) {C_n}^\alpha}{\sum_{n \notin C} {t_n}^\alpha + \sum_{n \in C} {C_n}^\alpha} \right) + \sum_{n \notin C} \ln(t_n).$$

We can now apply the following special purpose algorithm.

Algorithm 4 *Special purpose algorithm for the Weibull case*

1. *Solve the concave maximization problem (5.2.49) by a bisection method applied to the derivative and compute its optimal solution* α_*.
2. *Evaluate* $\beta(\alpha_*) = \frac{\|\bar{t}\|_\alpha}{|S|^{\alpha^{-1}}}$.
3. *Output* $(\alpha_*, \beta(\alpha_*))$.

5.3 Estimation of Weibull parameters

In the methodology part of the chapter, we worked out the Log-likelihood function for the censored lifetimes that are assumed to have Weibull distribution.

In this section this function is optimized (maximized) using the proposed bisection algorithm. Figures 5.8 and 5.9 depict the one-dimensional log-likelihood function for different values of the scale parameter to show the concaveness of the function. After obtaining the optimal shape parameter using Algorithm 4, we acquire the scale parameter.

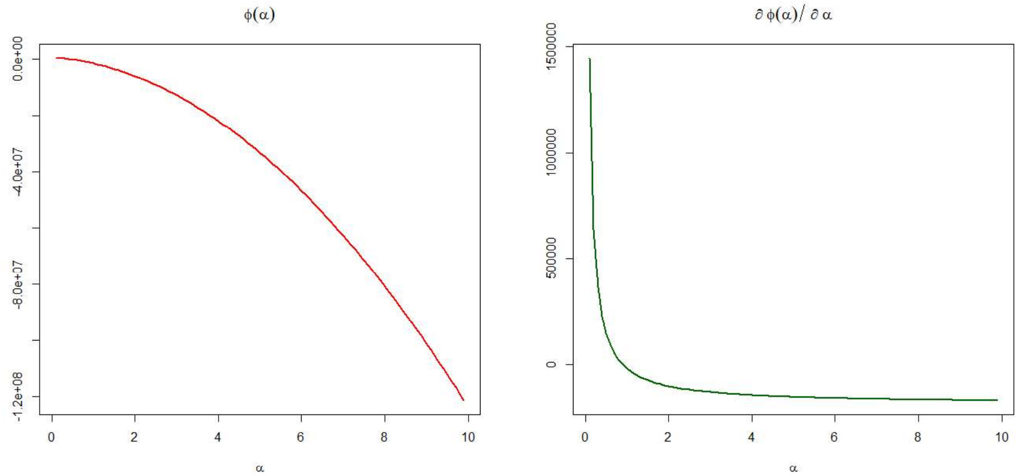


Figure 5.8: One-dimensional log-likelihood function and its gradient (Product A)

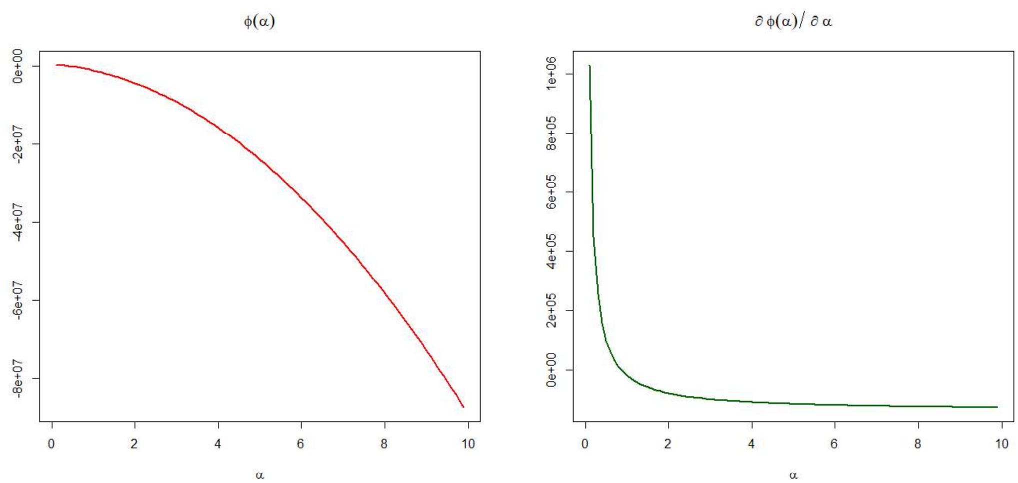
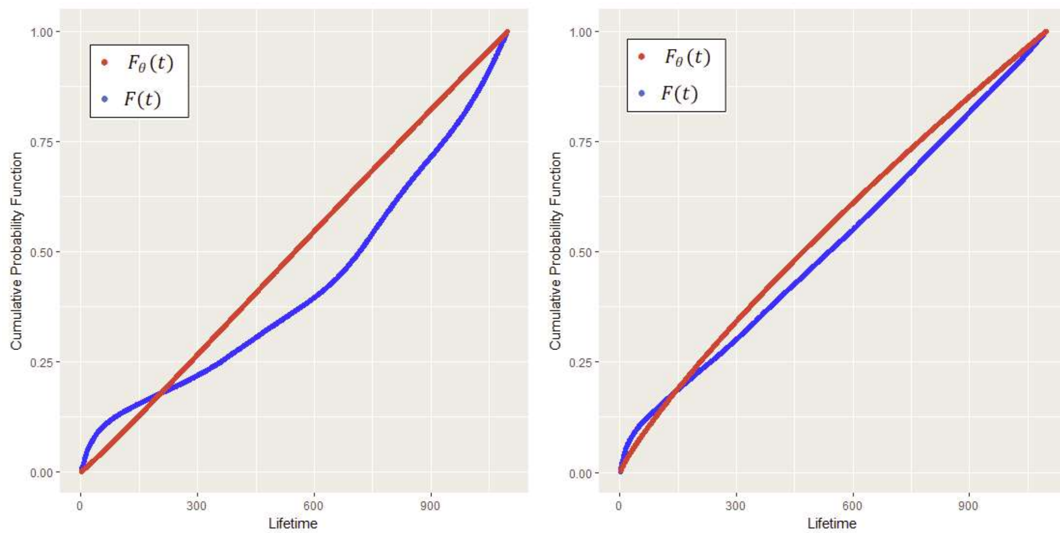


Figure 5.9: One-dimensional log-likelihood function and its gradient (Product B)

Using the provided functions, one can easily find the roots of the gradient function (shape parameter) as $\alpha = 1.07521$, and then from Algorithm 4 obtain the value of scale parameter as $\beta = 5190.92$ for product A. The same process yields $\alpha = 0.85987$ and $\beta = 11250.52$ for product B. It is important to choose the initial interval values properly in this approach. Otherwise, the bisecting rule will encounter errors.

Finally, we can assess the goodness-of-fit of Weibull distribution by means of Kolmogorov-Smirnov test which is based on inspecting the distance between the empirical distribution function (EDF). Figures (5.10a) and (5.10b) depict the empirical cdf of the data plotted together with Weibull(α, β) cdf fitted over our lifetime data.



(a) (b)

Figure 5.10: Empirical cdf and Weibull distribution with the estimated parameters for product A (a) and product B (b)

The Kolmogorov-Smirnov test statistic is : $D = Sup_x|F_\theta(x) - F(x)|$ where $F_\theta(x)$ is the cdf of the Weibull distribution with estimated α and β and $F(x)$ is the empirical distribution function of observed data. The test statistic and the corresponding values respectively 0.16 and 0.003 for product A and 0.06 and 0.004 for product B which means that Weibull distribution did not perform well in both product types. This can also be verified from the plotted $F_\theta(x)$ and $F(x)$. For the case of product A (Figure 5.10a), the EDF is higher than the CDF of the Weibull distribution for the lifetimes below 200 days and then it declines to the below of estimated Weibull CDF. This is mainly due to the fact that the distribution of lifetimes for product A has a multiplicative seasonality (See Figure 5.4). Also Weibull distribution is not able to explain the bathtub shape properly.

Chapter 6

CONCLUSION AND DISCUSSION

Predicting the number of items that will fail within a specific time interval in the future is of paramount importance for the manufacturers since this would help them maintain proper spare parts inventory and effective repair resource management. In this thesis, we aimed at providing a model for this. We succeeded in developing two stochastic sub-models to estimate the number of items sold up to a certain point in time, and the lifetimes of these items. To predict the number of failures, these two sub-models can be used separately and their results can be combined using numerical and/or computational methods. Combining the two sub-models analytically requires solving a difficult numerical integration problem, which we plan to address in future research.

We were fortunate to have access to a large dataset on the sale and failure times of two products. To provide a model that best explains this data, we used the MLE method assuming the lifetime data to follow a Weibull distribution. While there are other popular parametric classes of failure time distributions such as the lognormal, logistic and Gamma, these are not covered within this thesis partly due to the lack of a proper algorithm that can guarantee the optimal solution of our log-likelihood optimization problem. In addition, there are other approaches to estimate the cumulative cdf [29, 22]. Finally and of more importance is to test whether our failure data set satisfy the minimal repair assumption such that the counting process of failures can be represented by a non-homogeneous Poisson process. Statistical methods to test this assumption can be found in [18]. In addition, if we can obtain data on the costs of different repair types, we can test our cost estimation model as well. All of these suggest topics for future research.

Appendix A

In this Appendix we list the following well-known result for a probability Laplace-Stieltjes transform of a non-negative random variable $\mathbf{X}$. For completeness a proof of this result is listed. Observe $f^{(0)} := f$ and $f^{(n)}, n \in \mathbb{N}$ denotes the nth derivative of the function f.

Lemma 18 *If the random variable* $\mathbf{X}$ *is non-negative with cdf* F *satisfying* $F(\infty) := \lim_{x\uparrow\infty} F(x) = 1$ *then its pLSt* $\pi_{\mathbf{X}} : [0,\infty) \to [0,1]$ *is a continuous function on* $(0,\infty)$ *and* $\pi_{\mathbf{X}}(0) = 1$*. Also all its derivatives exists on* $(0,\infty)$ *and its* k*th derivative,* $k \in \mathbb{N}$ *is given by*

$$\pi_{\mathbf{X}}^{(k)}(s) = (-1)^n \mathbb{E}(\mathbf{X}^k e^{-s\mathbf{X}}), s > 0 \tag{A.1}$$

and

$$(-1)^n \pi_{\mathbf{X}}^{(k)}(0^+) := (-1)^n \lim_{s\downarrow 0} \pi_{\mathbf{X}}^{(k)}(s) = \mathbb{E}(\mathbf{X}^k) \leq \infty. \tag{A.2}$$

Proof. Since $F(\infty) = 1$ it is obvious that $\pi_{\mathbf{X}}(0) = 1$. Also by the monotone convergence theorem it follows that the pLSt function $\pi_{\mathbf{X}}$ is continuous on $(0,\infty)$ satisfying $\pi_{\mathbf{X}}(0^+) = \pi_{\mathbf{X}}(0)$. To show the existence of the derivatives we first verify the result for $k = 1$. It follows for any $n \in \mathbb{N}$ and $s > 0$ that

$$n(\pi_{\mathbf{X}}(s) - \pi_{\mathbf{X}}(s + n^{-1})) = n(\mathbb{E}(e^{-s\mathbf{X}}) - \mathbb{E}(e^{-(s+n^{-1})\mathbf{X}})) = \mathbb{E}(\mathbf{Y}_n)$$

with $\mathbf{Y}_n := ne^{-s\mathbf{X}}(1 - e^{-n^{-1}\mathbf{X}}), n \in \mathbb{N}$. Since for every $x \geq 0$

$$0 \leq n(1 - e^{-n^{-1}x}) = \int_0^x e^{-n^{-1}s} ds$$

it follows that the function $n \to ne^{-sx}(1 - e^{-n^{-1}x})$ is increasing for every $x \geq 0$. Hence the sequence of non-negative random variables $\mathbf{Y}_n$ satisfy $\mathbf{Y}_{n+1} \geq \mathbf{Y}_n$ for every $n \in \mathbb{N}$ and its limit random variable $\mathbf{Y}_\infty$ is given by

$$\mathbf{Y}_\infty = \lim_{n\uparrow\infty} \mathbf{Y}_n \uparrow \mathbf{X}e^{-s\mathbf{X}}.$$

So the conditions of the monotone convergence theorem are satisfied and we may conclude that

$$-\pi_{\mathbf{X}}^{(1)}(s) = \lim_{n\uparrow\infty} n(\pi_{\mathbf{X}}(s) - \pi_{\mathbf{X}}(s + n^{-1})) = \mathbb{E}(\mathbf{X}e^{-s\mathbf{X}}).$$

Since for every $s > 0$ the function $x \to xe^{-sx}$ has a finite upper bound on $[0, \infty]$ it follows that $\mathbb{E}(\mathbf{X}e^{-s\mathbf{X}})$ is finite. Applying now a standard induction and applying the same arguments yields the formula in relation (A.1). Again by the monotone convergence theorem using $\mathbf{Y}_n = \mathbf{X}^k e^{-n^{-1}\mathbf{X}}, k \in \mathbb{N}$ is an increasing sequence of random variables with limit

$$\mathbf{Y}_\infty = \lim_{n \uparrow \infty} \mathbf{Y}_n \uparrow \mathbf{X}^k$$

we obtain from relation (A.1) the result in relation (A.2). □

Appendix B

In this appendix we discuss part of the theory of the linear regression model. We start by making the following basic assumption. Let y be the endogenous variable and x be the k-dimensional vector of exogeneous variables and assume

$$\mathbf{y} = X\boldsymbol{\theta} + \boldsymbol{\epsilon}$$

for some $\boldsymbol{\theta} \in \mathbb{R}^k$ and $\boldsymbol{\epsilon}$ a k-dimensional vector of independent and identically distributed random variables with mean zero and unknown variance σ^2. To estimate the parameter $\boldsymbol{\theta}$, we use the mean square error function, and need to solve the optimization problem

$$\inf_{\theta \in \mathbb{R}^k} \parallel \mathbf{y} - X\boldsymbol{\theta} \parallel^2 = \inf_{\boldsymbol{\theta} \in \mathbb{R}^k} (\mathbf{y} - X\boldsymbol{\theta})^\intercal (\mathbf{y} - X\boldsymbol{\theta})$$

with $\parallel . \parallel$ denoting the Euclidean norm. Since

$$f(\boldsymbol{\theta}) = (\mathbf{y} - X\boldsymbol{\theta})^\intercal (\mathbf{y} - X\boldsymbol{\theta}) = \mathbf{y}^\intercal \mathbf{y} - 2\boldsymbol{\theta}^\intercal X \mathbf{y} + \boldsymbol{\theta}^\intercal X^\intercal X \boldsymbol{\theta}$$

and the gradient is given by

$$\nabla f(\boldsymbol{\theta}) = -2^\intercal X^\intercal \mathbf{y} + 2X^\intercal X \boldsymbol{\theta}$$

it follows that the optimal solution $\widehat{\boldsymbol{\theta}}$ satisfies the so-called first order conditions

$$\nabla f(\widehat{\boldsymbol{\theta}}) = \mathbf{0} \Longleftrightarrow X^\intercal X \widehat{\boldsymbol{\theta}} = X^\intercal \mathbf{y}$$

If the column rank of the $n \times k$ matrix equals k one can show that the $k \times k$ matrix $X^\intercal X$ has an inverse and so if this holds it follows that

$$\widehat{\boldsymbol{\theta}} = (X^\intercal X)^{-1} X^\intercal \mathbf{y}$$

Denoting now by $\widehat{\mathrm{y}}$ the part of y explained by the exogeneous variables defined by

$$\widehat{\mathbf{y}} = X\widehat{\boldsymbol{\theta}} = H\mathbf{y} \tag{B.1}$$

with $H := X(X^\intercal X)^{-1} X^\intercal$ the so-called hat matrix (by relation B.1 it transforms y to $\widehat{\mathrm{y}}$ and this creates the name). This matrix is clearly symmetric and it satisfies

$$HX = X \tag{B.2}$$

and it is idempotent. This means

$$H^2 = HH = X(X^\top X)^{-1}X^\top X(X^\top X)^{-1}X^\top = H. \tag{B.3}$$

Denoting now by e the unexplained part of the vector y given by

$$\mathbf{e} := \mathbf{y} - \widehat{\mathbf{y}}$$

it follows by relation (B.1) and (B.2) that

$$\mathbf{e} = \mathbf{y} - \widehat{\mathbf{y}} = (I - H)\mathbf{y} = (I - H)(X\boldsymbol{\theta} + \boldsymbol{\epsilon}) = X\boldsymbol{\theta} + \boldsymbol{\epsilon} - HX\boldsymbol{\theta} + H\boldsymbol{\epsilon} = (I - H)\boldsymbol{\epsilon} \tag{B.4}$$

Since by relation (B.3) we obtain

$$(I - H)^2 = (I - H)(I - H) = I - 2H + H^2 = I - H$$

the symmetric and idempotent matrix $I - H$ is a projection and by the previous relation this projection projects ϵ to e. By relation (B.4) and (B.2) we obtain

$$\widehat{\mathbf{y}}^\top \mathbf{e} = \widehat{\boldsymbol{\theta}}^\top X^\top \mathbf{e} = \widehat{\boldsymbol{\theta}}^\top X^\top (I - H)\boldsymbol{\epsilon} = 0. \tag{B.5}$$

This means that the explained part $\widehat{\mathbf{y}}$ of y is orthogonal on the nonexplained part of y. Now it follows by relation (B.5) that

$$\mathbf{y}^\top \mathbf{y} = (\widehat{\mathbf{y}} + \mathbf{e})^\top (\widehat{\mathbf{y}} + \mathbf{e}) = \widehat{\mathbf{y}}^\top \widehat{\mathbf{y}} + 2\widehat{\mathbf{y}}^\top \mathbf{e} + \mathbf{e}^\top \mathbf{e} = \widehat{\mathbf{y}}^\top \widehat{\mathbf{y}} + \mathbf{e}^\top \mathbf{e}$$

Similarly we can derive with

$$\overline{\mathbf{y}} := \frac{1}{n}\mathbf{y}^\top \mathbf{i}$$

with i denoting the vector of only ones and so $\overline{\mathbf{y}}$ denotes the average value and

$$\overline{\widehat{\mathbf{y}}} := \frac{1}{n}\widehat{\mathbf{y}}^\top \mathbf{i}$$

the average estimated value and using that the first column of the matrix X is given by i we obtain by relation (B.2) that

$$H^\top \mathbf{i} = \mathbf{i}.$$

This implies

$$\overline{\widehat{\mathbf{y}}} = \frac{1}{n}\widehat{\mathbf{y}}^\top \mathbf{i} = \frac{1}{n}(H\mathbf{y})^\top \mathbf{i} = \frac{1}{n}\mathbf{y}H^\top \mathbf{i} = \frac{1}{n}\mathbf{y}^\top \mathbf{i} = \overline{\mathbf{y}}.$$

Hence it follows

$$\sum_{i=1}^{n}(y_i - \overline{\mathbf{y}})^2 = \mathbf{y}^\intercal\mathbf{y} - n\overline{\mathbf{y}} = \widehat{\mathbf{y}}^\intercal\widehat{\mathbf{y}} + \mathbf{e}^\intercal\mathbf{e} - n\,\overline{\widehat{\mathbf{y}}} = \sum_{i=1}^{n}\left(\widehat{y}_i - \overline{\widehat{\mathbf{y}}}\right)^2 + \mathbf{e}^\intercal\mathbf{e}$$

The above relation relates the total variation of the original sample and the estimatıon and the error as follows

$$SS_{total} = SS_{explained} + SS_{error}$$

Now we define

$$R^2 = \frac{SS_{explained}}{SS_{total}}$$

BIBLIOGRAPGY

[1] ANDERSEN, P. K. Filtered renewal processess with a two-sided impact function. *Journal of Applied Probability 16*, 4 (1979), 813–821.

[2] APPLEBAUM, D. Lévy processes and stochastic calculus. *Cambridge university press* (2009).

[3] ASMUSSEN, S. *Applied Probability and Queues (second edition)*. Wiley, New York, 1987.

[4] AVEN, T., AND JENSEN, U. *Stochastic models in Reliability*. Springer Verlag, New York, 1998.

[5] AVEN, T., AND JENSEN, U. A general minimal repair model. *Journal of Applied Probability 37*, 1 (2000), 187–197.

[6] BASAK, P., BASAK, I., AND BALAKRISHNA, N. Estimation for the three-parameter lognormal distribution based on progressively censored data. *Computational Statistics and Data Analysis 53*, 10 (2009), 3580–3592.

[7] BISHOP, C. M. *Pattern Recognition and Machine Learning*. Springer Verlag, New York, 2006.

[8] BLOCK, H. W., BORGES, W. S., AND SAVITS, T. H. Age dependent minimal repair. *Journal of Applied Probability 22* (1985), 370–385.

[9] BOX, G. E., JENKINS, G. M., REINSEL, G. C., AND LJUNG, G. M. Time series analysis: forecasting and control. *Holden-Day* (1976).

[10] BREIMAN, L. *Probability*. Addison Wesley, Menlo Park, 1968.

[11] BRÉMAUD, P. *Point Processes and Queues*. Springer Verlag, New York, 1981.

[12] BROCKHOFF, K. A test for the product life cycle. *Econometrica 35*, 3-4 (1967), 472–484.

[13] ÇINLAR, E. *Introduction to Stochastic Processes*. Prentice-Hall, 1975.

[14] Çinlar, E. *Probability and Stochastics*. Springer Verlag, New York, 2011.

[15] Cha, J. H., and Finkelstein, M. Stochastic intensity for minimal repairs in heterogeneous systems. *Journal of Applied Probability 3* (2011), 868–876.

[16] Cheney, E. W., and Kincaid, D. R. *Numerical Mathematics and Computing (Edition 4)*. Brooks -Cole Publishing Company, Pacific Grove, 1999.

[17] Clottey, T., Benton Jr, W. C., and Srivastava, R. Forecasting product returns for remanufacturing operations. *Decision Sciences 43*, 4 (2012), 589–614.

[18] Cook, R. J., and Lawless, J. *The Statistical Analysis of Recurrent Events*. Springer Verlag, New York, 2007.

[19] Goh, T. N., and Varaprasad, N. A statistical methodology for the analysis of the life-cycle of reusable containers. *IIE transactions 18*, 1 (1986), 42–47.

[20] Hess, J. D., and Mayhew, G. E. Modeling merchandise returns in direct marketing. *Journal of Interactive Marketing 11*, 2 (1997), 20–35.

[21] Hosseini, M. M., Kerr, R. M., and Randall, R. B. An inspection model with minimal and major maintenance for a system with deterioration and poisson failures. *IEEE Transactions on Reliability 49*, 1 (2000), 88–98.

[22] Kalbfleisch, J. D., and Prentice, R. L. *The Statistical Analysis of Failure Time Data*. Wiley, New York, 2002.

[23] Karlin, S., and Taylor, H. E. *A First Course in Stochastic Processes*. Academic Press, New York, 1975.

[24] Keilson, J. *Markov Chain models-Rarity and Exponentiality*. Springer Verlag, New York, 1979.

[25] Kelle, P., and Silver, E. A. Forecasting the returns of reusable containers. *Journal of Operations Management 8*, 1 (1989), 17–35.

[26] Kim, T. Y., Dekker, R., and Heij, C. Spare part demand forecasting for consumer goods using installed base information. *Computers and Industrial Engineering 103* (2017), 201–215.

[27] Kuş, C., and Kaya, M. F. Estimation of parameters of the loglogistic distribution based on progressive censoring using the em algorithm. *Hacettepe Journal of Mathematics and Statistics 35*, 2 (2006).

[28] LAST, G., AND BRANDT, A.

[29] LAWLESS, J. J. *Statistical models and Methods for lifetime data (second edition)*. Wiley, New York, 2003.

[30] LEEMIS, L. M. Nonparametric estimation of the cumulative intensity function for a nonhomogeneous poisson process. *Management Science 37*, 7 (1991), 886–899.

[31] LINZ, P. *Analytical and Numerical Methods for Volterra Equations*. SIAM, Philadelphia, 1985.

[32] MADDALA, G. *Econometrics*. Mc-Graw Hill, Tokyo, 1977.

[33] MARX-GOMEZ, J., RAUTENSTRAUCH, C., NÜRNBERGER, A., AND KRUSE, R. Neuro-fuzzy approach to forecast returns of scrapped products to recycling and manufacturing. *Knowledge-Based Systems 15* (2002), 119–128.

[34] MENTZER, J. T., AND MOON, M. A. *Sales forecasting management: a demand management approach*. Sage, 2001.

[35] MINNER, S. Forcasting and inventory management for spare parts: An installed base approach: Chapter 8 service parts management demand forecasting and inventory control. *Springer Verlag* (2011), 157–169.

[36] MUNAKATA, S., TEZUKA, M., IIZUKA, S., AND URABE, S. Demand forecasting model of service parts with different failure rate. *TENCON 2010-2010 IEEE Region 10 Conference* (2010), 1427–1431.

[37] NICULESCU, . P. *On Banch Algebras, Renewal measures and Regenerative Processess*. CWI, Amsterdam, 1989.

[38] PARZEN, E. *Stochastic Processes*. Holden-Day, 1962.

[39] P.LANCASTER, AND M.TISMENETSKY. *The Theory of Matrices*. Computer Science and Applied Mathematics. Academic Press, 1985.

[40] ROBERTS, A. W. *Convex Functions*. Academic Press, 1973.

[41] ROSS, S. M. *Applied Probability Models with Optimization Applications*. Holden-Day, San Francisco, 1970.

[42] SCHOLZ, F. W., AND WORKS, B. P. Maximum likelihood estimation for type i censored weibull data including covariates. *Boeing Information and Support Services* (1996).

[43] Shorrock, R. A limit theorem for inter-record times. *Journal of Applied Probability 9* (1972), 219–223.

[44] Snyder, R. Forecasting sales of slow and fast moving inventories. *European Journal of Operational Research 140*, 3 (2002), 684–699.

[45] Steutel, F. W., and Van Harn, K. *Infinite divisibility of Probability Distributions on the Real Line*. Marcel Dekker, New York, 2004.

[46] Stummel, F., Hainer, K., and Everitt, W. N. *Introduction to Numerical Analysis*. Scottish Academic Press, 1980.

[47] Tijms, H. C. *Stochastic Models :An Algortihmic Approach*. Wiley, New York, 1995.

[48] Tijms, H. C. *A First course in Stochastic Models*. Wiley, New York, 2003.

[49] Toktay, L. B., van der Laan, E. A., and de Brito, M. P. Managing product returns: the role of forecasting. *Reverse Logistics* (2004), 45–64.

[50] Toktay, L. B., Wein, L. M., and Zenios, S. A. Inventory management of remanufacturable products. *Management Science 46*, 11 (2000), 1412–1426.

[51] Widder, D. V. *The Laplace Transform*. Princeton Mathematical Series. Princeton University Press, 1972.

[52] Xie, M., Tang, Y., and Goh, T. N. A modified weibull extension with bathtub-shaped failure rate function. *Reliability Engineering and System Safety 76*, 3 (2002), 279–285.