CUKUROVA UNIVERSITY

INSTITUTE OF NATURAL AND APPLIED SCIENCES

PhD THESIS

Parameter Estimation of Birnbaum-Saunders Distribution with Genetic Algorithm under Right Censored Data.

Ali ASSOUMANI RASSOUL

Statistics Department

September, 2023

CUKUROVA UNIVERSITY

INSTITUTE OF NATURAL AND APPLIED SCIENCES

PhD THESIS APPROVAL

Parameter Estimation of Birnbaum-Saunders Distribution with Genetic Algorithm under Right Censored Data.

Ali ASSOUMANI RASSOUL

Statistics Department

This Doctorate Thesis was evaluated by the following Jury Members on 12/09/2023 and was approved by unanimity / majority of votes.

Jury : Prof. Dr. Güzin YÜKSEL (Advisor) ……

: Prof. Dr. Ali İhsan GENÇ ……

: Prof. Dr. Gülsen KIRAL ……

: Prof. Dr. Nazif ÇALIŞ ……

: Asst. Prof. Dr. Murat GENÇ ……

This Thesis was written in the Department of Statistics, Institute of Natural and Applied Sciences.

Thesis Number:

Prof. Dr. Sadık DİNÇER
Director
Institute of Natural and Applied Sciences

This work was supported by the YTB.
Project ID: 17KM-0000-67

CONTENTS

CUKUROVA UNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCES

PhD THESIS

Parameter Estimation of Birnbaum-Saunders Distribution with Genetic Algorithm under Right Censored Data.

Ali ASSOUMANI RASSOUL

Advisor: Prof. Dr. Güzin YÜKSEL

Department of Statistics

ABSTRACT

The Birnbaum-Saunders (BS) distribution is a common reliability distribution used in scientific studies. There have been studies in the literature on parameter estimates for this distribution. In addition, in many studies, it is recommended to use genetic algorithm (GA) optimization methods for parameter estimation in modelling. This thesis focuses on the analysis and estimation of model parameters for a two-parameter Birnbaum-Saunders distribution for right-censored reliability data. For the estimation of Birnbaum-Saunders distribution parameters, we propose the genetic algorithm (GA) method as an alternative to the maximum likelihood estimation (ML) method. Psi31 data are often used as an example to show the limitations of prediction methods when using censored data. In addition, the performance of the ML and GA methods were studied by Monte Carlo simulation with different sample sizes and censorship rates.

Keywords: Birnbaum-Saunders distribution, Genetic algorithm, Maximum likelihood, Right censored data.

ÇUKUROVA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

DOKTORA TEZİ

Sağ Sansürlü Veriler Altında Birnbaum-Saunders Genetik Algoritma ile Dağılımının Parametre Tahmini.

Ali ASSOUMANI RASSOUL

Danışman: Prof. Dr. Güzin YÜKSEL

İstatistik Anabilim Dalı

ÖZ

Birnbaum-Saunders (BS) dağılımı bilimsel çalışmalarda kullanılan yaygın bir güvenilirlik dağılımıdır. Literatürde bu dağılım için parametre tahminleri üzerine çalışmalar yapılmıştır. Ayrıca birçok çalışmada modellemede parametre tahmini için genetik algoritma (GA) optimizasyon yöntemlerinin kullanılması önerilmektedir. Bu tez, sağ sansürlü güvenilirlik verileri kullanak iki parametreli Birnbaum-Saunders dağılımında model parametrelerinin analizi ve tahminine odaklanmaktadır. Birnbaum-Saunders dağılım parametrelerinin tahmini için, maksimum olabilirlik tahmini (ML) yöntemine alternatif olarak genetik algoritma (GA) yöntemini öneriyoruz. Psi31 verileri, sansürlü veriler kullanıldığında tahmin yöntemlerinin sınırlamalarını göstermek için sıklıkla bir örnek olarak kullanılır. Ayrıca, ML ve GA yöntemlerinin performansı, farklı örneklem büyüklükleri ve sansür oranları ile Monte Carlo simülasyonu ile incelenmiştir.

Anahtar Kelimeler: Birnbaum-Saunders dağılımı, Genetik algoritma, Maksimum olasılık, Sağdan sansürlü veriler.

GENİŞLETİLMİŞ ÖZET

Yorulma ömrü (fatigue life) dağılımı olarak da adlandırılan Birnbaum-Saunders dağılımı, başarısızlık sürelerini modellemek için güvenilirlik uygulamalarında yaygın olarak kullanılan bir olasılık dağılımıdır. Bu dağılım, ilk olarak Birnbaum ve Saunders (1969) tarafından yapılan çalışma ile istatistik literatürüne kazandırılmıştır. Makinalardaki metalik parçaların yüzeylerinde oluşan kırılmalardan ve çatlaklardan dolayı gerçekleşen arızalar problem yarattığı için bu arızaların modellenmesine ihtiyaç duyulmuş ve Birnbaum Saunders dağılımı önerilmiştir. Başarısızlık sürelerini modellemek için güvenilirlik uygulamalarında da yaygın olarak kullanılan bir olasılık dağılımıdır. Günümüzde iş dünyasında kullanılan makinaların parçalarının yorulma ve bozulma ömürleri ile dayanıklılıklarının modellenmesi konusunda tercih edilen dağılımlardan biridir. Mühendislik alanlarının dışında işletme, ekonomi, tarım, tıp, hava kirliliği gibi alanlarda da kullanımına rastlanır. Bu sebeplerden dolayı parametrelerinin tahminin etkin bir şekilde yapılması önem arz etmektedir.

Yaygın olarak kullanılan istatistiksel dağılım olan Normal dağılım üzerinde bazı dönüşümler yapılarak çeşitli dağılımlar geliştirilmiştir. Literatürde bu dağılımın birkaç alternatif formülasyonuincelenmiştir. Bunlardan biri olan. iki parametreli Birnbaum-Saunders(BS) dağılımı, standart normal dağılım üzerinde bir monoton dönüşüm kullanılarak geliştirilen bu tür dağılımlardan biridir. Bu dağılım Kundu ve ark (2010) tarafından literature kazandırılmıştır..

Hayatta kalma verileri olarak da bilinen sansürlü veri analizinde, bir olayın zaman içinde gerçekleşmesi için geçen süre ile ilgilenilir. Gerçekten de, hayatta kalma verileri sadece biyoistatistikçilerin alanı değildir. Aynı zamanda güvenilirlik, ekonomi, sigortacılık, psikoloji ve pek çok başka alanda da kullanımı yaygındır. Endüstriyel güvenilirlikte, ilgilenilen olaylar, örneğin bir sistemin bileşenlerinin yaşam süresidir. Diğer alanlarda da bu konu ile ilgili uygulamalardan bahsedecek olursak; ekonomistler işsizlik dönemlerinin süreleriyle ilgilenirken, sigortacılar bir talebin geldiği an ile, psikologlar ise bir öznenin belirli bir görevi yerine getirmesi için geçen süre ile ilgilenirler. Biyomedikal alanda, hayatta kalma verileri esas olarak, tedavi edici deneylerdeki klinik araştırmalarda ve epidemiyolojideki kohort çalışmalarında bulunur. İlk durumda, iki tedavinin etkililiğini ilgili bir olaya (nüksetme, ölüm...) zaman kriteriyle karşılaştırmak genellikle gereklidir; ikinci durumda, takip süresi boyunca hastalığın ortaya çıkabileceği veya çıkmayabileceği bir denek kohortu zaman içinde takip edilir ve bu hastalık üzerindeki risk faktörlerinin önemi değerlendirilmeye çalışılır.

Hayatta kalma sürelerinin belirleyiciliği, pozitif rastgele değişkenlere karşılık gelmesi ve eksik gözlemlere sahip olmasından gelir. Aslında, ileriye dönük çalışmaların çoğunda, bireyler önceden belirlenmiş bir gözlem süresi boyunca izlenir. İlgilenilen olayın gözlem periyodu sırasında meydana geldiği denekler için, bu ilgilenilen olayın tam ortaya çıkış zamanını bilebiliriz ve bu, açık bir şekilde belirtilmesi gereken bir başlangıç tarihinden itibaren ölçülür (örneğin, klinik deneylerde

randomizasyon tarihi) . Ancak gözlem süresinin sonunda bazı denekler olayı yaşamamış olacaktır. O zaman bu denekler için sadece eksik bilgiye sahip olacağız. Daha açık bir şekilde ifade edersek, olayın meydana gelme zamanının gözlem zamanından daha uzun olduğu durum söz konusu olacaktır. Bu tür verilere sağdan sansürlü adı verilir ve gözlem süresi sansür dönemini oluşturur. Denekler çalışmaya farklı tarihlerde girdiğinden, her deneğin kendi zaman çerçevesi ve sansür süresi vardır. Diğer sansür türleri ile karşılaşılabilir. Örneğin olayın meydana gelme zamanının yalnızca bir üst sınırı biliniyorsa, verilerin soldan sansürlü olduğu durum söz konusudur. Ya da bir olayın yalnızca bilinen iki tarih arasında meydana geldiği biliniyorsa (örneğin, iki hasta ziyareti arasında bir biyolojik parametrenin kritik eşiğinin geçmesi) aralıklı sansürlü veri söz konusudur. Sansürlü verilerin analizinde sansür döneminde yer alan bilgileri dikkate alan metodoloji kullanılır. Sansürü kullanan olağan prosedürlerde, neredeyse her zaman hayatta kalma süreleri ve sansür süreleri değişkenlerinin bağımsızlığı varsayılır.

Başarısız olan deneklerin ömür dağılımlarıyla ilgili istatistiksel ve olasılıksal problemler, güvenilirlik teorisine yer alır. Bu teoriye dayanarak, rastgele ömür değişkenleri aracılığıyla nesnelerin performansı ve bozulması olasılıksal yöntemlerle açıklanır. Bu elemanların güvenilirlik (veya hayatta kalma) fonksiyonu, başarısızlık oranı (veya risk) ve ortalama ömürlerinin belirlenmesi için ömür dağılımları kullanılır. Bu teorinin istatistiksel yönleri, yaşam dağılımı parametrelerinin tahmin ve çıkarım yönlerini çözmek için kullanılır. Kısaca, bir olayın güvenilirliği, rastgele bir yaşam değişkeni olan T > 0'ın sabit bir t süresini aşma olasılığıdır. Güvenirlik hesaplamalarında dağılımların parametrelerinin tahmininde birçok yöntemden yararlanılır. Bunlardan bir de ML tahmin yöntemidir. Ancak Genetik Algoritma yönteminin de bazı dağılımların parameter tahminin de kullanıldığı literatürde görülmüştür.

Genetik Algoritma yönteminin (GA) problem çözmek için kullanımı, genetik algoritma yöntemini öneren John Holland (1960) ve Michigan Üniversitesi'ndeki meslektaşları ve öğrencileri tarafından yapılan araştırmalardan günümüze kadar gelmiştir. Bu yeni araştırma grubunun literatüre eklediği yenilik, çaprazlama operatörünün mutasyon modellerine eklenmesiydi. Bu operatör genellikle popülasyondaki farklı bireylerin genlerini birleştirerek bir fonksiyonun optimumunu bulmak için kullanılır. Bu araştırmanın ilk sonucu Holland, J.H. (1975) tarafından “Adaptation in natural and artificial systems” adlı yayının yayınlanması ile açıklanmıştır.. GA'lar genellikle makine öğrenimi ve optimizasyon uygulamalarında kullanılır. GA, birçok olası çözüm arasından en iyi çözümü bulmak için tasarlanmıştır. Genel olarak algoritmaya, her biri bir kromozoma karşılık gelen karakter dizilerinden oluşan bir popülasyonla başlanır. Genetik algoritmalarda, belirli bir popülasyonda depolanan genler, türün çevresel ihtiyaçlarına en uygun olanıdır. Aslında, belirli gen varyasyonları, onlara sahip olan bireylere, popülasyonun geri kalanı üzerinde rekabet avantajı sağlar. Bu rekabet avantajı, birbirini takip eden nesillerde, tüm popülasyona aktarılabilmiş olan bu bireylerin daha iyi üremesiyle sonuçlanır. Genetik algoritmaların prensibi, hangisinin en iyi sonuçları verdiğini görmek için farklı işlemleri

denemektir. Genetik algoritmalar, problemleri etkili bir şekilde çözmek için kullanılabilir. Bununla birlikte, bu yöntemlerin kullanımında, sorun belirli özelliklerine göre koşullandırılmalıdır. Değerlendirmede uygunluk fonksiyonunun hesaplama süresi oldukça kısa olmalıdır.

Evrim teorisine dayanan Genetik Algoritma (GA), stokastik optimizasyon uygulamalarında yaygın olarak kullanılmaktadır. Algoritma, her biri bir gen (vektörün bir bileşeni) ile karakterize edilen N birey popülasyonu ("kromozomlar" olarak da bilinir) üzerinde yürütülür. İşlevsel olarak en üretken bireyi belirlemek, çeşitli faktörlerin dikkatle değerlendirilmesini gerektiren karmaşık bir görevdir. Bireylerin nesilleri birbirini takip eder, en üretken üyeler bir sınıra ulaşılana kadar üretmeye devam eder. Bir nesilden diğerine geçiş, "operatörler" aracılığıyla gerçekleştirilir.

Son çalışmalar, GA yönteminin istatistiksel modellerdeki parametreleri tahmin etmek için yararlı bir araç olduğunu göstermiştir.Bu konu ile ilgili, Lee ve ark.(2006) çalışmalarında bir orta ölçekli meteorolojik modelde iki parametre için bir GA yaklaşımını kullandılar.

Genetik algoritmanın parameter tahmininde kullanımı ile ilgili çalışmalardan biri de Tekeli ve Yüksel (2020) makalesidir. Bu çalışmada iki katlı Weibull karma modelinin parametrelerini tahmin etmek için sağdan sansürlü veri ile genetik algoritmayı kullandılar. Aynı yıl, Akay ve ark.(2020), GA yaklaşımını kullanarak yeni uygunluk fonksiyonlu optimal kümelemeyi önerdiler.

Bu nedenle bu tez çalışmasında da Genetik Algoritma yönteminin Birnbaum Saunders dağılımında kullanımı üzerinde yoğunlaşılmıştır.

Bu tez, giriş bölümü dahil 7 bölümden oluşmaktadır. Bölüm 2'de Birnbaum-Saunders, sansürlü veri ve genetik algoritma hakkında bilgiler bulunmaktadır. Bölüm 3'te, bir olayın belli bir süre içinde meydana geldiği zamana bakarak sansürlü veriler ve sansürlü veri türleri üzerine bilgiler bulunur. Bununla birlikte, 4. bölümde Birnbaum-Saunders dağılımının ortaya çıkışı ve matematiksel türevinin ayrıntılı bir incelemesiniyer almaktadır. Ek olarak, olasılık yoğunluğu, kümülatif dağılım ve kuantum fonksiyonları gibi olasılık özelliklerinin daha kapsamlı bir detayı sunulmaktadır. Bölüm 5, Birnbaum-Saunders dağılımının parametrelerini tahmin etmek için bir teknik sağlar. Bölüm 6, iki parametreli Birnbaum-Saunders dağılımının parametre tahmini için genetik algoritmayı kullanıyoruz. Ayrıca, bölümde gerçek very ile bir uygulama verilmiştir. Göreceli değerlerini değerlendirmek için Genetik Algoritmanın ve mevcut yöntem olan ML'nin performansını karşılaştırabilmek için bir sümulasyon çalışması yer almaktadır. Son olarak, bölüm 7'de bu çalışmalardan çıkarılan sonuçlar ve yorumlar sunulmaktadır.

ACKNOWLEDGEMENTS

My sincere gratitude goes straight to my supervisor Prof.Dr. Güzin YÜKSEL, for her valuable knowledge, expertise and encouragement, without which it would be impossible for me to reach the doctorate. Her requirements have greatly stimulated and edified me. I would also like to thank Prof. Ali Arslan Özkurt for his encouragement to me. This thesis would not have looked the same without the help and encouragement of Prof. Dr. Ali İhsan GENÇ, Prof. Dr. Gülsen KIRAL and Dr. Erkut Tekeli.

I would like to thank my wife Anrafa AHAMADA, who has always been present, for her love and her unfailing support. Thanks also to my son Moudjib for his autonomy which allowed me to free up time to carry out this synthesis work. I express my deep gratitude to my older brother Sir Younoussa Said M'madi and my first Teacher Sir Said Soillih for the advice and almost daily follow-up in the realization of this Memory. Without financial support, I would have a hard time making a living, and I thank YTB for all it brought to me. I also thank the Department of Mathematics and Statistics at Çukurova University for its support.

Finally, I express my deep gratitude to:

My dear parents, Assoumani Rassoul and Riama Mohamed, who have always been there for me. << You have sacrificed everything for your children, sparing neither health nor strength. You have given us a magnificent model of toil and perseverance. I am indebted to an education of which I am proud. >>

All my brothers and sisters.

I offer my sincere thanks to all those who, from near and far, have contributed to my formation or to the development of this memory.

LIST OF TABLES

LIST OF FIGURES

SYMBOLS AND ABBREVIATIONS

BS : Birnbaum-Saunders
CR : Censoring rate
DFR : Decreasing failure rate
DLSV : Dummy least squares variable
EBIAS : Estimated bias
EM : Expectation-Maximization
EMSE : Estimated mean square error
Fig : Figure
GA : Genetic algorithm
IFR : Increasing failure rate
iid : Independent and identically distributed
KS : Kolmogorov-Smirnov
ML : Maximum likelihood
MLE : Maximum likelihood Estimator
MM : Modified moment
MMEs : Modified moment estimators
OLS : Ordinary least squares
PP : Probability-probability
QQ : Quantile-quantile
RV : Random variable
SHN : Sinh-normal
TBS : Two-parameter Birnbaum-Saunders
TTT : Total Time on Test
UMLE : Unbiased maximum likelihood estimators
UMME : Unbiased modified moment estimators

1. INTRODUCTION

In the area of censored data analysis, also known as survival data, we are interested in the time it takes for an event to occur over time. The first recorded event is death by demographers, which explains why the term survival is still used generally, although many other events besides death are now being studied. Indeed, survival data are not the preserve of biostatisticians and are also present in areas such as reliability, economics, insurance, psychology, and many others to name but a few.

In industrial reliability, the events of interest are for example the lifetime of the components of a system; economists are interested in the durations of unemployment episodes, and insurers at the time of arrival of a claim while psychologists measure the time it takes a subject to accomplish a given task. In sociology, the choice is vast with the successions of life events like weddings, the birth of the first child, divorce, etc. In the biomedical field, survival data are mainly found in clinical research in therapeutic trials and cohort studies in epidemiology. In the first case, it is often necessary to compare the effectiveness of two treatments with the criterion of time to an event of interest (relapse, death. . .); in the second case, a cohort of subjects is followed over time for which the disease may or may not appear during the follow-up period and an attempt is made to assess the importance of risk factors on this disease.

The specificity of survival times is to correspond to positive random variables and to have incomplete observations. Indeed, in most prospective studies, individuals are monitored for a predetermined observation period. For the subjects for which the event of interest takes place during the observation period, we have the exact time of appearance of this event of interest, measured from an initial date that must be specified unambiguously (date of randomization in clinical trials, for example). However, some subjects will not have had the event, at the end of the observation period. We will then have for these subjects only incomplete information, namely that the time of occurrence of the event is greater than the time of observation. Such data are said to be censored on the right and the observation period constitutes the censorship period. As subjects enter the study on different dates, each subject has its own time frame and censorship period.

The time observed is therefore the minimum of the time of occurrence of the event and the time of censorship. This censorship mechanism constitutes right-wing random censorship, which is the most common case in prospective studies and which is the case we consider all the time. Subjects who did not have the event of interest during the observation period are called “living excluded” at the end of the study.

Another cause of censorship on the right, which we try to limit to the maximum in the studies, corresponds to the so-called «lost of sight» subjects, who are the individuals who left the study before the occurrence of the event of interest and are no longer new to the study completion date. Other types of censorship can be encountered: data are said to be censored on the left if only

an upper limit of the time of occurrence of the event is known. They are censored by intervals if it is known only that the event took place between two known dates (for example, the passage of the critical threshold of a biological parameter between two patient visits). The analysis of censored data requires an adapted methodology to take into account the information contained in the censorship period. The usual procedures that take into account censorship almost always assume the independence of the variables survival times and censorship times. Censorship is also assumed most of the time not to be informative, in the sense that its distribution does not depend on the parameters involved in the distribution of the survival duration variable. These assumptions are retained throughout. These assumptions are reasonable if the censored data are due to subjects who did not experience the event at the end of the study ("excluded alive"). They are less obviously verified when the censored data correspond to subjects who left the study before the end ("lost sight").

The Birnbaum-Saunders (BS) distribution is indexed in terms of shape (α) and scale (β) parameters. However, when a random variable (RV) T follows the Birnbaum-Saunders distribution with parameters $\alpha > 0$ and $\beta > 0$, the notation $T \sim BS(\alpha, \beta)$ is used and the corresponding probability density function (PDF) is given by

$$f(T;\alpha,\beta) = \frac{1}{2\alpha\beta\sqrt{2\pi}}\left[\left(\frac{T}{\beta}\right)^{-1/2} + \left(\frac{T}{\beta}\right)^{-3/2}\right]\exp\left[-\frac{1}{2\alpha^2}\left(\frac{T}{\beta} - 2 + \frac{\beta}{2}\right)\right], T > 0, \alpha > 0, \beta > 0. \quad (1)$$

Besides, the objective of this thesis is:

1. Use a data set and simulation study to estimate the Birnbaum-Saunders two-parameter model parameters.
2. Propose a genetic algorithm method to compare it with the maximum likelihood method used in the literature.
3. Conclude the result displayed by the genetic algorithm relative to the sample size.

This work contains 7 chapters taking into account this introductory chapter. In Chapter 2, we present information about Birnbaum-Saunders, the censorship and genetic Algorithm. In Chapter 3 we conduct a study of censored data and types of censored data that looks at the time an event occurs over time. However, in chapter 4 we present a detailed study of the genesis and the mathematical derivation of the Birnbaum-Saunders distribution. In addition, a more extensive detail of the probabilistic characteristics, namely probability density, cumulative distribution and quantum functions, is presented. Chapter 5 provides a technique for estimating the distribution of Birnbaum-Saunders. In Chapter 6, we use the genetic algorithm for parameter estimation of the two-parameter Birnbaum-Saunders Distribution. Furthermore, we give a numerical example. We simulate also the performance of GA and the current method, MLE, to assess their relative merits. Finally, in Chapter 7 we present a general conclusion of this work.

2. PRELIMINARY WORK

Known as the fatigue life distribution, the Birnbaum-Saunders distribution is a statistical model that focuses on studying the fatigue life of structures through recycling. Birnbaum and Saunders derived the distribution of Birnbaum-Saunders from the study of ruptures from a fissure progression (Birnbaum and Saunders (1969)). This distribution is unimodal, parametric positive and asymmetrical. For more details (Marshall, A., Olkin, I., (2007) and Saunders (2007)).

Two parameters with shape and scale parameters are present in this distribution. Data with values above zero are studied by this distribution (Birnbaum and Saunders (1969a)). In addition, the distribution of Birnbaum_Saunders studies the total time produced by the failure of a normally distributed type of cumulative damage. This idea for the distribution of Birnbaum_Saunders is the result of its attractive properties, its physical theoretical arguments and its existing link with the normal distribution. Today, the distribution of Birnbaum_Saunders takes its usefulness in various fields including agriculture, air contamination, finance, business, economy, food, forestry and textiles, human mortality, engineering, computer science, environment, insurance, medicine, microbiology, nutrition, pharmacology, psychology, quality control, queue theory, toxicology, water quality and wind energy and others. For more details, please refer to Balakrishnan et al. (2007, 2009b, 2011), Leiva et al. (2007, 2008a,c, 2009b, 2010, 2014), Podlaski (2008), Barros et al. (2008, 2009, 2014), Kotz et al. (2010), Bhatti (2010), Ahmed et al. (2010), Vilca et al. (2010, 2011),Villegas et al. (2011), Azevedo et al. (2012), Ferreira et al. (2012), Marchant et al. (2013a,b) and Santos-Neto et al. (2012).

A random variable following the distribution of Birnbaum_Saunders can be transformed into another random variable through a standard normal distribution. For the maintenance of such a transformation, various arguments can be put forward for the construction of more general model classes. Very recently, Samet Kaya (2022) worked on parameters Estimation of the Birnbaum-Saunders distribution by the Bayesian method. In addition, a group of international and transdisciplinary researchers have advanced methods of extensions and generalizations of the Birnbaum_Saunders distribution. However, the first extension of this distribution was attributed to Volodin and Dzhungurova (2000). Five years later, a generalization of the distribution of Binbaum_Saunders was introduced by Díaz-García et al.(2005) based on elliptical distributions. Thus Owen(2006) presented the extension of the Birnbaum_Saunders distribution to three parameters. However, Vilca and Leiva (2006) presented a Birnbaum_Saunders distribution based on skew-normal models.

Leiva et al. (2009b) provided a bias in the length of the BS distribution. Gómez et al. (2009) expanded the distribution of Birnbaum_Saunders to a slash-elliptical model. A dispersed version of the Birnbaum_Saunders distribution was proposed by Guiraud et al.(2009). Birnbaum_Saunders distribution mixing models were developed by Vilca et al. (2010). Through

Castillo et al. (2011), the epsilon-skew version of the Birnbaum_Saunders distribution was introduced. Subsequently, Cordeiro and Lemonte (2011) introduced the beta-Birnbaum_Saunders distribution. A demonstration of the Kumaraswamy-Birnbaum_Saunders distribution was presented by Saulo et al. (2012).

A Marshall-Olkin-Birnbaum_Saunders. (MOBS) distribution study was conducted by Lemonte (2013). The following year, an alpha-power extension of the Birnbaum_Saunders distribution was presented by Martínez-Flórez et al (2014).

In reliability, censorship consists of taking into account non defaulting systems for establishing the law of reliability. Generally, censorship applies when the failure date is not known exactly, either because the failure has not yet occurred or because it has not been accurately recorded. Although censorship information is less rich than a definite moment of failure, censorship is therefore information that must be integrated into the reliability model. Moreover, in survival analysis, the best idea is to consider censored patients as having similar survival prospects to those of participants who continued to be followed (Bland JM, Altman DG (1998)). Thus, participants who drop out of the study must do so for reasons unrelated to the study. In this case, censorship in survival analysis is called «non-informative». Unlike informative censorship which occurs when participants are lost for study-related reasons. For example, in a study comparing psychological monitoring after an earthquake that shook a country. Patients in the intervention arm can be completely cured by effective treatment and no longer feel the need for follow-up. Therefore, if these participants are systematically censored, the true effect of the treatment will not be detected and the results of the study will be biased. Survival rates without tromatism would be based on patients who continued to be followed in the study, and would be over-estimated for the control arm and under-estimated for the treatment arm.

In addition, methods are described to correct the problem of informative censorship. These include imputation techniques for missing data, sensitivity analyses to mimic the best and worst scenarios, and the use of the abandonment event as a study endpoint (Shih W (2002)). For an impartial analysis of survival curves, censorship due to loss of follow-up must be minimal and truly «non-informative». Not understanding these aspects of survival analysis could lead to grossly erroneous results from well-conducted studies.

We find that genetic algorithms (GAs) have experienced a resurgence of interest in recent years thanks to their flexibility in dealing with non-regular objective functions or functions defined on non-standard search spaces and their fields of application in several fields. Thus, in our thesis, a basic mechanism of these algorithms and an overview of their application are presented. Moreover, the genetic algorithm is a method inspired by the theory of evolution as defined by the British naturalist Charles Darwin in the last century in his book entitled “The Origin Of Species”. In addition, genetic algorithms (GAs) are stochastic optimization algorithms based on the mechanisms of natural selection and genetics. Their operation is extremely simple. We start with a population of

potential solutions (chromosomes) arbitrarily chosen. Their relative performance (fitness) is evaluated. Based on these performances, a new population of potential solutions is created using simple evolutionary operators: selection, crossing and mutation. We repeat this cycle until we find a satisfactory solution. John Holland in their work on evolutionary systems in 1975 found the first formal model of genetic algorithms (the canonical genetic algorithm GAC) in his book “Adaptation in Natural and Artificial Systems”. He explained how to add intelligence into a computer program with crossovers and mutation. This model will serve as a basis for further research and will be more particularly taken up by Goldberg D.(1989). These genetic algorithms are applied in various fields including economics, where they are used for function optimization (De Jong (1980)), finance (Robert Pereira (2000)), optimal control theory (Krishnakumar K and Goldberg DE (1992), Michalewicz, Janikow and Krawczyk (1992) and Marco et al. (1996), or in the theory of repeated games (Axelrod (1987)) and differentials (Özyildirim (1996, 1997) and Alemdar, N. et al. (1998)). Thus, genetic algorithms (AGs) are stochastic optimization algorithms, metaheuristic, belonging to the family of evolutionary algorithms, based on the mechanisms of natural selection and genetics, hence their name << genetic algorithm >>. In addition, genetic algorithms are an original approach: it is not a question of finding an exact analytical solution or a good numerical approximation, but of finding solutions that best meet different, often contradictory criteria. If they do not allow us to find for sure the optimal solution for the research space, at least it can be found that the solutions provided are generally better than those obtained by conventional methods.

3. CENSORING TYPES AND ANALYSIS

3.1. Concepts and Interest Functions

We designated by T the time of the event of interest. T is a known positive random variable. To make it easier to explain these basic concepts, the variable T will be set as the representative of survival time, which means that the observed event is death. In survival analysis, the most commonly used functions that best characterize the T distribution are the instantaneous risk function h, the survival function S and the cumulative risk function H.

The T probability density rated f is defined by:

$$f(t) = \frac{d}{dt}F(t), \tag{2}$$

where $F(t)$ is its distribution function representing the probability of death of the individual before time t defined by:

$$F(t) = P(T \leq t) = \int_0^t f(u)du. \tag{3}$$

However, the survival function is defined by:

$$R(t) = S(t) = P(T > t) = 1 - \int_0^t f(u)du. \tag{4}$$

The instantaneous risk of death will therefore be defined as

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t < T \leq t + \Delta t | T > t)}{\Delta t}. \tag{5}$$

$h(t)$ represents the probability of death between t and $t + \Delta t$ for a subject, conditional on the fact that the subject was still alive just before t. The survival risk function $h(t)$ checks the following relationships:

$$h(t) = \frac{f(t)}{S(t)} = -\frac{S'(t)}{S(t)} = \frac{d}{dx} ln \frac{1}{S(t)}. \tag{6}$$

In addition, the cumulative risk function H is defined by

$$H(t) = \int_0^t h(x)dx \tag{7}$$

which is also defined by

$$H(t) = ln\frac{1}{S(t)}. \tag{8}$$

3.2. Censorship to Right

A patient population whose survival time is described by a random variable T. A sample of n individuals is randomly drawn from this population. The random variables T_i (where $i \in (1, n)$) describing the survival time of n patients are assumed to be independent and identically distributed (iid). The effective life of the patient i is noted t_i. It is censored if $t_i > c_i$. In general, we observe $\tau_i = min\{t_i, c_i\}$ the time frame observed and

$$\delta_i = \mathbf{1}_{(t_i \le c_i)} = \begin{cases} 1 & \text{if } t_i \le c_i \text{ detection,} \\ 0 & \text{if } t_i > c_i \text{ censored data} \end{cases}$$

the censorship indicator. In addition, c_i denotes the threshold for uncensored data. The n-sample $(\tau_i, \delta_i)_{i=1,2,\dots,n}$ thus constitutes the observed data. We fix at $t_{(1)} < t_{(2)} < \cdots < t_{(k-1)} < t_{(k)}$ the observed unobscribed separate times of death ordered and $\gamma_{(i)}$ the number of times of death equal to $t_{(i)}$. $\zeta(t_{(i)})$ then designates all subjects at risk of death just before $t_{(i)}$.

3.3. Censorship to Left

Left censorship occurs if the individual has already realized the event on the date of observation. As a result, there is a lack of information on the exact date of occurrence; the event already occurred on the date of review.

For example, in a study on the side effects of the BioTech vaccine for COVID-19, if the patient no longer remembers his first disorder, it can only be inferred that the age of onset is younger than the current age of the individual. In the presence of censoring on the left, the information available for each individual is $\{\tau_i; \delta_i\}$ with:

- τ_i the actual duration observed: $\tau_i = max(t_i; c_i)$.

- $\delta_i = \mathbf{1}_{(t_i \geq c_i)} = \begin{cases} 1 & \text{if } t_i \geq c_i \text{ detection,} \\ 0 & \text{if } t_i < c_i \text{ censored data} \end{cases}$

3.4. Interval Censorship

In this case, the event is not observed, but it occurs between two observation dates. The only information available is that it occurred between two known dates. It is common in longitudinal studies that patients are not followed continuously. However, the exact date of occurrence of the event will never be known. Therefore, right censorship and left censorship are special cases of interval censorship. Now let's look at the follow-up study of patients with COVID-19. The first symptoms of COVID-19 in a patient can only be seen during a medical visit. On the graph below, Individual 2 made two visits: the first in a_1 and the second in a_3. If on the second visit, it is declared positive for COVID-19, the only information on the patient's positivity date is a time interval between the two visits, namely $T_2 \in [a_1; a_3]$.

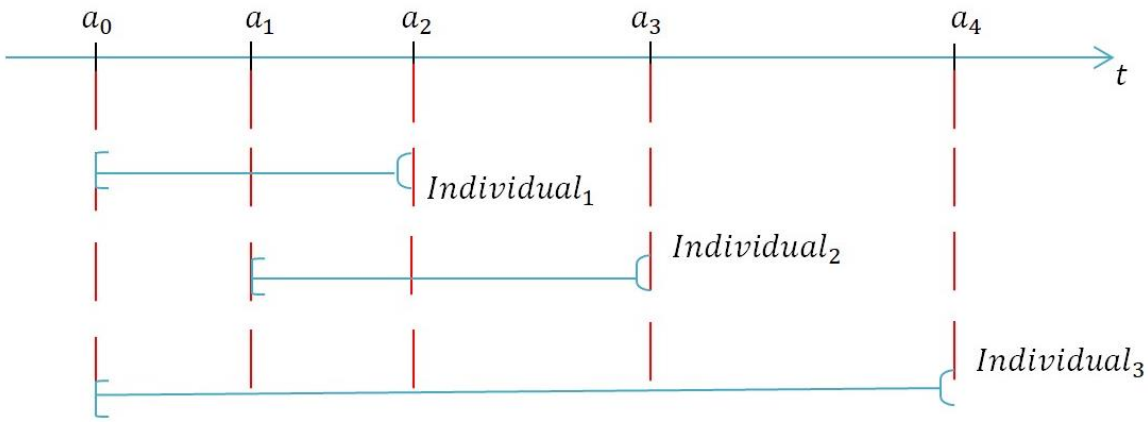

Figure 3.1. Real representation of interval censored data.

3.5. Type I Interval Censorship

This censorship type refers to studies where the individual is observed only once during the study. This type of data is very common in demography, epidemiology and reliability studies. Again, the date of occurrence of the event of interest is not known. Only the current status of the subject is known on the single visit date. All intervals contain either zero or infinity, that is, observations are either censored to the left or censored to the right. For each patient, two pieces of information are available: the follow-up time C, and the status of the patient at this date δ. Thus, it is common to represent censors by type 1 interval by:

$$\{C, \delta = \mathbf{1}_{(T \geq C)}\}$$

3.6. Type II Interval Censorship

The random variable T is not always known accurately. Thus, the information on T can be expressed by the following interval:

$$T \in (D_1, D_2],\ D_1 \leq D_2,$$

with D_1 and D_2 random variables, independent of T.

Type II interval censorship, often simply referred to as "interval censorship", includes the case of an event date belonging to an interval, such as $T \in (D_1, D_2]$. Typically, type II censorship includes only studies in which each patient performs only two visits. Information censored by interval may be expressed as:

$$\{D_1, D_2, \delta_1 = \mathbf{1}_{(T \leq D_1)}, \delta_2 = \mathbf{1}_{(D_1 \leq T \leq D_2)}, \delta_3 = 1 - (\delta_1 + \delta_2)\}$$

The δ_1 represent the indicators of censorship left, the δ_2 represent the indicators of censorship by interval and the δ_3, the indicators to censorship right.

A natural extension of type II censorship, is type k censorship where k is a fixed number of visits. However, knowing that at most two intervals of observations among the $(k\ +1)$ will be necessary to frame the variable T, it is possible to ignore other intervals that do not provide more information about T and revert to a type II censorship case. This is the most common type of interval censorship.

3.7. Non-Parametric Maximum Likelihood Estimate

An empirical estimator $\tilde{S}_{KM}(t)$ of $S(t)$ is the fraction of t_i values strictly greater than. If the t_i are in ascending order $\left(t_{(1)} \leq t_{(2)} \leq \cdots t_{(n-1)} \leq t_{(n)}\right)$, we have

$$\hat{S}_{KM}(t) = \begin{cases} 1 & \text{if } t < t_{(1)} \\ \prod_{i:t_{(i)} \leq t} \left(1 - \frac{\gamma_{(i)}}{d_{(i)}}\right) & \text{if } t \geq t_{(1)} \\ 0 & \text{if } t \geq t_n \end{cases} \tag{9}$$

where $d_{(i)}$ is the number of items in a set $\zeta\left(t_{(i)}\right)$. The Kaplan-Meier estimator converts almost surely and uniformly to $S(t)$, for more details, we can refer to the book (Gill,R. 1980).

i.e.

$$\sup_{0\le t\le \tau_i} \left|\hat{S}_n(t) - S(t)\right| \xrightarrow[n\to\infty]{} 0$$

Under certain regularity conditions, it converges in law to a Gaussian process. In other words, the Kaplan-Meier estimator (1958) is the only coherent survival estimator. However, the probability of dying beyond t is the sum of the probabilities of being at risk at t and being censored before t.

i.e.

$$\hat{S}(t) = \frac{1}{n}\left[\sum_{i=1}^{n} \mathbf{1}\{t_i > t\} + \sum_{i=1}^{n} \mathbf{1}\{t_i \le t, \delta_i = 0\} \cdot \frac{\hat{S}(t)}{\hat{S}(t_i)}\right]. \quad (10)$$

Moreover, this estimator is that of the maximum non-parametric likelihood of survival in the space of survival functions. Using (8), the Kaplan-Meier estimator can be derived to obtain the cumulated risk estimator $\widehat{H}_{Br}(t)$ known as the Breslow cumulative risk estimator (Breslow, N. E. 1972).

i.e.

$$\widehat{H}_{Br}(t) = \ln \frac{1}{\hat{\hat{S}}_{KM}(T)}$$

A more well-known cumulative risk estimator known as the Nelson-Aalen (1978) estimator is also defined as:

$$\widehat{H}_{NA}(t) = \begin{cases} 0 & \text{si } t < t_{(1)}, \\ \sum_{i:t_{(i)}} \frac{\gamma_{(i)}}{d_{(i)}} & \text{si } t \ge t_{(1)}. \end{cases} \quad (11)$$

Another survival function estimator S known as the Harrington and Fleming estimator can be obtained by having the relation $S(t) = e^{-H(t)}$ and the expression $\widehat{H}_{NA}(t)$.

3.8. Comparison of Survival of Two or More Groups

Various tests are proposed to compare two or more survival curves obtained by the nonparametric method (Kaplan-Meier or Actuarial): Log-Rank test, Mantel-Haenzel test and Wilcoxon test. To compare the survival functions of two or more groups, weighted log-rank tests are best suited. Like any statistical test, the Log-Rank test aims to test a null hypothesis (H_0)

against an alternative hypothesis (H_1). The hypothesis formulation for the comparison of two curves associated with two groups (A and B) is as follows:

Null hypothesis (H_0): the survival curves are the same, $S_A(t) = S_B(t)$ for any $t > 0$ which means for any t value the observed deaths should be distributed in both groups proportionally to the persons at risk.

Alternative hypothesis (H_1)*:* the survival curves are different in either $S_A(t) \neq S_B(t)$ which means that the observed deaths do not occur with the same frequency and/or at the same time in both groups proportionally to the people at risk. Based on the null hypothesis, the statistics of the weighted log-rank tests are written:

$$L_r = \frac{\sum_{i=0}^{n} w_i \left(\gamma_{B_{(i)}} - \gamma_{(i)} \frac{d_{B_{(i)}}}{d_{(i)}} \right)}{\left(\sum_{i=0}^{n} w_i^2 \gamma_{(i)} \frac{d_{(i)} - \gamma_{(i)}}{d_{(i)} - 1} \frac{d_{A_{(i)}} d_{B_{(i)}}}{d_{(i)}^2} \right)^{1/2}} \tag{12}$$

In general, it has been shown (Gill 1980) that the asymptotically optimal weights for a given alternative are proportional to $ln\big(h_B(t)\big) - ln\big(h_A(t)\big)$, where h_A and h_B are the instantaneous risks of death in groups A and B respectively. The w_i weights give more importance to certain deaths. The choice of w_i gives a different test:

a. if $w_i = 1, \forall i = 1, \dots, n$ gives the lograng test.
b. if $w_i = d_{A_{(i)}} + d_{B_{(i)}}, \forall i = 1, \dots, n$ gives the Gehan-Wilcoxon test. It gives more weight to events (the t_i for which $\Delta_i = 1$) that occur at early times.
c. if $w_i = \hat{S}_{KM}, \forall i = 1, \dots, n$ gives the Peto/Prentice test. It is also called the generalized log-rang test. It also gives more weight to events (t_i for which $\Delta_i = 1$) that occur at early times.

Moreover, the log-rank test defined by (a) is optimal for proportional risk alternatives, which corresponds to the fundamental hypothesis of Cox regression. So:

- The Log-rank test only involves the rank of observations
- This test easily extends to more than two groups. The test statistic follows asymptotically a law of χ^2 whose number of degrees of freedom is equal to the number of groups minus 1.
- The choice of w_i weights influences the power of the tests.

3.9. Cox Regression

The British statistician David Cox (1972) presented the proportional risk model, regression models and survival tables in the Journal of the Royal Statistical Society. This statistical model, Cox's proportional risk model, does not impose any form specific to the survival function, which makes it possible to model the censored data in a very flexible way. Cox's regression makes no assumptions about the underlying distribution of data in which the predictors are multiplicatively related to lifetime. Cox's proportional risk model is modelled by the following product:

$$h(t|X) = h_0(t)\exp(\beta' X) \tag{13}$$

with X a p −vector of covariates, β the p −vector of regression parameters and h_0 an unknown basic risk function. Being a log-linear model, Cox's regression focuses on the proportional risks hypothesis, which generates a constant result of the ratio of the instantaneous risks of two individuals over time as well as a dependence on their covariates. Under the model (13) it can be noted that time does not influence on the instantaneous risk ratio for two subjects with covariates X_1 and X_2.
i.e.

$$\frac{h(t|X_1)}{h(t|X_2)} = \frac{h_0(t)\exp(\beta' X_1)}{h_0(t)\exp(\beta' X_2)} = [\exp(\beta' X_1)][\exp(-\beta' X_2)] = \exp[\beta'(X_1 - X_2)]. \tag{14}$$

Cox (1972) offered heuristically to estimate the vector of the β parameters and to carry out tests of significance by maximizing what he called partial likelihood, by considering the basic risk function h_0 as a nuisance parameter. Later, Gill (1980) demonstrated these asymptotic properties in the context of the theory of point processes, which makes it possible to use Cox's partial likelihood as a usual likelihood function (Wald tests, score or likelihood ratio on model parameters). Cox's partial likelihood is written as a conditional likelihood product calculated at each time $t_{(i)}$, the $t_{(i)}$ being assumed to be fixed. The V_i contribution to the likelihood of the subject i whose death was observed in t_i is equal to the probability, conditional on $\zeta(t_{(i)})$, that it is precisely this subject who dies in $t_{(i)}$ among the set $\zeta(t_{(i)})$ of subjects at risk in $t_{(i)}$. It was therefore

$$V_i = \frac{h(t_{(i)}, X_i)}{\sum_{j\in\zeta(t_{(i)})} h(t_{(i)}, X_j)}.$$

Under the Cox model, check that the nuisance term $h_0(t_{(i)})$ eliminate in this expression and it remains

$$V_i = \frac{\exp(\beta' X_i)}{\sum_{j \in \zeta(t_{(i)})} \exp(\beta' X_j)}.$$

The Cox partial likelihood is then calculated as the product on i of all the contributions of the deceased subjects and in the absence of equal lifetime, the partial log-likelihood function is then valid

$$L(\beta) = \sum_{i=1} \left(\beta' X_i - \ln \left(\sum_{j \in \zeta(t_{(i)})} e^{\beta' X_j} \right) \right). \quad (15)$$

We can refer to Breslow (1974) and Efron (1977) for more details on partial log-likelihood. Moreover, an overview of many extensions of Cox's regression can be found in Therneau et al.(2000).

4. BIRNBAUM-SAUNDERS DISTRIBUTION

In this section, we have provided a context and history of life distributions. However, a detailed study of the fatigue process, the genesis and the mathematical derivation of the Birnbaum-Saunders distribution is also developed. In addition, using the standard normal distribution, the most useful characteristics and properties, moments and aspects of simulation of the Birnbaum Saunders distribution are studied. In particular, a more extensive detail of the probabilistic characteristics, namely probability density, cumulative distribution and quantum functions, is presented. In addition, lifetime indicators including reliability function and failure rate are also detailed. Using probabilistic characteristics, the mode and median of the distribution are determined. Recall that the failure rate of the Birnbaum-Saunders distribution has an unimodal form.

In addition, a series of moment calculations was performed. By using the relationship between the Birnbaum Saunders and log-Birnbaum Saunders distributions as well as the Bessel function, we obtain moments of all kinds, including negative moments. In addition, a calculation of mean, variance coefficients of variation, asymmetry and finally flattening is made. At the same time, an overview of Birnbaum-Saunders parameter estimation methods is provided. Subsequently, a discussion on the modified moment method for parameter estimation was discussed. A graphical descriptive approach to facilitate the estimation of corresponding parameters using the least squares method used in the regression is presented. In light of the above, regression models and their diagnostics based on the Birnbaum-Saunders distribution with uncensored and censored data are presented. At the end of this section, adjustment methods and model selection criteria for the Birnbaum-Saunders distribution is discussed.

4.1. Genesis of the Birnbaum-Saunders Distribution

In all cases, the methods and statistical models for lifetime data recognize the concept of "lifetime" as a positive continuous random variable representing time until an interesting event is created. Thus, in particular cases such as the ageing of elements, it is not measured in chronological terms. In such cases, other random variables are used to measure lifetime. Examples include the level of degradation, the number of kilometres travelled, the strength of the material samples until it ruptured, the flexibility of an adhesive, and the number of cycles until a fatigue (fatigue-life) material sample ruptured.

In addition, the notion of a “random lifetime variable”, hereafter referred to as a T, is the designation of any positive continuous random variable (for example, the concentration of the COVID-19 virus epidemic contaminant in a country). In addition, the term life distribution is the probability associated with a random life variable. For more details on life distributions, (Marshall and Olkin (2007)). Statistical and probabilistic problems related to the lifetime distributions of

failing subjects are referred to as the theory of reliability. Based on this theory, the performance and degradation of objects using random lifetime variables are described by probabilistic methods. For the determination of the reliability (or survival) function, the failure rate (or risk) and the average lifetime of these elements, lifetime distributions are used. The statistical aspects of this theory are used to solve the aspects of estimation and inference of life distribution parameters. In short, the reliability of an event is the probability that a random life variable $T > 0$ exceeds a fixed time t. Therefore, over time t, the reliability function of T is defined by the equation (4). Note that the reliability function depends on the continuous lifetime distribution when parametric analysis is performed. In addition, such a distribution for T is characterized either by its probability density function $f_T(\cdot)$, its cumulative distribution function $F_T(\cdot)$ or by its failure or chance rate given by equation (6), where we define its cumulative failure by equation (7).
with,

$$\frac{f_T(v)}{1 - F_T(v)} = h_T(t)$$

Normal (or Gaussian) distribution is the most commonly used probabilistic model. However, two-parameter probabilistic models with right asymmetry (positive asymmetry), unimodality and positive support $(T > 0)$ are often life distributions. In addition, the normal distribution is not suitable for modelling lifetime data.

Based on parametric lifetime analyses, the key distribution is the exponential model, known as the negative exponential distribution. Using the exponential distribution, the modelling of the sum of a lifetime with this distribution is done using the gamma distribution. A good thing about the distribution of gamma life is that it naturally results from the convolution of exponential distributions, but its disadvantage is that the gamma model is not algebraically treatable.

Several researchers have conducted studies that have found life data that are better modelled by a different life distribution than the exponential model. For example, Lieblein and Zelen (1956) and Kao (1959) worked on the distribution of Weibull, as this model corresponds well to the failure data. For a description of the breaking strength of the materials, the Weibull distribution was developed in 1951. Due to the work of Zelen and Dannemiller (1961), many lifetime testing procedures based on exponential distribution are inadequate.
This is at the origin of the increase in interest in the Weibull distribution of life. Because distribution is flexible to describe different types of hazardous situations and, in addition, it is mathematically treatable, this distribution is therefore very useful. In 1968, Birnbaum and Saunders used material fatigue in commercial aircraft caused by their constant vibrations and the problems they cause, to introduce a probability model describing the lifetime associated with samples of materials exposed to fatigue. It took Birnbaum and Saunders (1969a,1969b) to introduce the

lifetime distribution that would later be named after them. From this distribution, a two-parameter parameter estimation method was proposed. In addition, Birnbaum and Saunders (1958) erred on a model describing the fracture time associated with samples of fatigue materials. The main results of these investigations have been established in the following three works:

(1) Birnbaum et al. (1966) developed a stochastic wear pattern.
(2) Birnbaum and Saunders (1969a) calculated the fatigue lifetime distribution.
(3) Esary et al. (1973) proposed the shock model.

The Birnbaum-Saunders distribution focused on the cumulative material damage that causes material fatigue. This notion is the fruit of Miner; see Miner (1945). Birnbaum and Saunders (1968) revealed the probabilistic significance of this fruit. In addition, the fatigue life distribution of Birnbaum and Saunders (1969a) is extracted from a model for which the time consumed until the cumulative damage, due to the evolution of a dominant crack, exceeds a threshold value and results in the stranded material sample.

Desmond (1985) improved the usefulness of this distribution by facilitating the use of some of the first hypotheses made by Birnbaum and Saunders (1969a). In addition, a more general derivation of the Birnbaum-Saunders distribution was provided by Desmond (1985) from a biological model introduced by Cramér (1946). Noting that the life fatigue distribution was known as the Birnbaum-Saunders model, Freudenthal and Shinozuka (1961) presented a similar model with a different parameter. In addition, Fréchet (1927) proposed ideas related to the Birnbaum-Saunders distribution, but also with a parameter different from that considered by Birnbaum and Saunders (1969a). This was done occasionally by Birnbaum and Saunders (1969a) and appears to have been introduced in Desmond (1985), Johnson et al. (1995).

4.2. Characteristics of the Birnbaum-Saunders Distribution

Recall that when a random variable T follows a Birnbaum-Saunders (BS) distribution of form parameter α and scale parameter $\beta > 0$, we note $T \sim BS(\alpha, \beta)$. In addition, choosing a random variable

$$Z = \frac{1}{\alpha}\left(\left(\frac{\beta}{T}\right)^{-\frac{1}{2}} - \left(\frac{T}{\beta}\right)^{-\frac{1}{2}}\right). \tag{16}$$

According to the normal distribution of mean $\mu = 0$ and variance $\hat{\sigma}^2 = 1$, then a random variable Z follows the standard normal distribution and we note $Z \sim N(0,1)$.
i.e.

$$Z = \frac{1}{\alpha}\left(\left(\frac{\beta}{T}\right)^{-\frac{1}{2}} - \left(\frac{T}{\beta}\right)^{-\frac{1}{2}}\right) \sim \mathcal{N}(0,1). \tag{17}$$

Therefore, for a Birnbaum-Saunders distribution, any random variable T can be the result of a random variable Z of a standard normal distribution. However, we get this one

$$T = \frac{\beta}{2}\left(\alpha Z + \sqrt{(\alpha Z + 2)^2 - \frac{\alpha Z}{2}}\right)^2 \tag{19}$$

4.3. Probability Functions and Properties

Let T be a random variable according to the Birnbaum-Saunders distribution of shape parameter α and scale parameter $\beta > 0$. That is $T \sim BS(\alpha, \beta)$. Then, the probability density function is defined by

$$f_T(t) = \frac{1}{2\alpha\beta\sqrt{2\pi}}\left(1 + \frac{\beta}{t}\right)\sqrt{\frac{\beta}{t}}.\exp\left[-\frac{1}{2\alpha^2}\left(\frac{t}{\beta} + \frac{\beta}{t} - 2\right)\right] \tag{20}$$

with α and β strictly above zero. The following probability density function can be easily obtained using the transformation theorem of random variables.

We illustrate the probability density function with the figure below by setting its scale parameter to the unit ($\beta = 1$) and assigning its shape parameter different values.

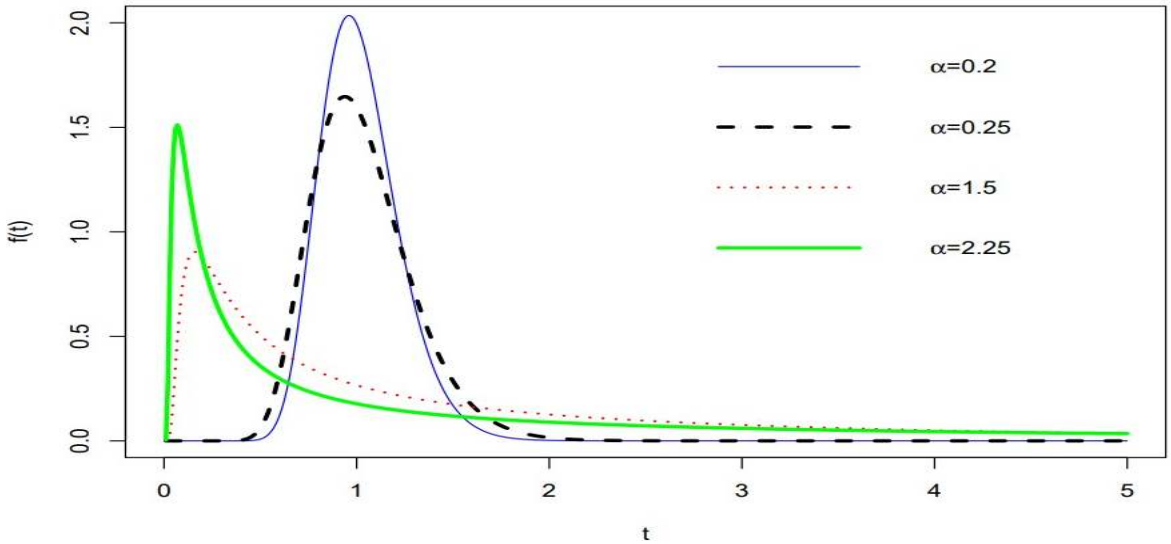

Figure 4.1. Plot of the Birnbaum-Saunders probability density function for the indicated values of $\alpha = 0.2, 0.25, 1.5, 2.25$ with $\beta = 1{,}0$.

Using monotone transformation (19), we can obtain the cumulative distribution function of $T\sim BS(\alpha,\beta)$ as

$$F_T(t|\alpha,\beta)=P(T\leq t)=P\left(\frac{\beta}{2}\left(\alpha Z+\sqrt{(\alpha Z+2)^2-\frac{\alpha Z}{2}}\right)^2\leq t\right)$$

$$=P\left(Z\leq\frac{1}{\alpha}\left(\left(\frac{\beta}{T}\right)^{-\frac{1}{2}}-\left(\frac{T}{\beta}\right)^{-\frac{1}{2}}\right)\right).$$

where,

$$F_T(t)=\Psi\left(\frac{1}{\alpha}\varphi\left(\frac{\beta}{t}\right)\right),\alpha>0,\beta>0\text{ and }t>0 \tag{21}$$

with, $\varphi(\varepsilon)=\varepsilon^{\frac{1}{2}}-\varepsilon^{-\frac{1}{2}}$ and

$$\Psi(z)=\int_{-\infty}^{z}\psi(\varepsilon)d\varepsilon\,,\qquad z\in\mathbb{R} \tag{22}$$

represents the standard normal cumulative distribution function, where

$$\psi(\varepsilon)=\frac{1}{\sqrt{2\pi}}exp\left(-\frac{1}{2}\varepsilon^2\right),\qquad \varepsilon\in\mathbb{R}. \tag{23}$$

In addition, we can also retain this by integrating the function of probability density $f_T(t|\alpha,\beta)$ described by equation (20), we obtain equation (21). For statistical analyses, the quantile function is another useful indicator. Note that if the aleatory variable $T\sim BS(\alpha,\beta)$ then,

$$F_T(t)=\Psi\left(\frac{1}{\alpha}\varphi\left(\frac{\beta}{t}\right)\right),\alpha>0,\beta>0\ \text{ and }\ t>0. \tag{24}$$

By definition, the q-th quantile is the value $t_q>0$ satisfying the following condition

$$F_T(t_q)=\Psi\left[\frac{1}{\alpha}\left\{\left(\frac{t_q}{\beta}\right)^{1/2}-\left(\frac{\beta}{t_q}\right)^{1/2}\right\}\right]=q. \tag{25}$$

As Ψ is invertible, we get

$$\frac{1}{\alpha}\left\{\left(\frac{t_q}{\beta}\right)^{1/2}-\left(\frac{\beta}{t_q}\right)^{1/2}\right\}=\Psi^{-1}(q), \tag{26}$$

because $\Psi^{-1}(q)$ is the q-th quantile of a standard normal random variable. To make the algebraic calculation easy, we first observe that

$$\left(\frac{t_q}{\beta}\right)^{1/2}-\left(\frac{\beta}{t_q}\right)^{1/2}=\frac{1}{\sqrt{\beta}}\left(\frac{t_q-\beta}{\sqrt{t_q}}\right).$$

From the equation (26) we have,

$$\alpha\sqrt{\beta}z_q=\frac{t_q-\beta}{\sqrt{t_q}}$$

i.e.

$$t_q-\alpha\sqrt{\beta}z_q\sqrt{t_q}-\beta=0.$$

Let us assume that $u=\sqrt{t_q}$, then we have the quadratic $u^2-\alpha\sqrt{\beta}z_q u-\beta=0$.
From the quadratic formula, it comes

$$u=\frac{\alpha\sqrt{\beta}z_q\pm\sqrt{\alpha^2\beta z_q^2+4\beta}}{2}. \tag{27}$$

Assuming that $\sqrt{t_q}\geq 0$ and based on equation (27), the quantile function of $T\sim BS(\alpha,\beta)$ is given by

$$t(q)=F_T^{-1}(p|\alpha,\beta)=\beta\left(\frac{\alpha z(q)}{2}+\sqrt{\left(\frac{\alpha z(q)}{2}\right)^2+1}\right)^2, \tag{28}$$

with $F_T^{-1}(\cdot)$ the inverse function of $F_T(\cdot)$ defined by equation (21) and for 0 < q ≤ 1, $z(q)=\Psi^{-1}(q)$ is the $q\times$100th quantile of the standard normal distribution. Given $\Psi^{-1}(\cdot)$ the inverse function of $\Psi(\cdot)$, if $q=0.5$, then $z(0.5)=0$, which corresponds to the mean, median and mode of $Z\sim\mathcal{N}(0,1)$. Suffice it to note that by fixing $q=0.5$, it results from the quantile function defined in equation (28) that $t(0.5\,;\,\alpha,\beta)=\beta$.

This proves that the $\beta > 0$ scale parameter is also the median of the Birnbaum-Saunders distribution. Equation (28) can be used to generate random numbers in the simulation processes of the Birnbaum – Saunders distribution but also to derive associated quality fit tools. In lifetime analysis, the survival or reliability function and the failure rate or hazard function are two useful indicators.

From equations (20) and (24), the reliability function and the failure rate of $T \sim BS(\alpha, \beta)$ are, respectively, given by

$$R_T(t) = 1 - F_T(t) = 1 - \Psi\left(\frac{1}{\alpha}\varphi\left(\frac{\beta}{t}\right)\right), \quad t > 0, \tag{29}$$

$$h_T(t) = \frac{f_T(t)}{R_T(t)} = \frac{\frac{1}{2\alpha\beta\sqrt{2\pi}}\left(1 + \frac{\beta}{t}\right)\sqrt{\frac{\beta}{t}}.\exp\left[-\frac{1}{2\alpha^2}\left(\frac{t}{\beta} + \frac{\beta}{t} - 2\right)\right]}{\Psi\left(-\frac{1}{\alpha}\varphi\left(\frac{\beta}{t}\right)\right)}, t > 0, \tag{30}$$

for $0 < R_T(t|\alpha, \beta) < 1$.

It may be noted that when $\boldsymbol{t}$ tends to infinity, the failure rate of the Birnbaum-Saunders distribution is reduced to $\frac{1}{2\alpha^2\beta}$. On the other hand, when $\boldsymbol{\alpha}$ tends towards zero and the change point t_c of $h_T(t|\alpha, \beta)$ is large, the failure rate decreases. For more details, (Chang and Tang (1993) and Kundu et al. (2008)). At the same time, an illustration of the theories described above is made by a graphic display of the failure rate of the Birnbaum-Saunders Distribution for different values of α and $\beta = 1$.

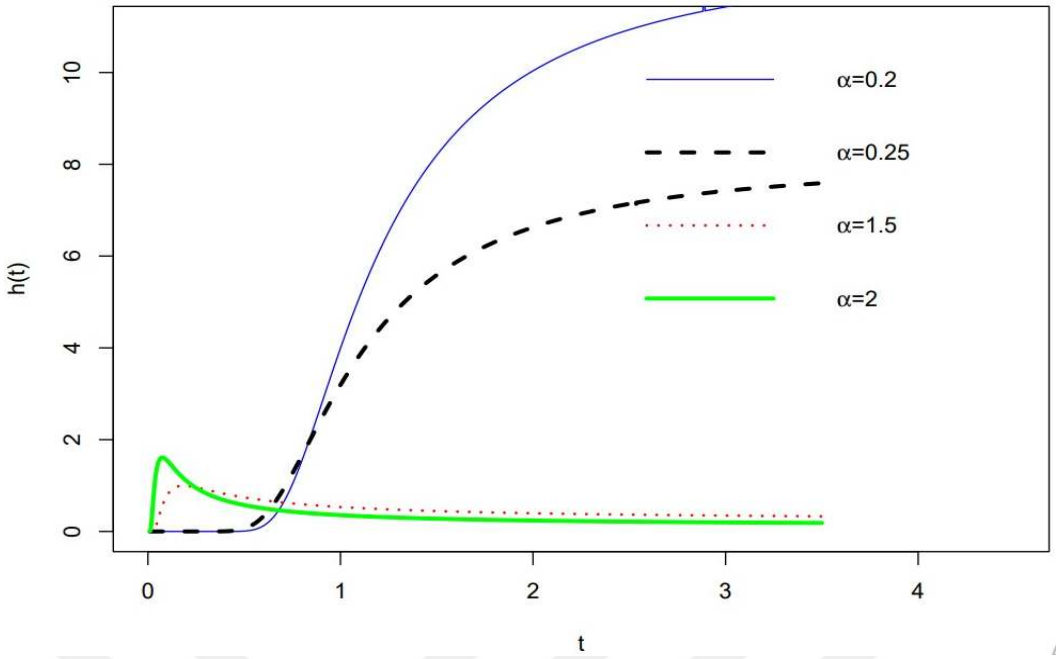

Figure 4.2. Plot of the Birnbaum–Saunders failure rate for the indicated values of $\alpha = 0.2, 0.25, 1.5, 2$ with β = 1.0.

When a variable $T \sim BS(\alpha, \beta)$, then the random variable $X = aT \sim (\alpha, a\beta)$ whatever has positive. It should be borne in mind that the distribution of Birnbaum-Saunders can be defined by the property according to which

$$z^2 = \frac{1}{\alpha^2}\left(\frac{t}{\beta} + \frac{\beta}{t}\right) \sim \chi^2(1),$$

with, $\mathrm{E}(z^2) = 1$ and $\mathrm{Var}(z^2) = 2$.

Let T be the variable defined in paragraph (4.2) with the probability density function defined by equation (20). Then the Fourier transform of T is defined by

$$\Xi_T(s) = E[\exp(isT)] = \int_0^{\infty} \exp(isT) f_T(t) dt,$$

where, $\exp(isT) = \cos(sT) + isin(sT)$ et $i^2 = -1$ is the imaginary unit. We can establish the probability density function of T as soon as the characteristic function has been determined. Thus, one can determine the moments of order r of the random variable T by

$$E(T^r) = \frac{\Xi_T^{(r)}(0)}{i^r}, \qquad (31)$$

with $\Xi_T^{(r)}(0)$ is r-th derived from the characteristic function $\Xi_T(\cdot)$ valued at $r = 0$. Now, $T \sim BS(\alpha, \beta)$. Using

$$\Xi_T(s) = \int_0^{\infty} \exp(isT) \frac{1}{2\alpha\beta\sqrt{2\pi}} \left(1 + \frac{\beta}{t}\right) \sqrt{\frac{\beta}{t}} . \exp\left[-\frac{1}{2\alpha^2}\left(\frac{t}{\beta} + \frac{\beta}{t} - 2\right)\right] dt. \tag{32}$$

Under the above and by taking $\beta = 1$ in equation (32) it comes,

$$\Xi_T(s) = \int_0^{\infty} \exp(isT) \frac{1}{2\alpha\sqrt{2\pi}} \left(1 + \frac{1}{t}\right) \sqrt{\frac{1}{t}} . \exp\left[-\frac{1}{2\alpha^2}\left(t + \frac{1}{t} - 2\right)\right] dt.$$

At present, the r-th moment of the continuous random variable T is generally defined as

$$E(T^r) = \int_0^{\infty} t^r f_T(t) dt, \tag{33}$$

Moreover, the existence of moments depends on the convergence of the integral defined in equation (33). Regarding the Birnbaum – Saunders distribution, all positive and negative moments exist. For more details (Rieck (1999)). From the derivative of the characteristic function defined by (31), positive whole moments are obtained. Using the relationship established in equation (19), we obtain the moments of the Birnbaum-Saunders distribution. Thus, the random variable $\left(\frac{T}{\beta}\right)^r$ has as expected value the expression

$$E\left[\left(\frac{T}{\beta}\right)^r\right] = E\left(\frac{1}{2^r}\left(\alpha Z + \sqrt{(\alpha Z + 2)^2 - \frac{\alpha Z}{2}}\right)^{2r}\right)$$

$$= E\left(\left(\frac{\alpha Z}{2} + \sqrt{\frac{\alpha Z}{2} + 1}\right)^{2r}\right). \tag{34}$$

Applying Newton's binomial theorem to Equation (34), it comes

$$E\left[\left(\frac{T}{\beta}\right)^r\right] = \sum_{k=0}^{2r} \binom{2r}{k} E\left[\left(\frac{\alpha Z}{2}\right)^{2r-k} \left(\frac{\alpha Z}{2} + 1\right)^{\frac{k}{2}}\right] \tag{35}$$

Using parity on the exponent of equation(35) and by the further development of Newton's binomial, it will follow

$$E\left(\frac{T}{\beta}\right)^{r}=\sum_{\tau=0}^{r}\binom{2r}{2\tau}\sum_{k=0}^{\tau}\binom{\tau}{k}\left(\frac{\alpha}{2}\right)^{2(r-\tau+k)}E\left(Z^{2(r-\tau+k)}\right). \tag{36}$$

It suffices to remember that $Z\sim\mathcal{N}(0,1)$ to conclude that

$$E\left(Z^{2(r-\tau+k)}\right)=\frac{2(r-\tau+k)!}{2^{r-\tau+k}(r-\tau+k)!}.$$

From the above, it comes

$$E\left(\left(\frac{T}{\beta}\right)^{r}\right)=\sum_{\tau=0}^{r}\binom{2r}{2\tau}\sum_{k=0}^{\tau}\binom{\tau}{k}\frac{2(r-\tau+k)!}{2^{r-\tau+k}(r-\tau+k)!}\left(\frac{\alpha}{2}\right)^{2(r-\tau+k)}. \tag{37}$$

Then, by fixing r to the unit in equation (37), we obtain

$$E\left(\frac{T}{\beta}\right)=\frac{2}{21!}\binom{2}{0}\left(\frac{\alpha}{2}\right)^{2}+2\binom{2}{2}\sum_{k=0}^{1}\frac{k!}{2^{k}k!}\binom{1}{k}\left(\frac{\alpha}{2}\right)^{2k}$$

i.e.

$$E\left(\frac{T}{\beta}\right)=1+\frac{\alpha^{2}}{2}.$$

Consequently,

$$E(T)=\beta\left(1+\frac{\alpha^{2}}{2}\right). \tag{38}$$

Since $T\sim B(\alpha,\beta)$ and by setting r to 2 in equation (37), the variance of the variable T is calculated by the formula

$$Var(T)=E(T^{2})-\left(E(T)\right)^{2},$$

where,

$$E\left(\left(\frac{T}{\beta}\right)^{2}\right)=\sum_{\tau=0}^{2}\binom{4}{2\tau}\sum_{k=0}^{\tau}\binom{\tau}{k}\frac{2(2-\tau+k)!}{2^{2-\tau+k}(2-\tau+k)!}\left(\frac{\alpha}{2}\right)^{2(2-\tau+k)},$$

i.e.

$$E\left(\left(\frac{T}{\beta}\right)^2\right) = \frac{3}{2}\alpha^4 + 2\alpha^2 + 1.$$

Consequently,

$$Var(T) = (\alpha\beta)^2\left(1 + \frac{5}{4}\alpha^2\right). \tag{39}$$

According to equation (37), the mean and variance of the variable $T \sim BS(\alpha, \beta)$ are easily reduced. In addition, if the random variable $T \sim BS(\alpha, \beta)$ then the variable $T^{-1} \sim BS(\alpha, \beta^{-1})$. For more details (Birnbaum and Saunders, 1969a). From equations (38) and (39), we deduce the mean and the variance of the variable T^{-1}. i.e.

$$E\left(T^{-1}\right) = \beta^{-1}\left(1 + \frac{\alpha^2}{2}\right), \tag{40}$$

and

$$Var(T^{-1}) = (\alpha\beta^{-1})^2\left(1 + \frac{5}{4}\alpha^2\right). \tag{41}$$

4.4. Modelling Based on the Birnbaum-Saunders Distribution

A study by Rieck and Nedelman (1991) and Tsionas (2001) is the result of statistical modelling based on the Birnbaum–Saunders distribution. In addition, aspects related to influence diagnostics in log-Birnbaum-Saunders regression models with uncensored and censored data were examined by Galea et al. (2004), Leiva et al. (2007), and Xie and Wei (2007). This allowed Rieck and Nedelman (1991) to model Birnbaum-Saunders distribution using the Sinh-normal distribution.

As a reminder, the sinh-normal distribution is the result of the application of the sinh-arcsinh transformation to the normal distribution. However, Sinh-normal distribution therefore has a major role in data analysis where robustness to normality assumptions is a concern. The log-Birnbaum–Saunders distribution is a special case of the sinh-normal distribution. In addition, Rieck and Nedelman (1991) introduced the sinh-normal distribution (SHN) as a three-parameter symmetric model.

4.5. Log-Birnbaum–Saunders Distribution

A random variable Y follows a sinh-normal distribution law of three parameters α, η and σ if we have the following cumulative distribution function (CDF)

$$F_{Y(y;\alpha,\eta,\sigma)} = \Psi\big(\emptyset(y;\alpha,\eta,\sigma)\big);\ \ y \in \mathbb{R}. \tag{42}$$

Where $\Psi(\cdot)$ is the CDF of a standard normal random variable,

$$\emptyset(y;\alpha,\eta,\sigma) = \frac{2}{\alpha}\sinh\left(\frac{y-\sigma}{\eta}\right),$$

$\alpha > 0, \eta > 0,\ \sigma \in (-\infty,\infty)$, where $\sinh(x) = \frac{e^{x}-e^{-x}}{2}$ and $\sinh(x)$ is the hyperbolic sinus function of x. Note that $\emptyset(y;\alpha,\eta,\sigma)$ is an increasing function of y. Moreover, $\emptyset(y;\alpha,\eta,\sigma)$ has the same convergence as y when it tends towards $\pm\infty$. In addition, from a transformation simple

$$Y = \eta + \sigma\text{arcsinh}\left(\frac{\alpha Z}{2}\right),$$

we obtain the sinh-normal distribution, with $Z\sim\mathcal{N}(0,1)$ and we therefore note $Y\sim SHN(\alpha,\eta,\sigma)$. The probability density function of y is given by

$$f_{Y(y;\alpha,\eta,\sigma)} = \frac{2cosh\left(\frac{y-\eta}{\sigma}\right)}{\alpha\sigma}\phi\left(\frac{2}{\alpha}\sinh\left(\frac{y-\eta}{\sigma}\right)\right),\ \ y\in\mathbb{R}. \tag{43}$$

With $\alpha > 0$ the shape parameter, $\eta \in \mathbb{R}$ a location parameter. The reliability function and failure rate of $Y\sim SHN(\alpha,\eta,\sigma)$ are defined by

$$R_Y(y;\alpha,\eta,\sigma) = \Psi\left(-\frac{2}{\alpha}\sinh\left(\frac{y-\eta}{\sigma}\right)\right), \tag{44}$$

and

$$h_Y(y;\alpha,\eta,\sigma) = \frac{2\cosh\left(\frac{y-\eta}{\sigma}\right)\phi\left(\frac{2}{\alpha}\sinh\left(\frac{y-\eta}{\sigma}\right)\right)}{\alpha\sigma\Psi\left(-\frac{2}{\alpha}\sinh\left(\frac{y-\eta}{\sigma}\right)\right)}. \tag{45}$$

The distribution is strongly unimodal for $\alpha \leq 2$, and it is bimodal if $\alpha > 2$. As α increases, the distribution begins to accentuate its bimodality, with more separated modes, and its flattening is greater than the flattening of the normal distribution. Another important property of the distribution SHN is that if $Y \sim SHN(\alpha,\eta,2)$ proved that the random variable $T\sim BS(\alpha,\beta)$, Y is said to follow a log-Birnbaum-Saunders distribution.

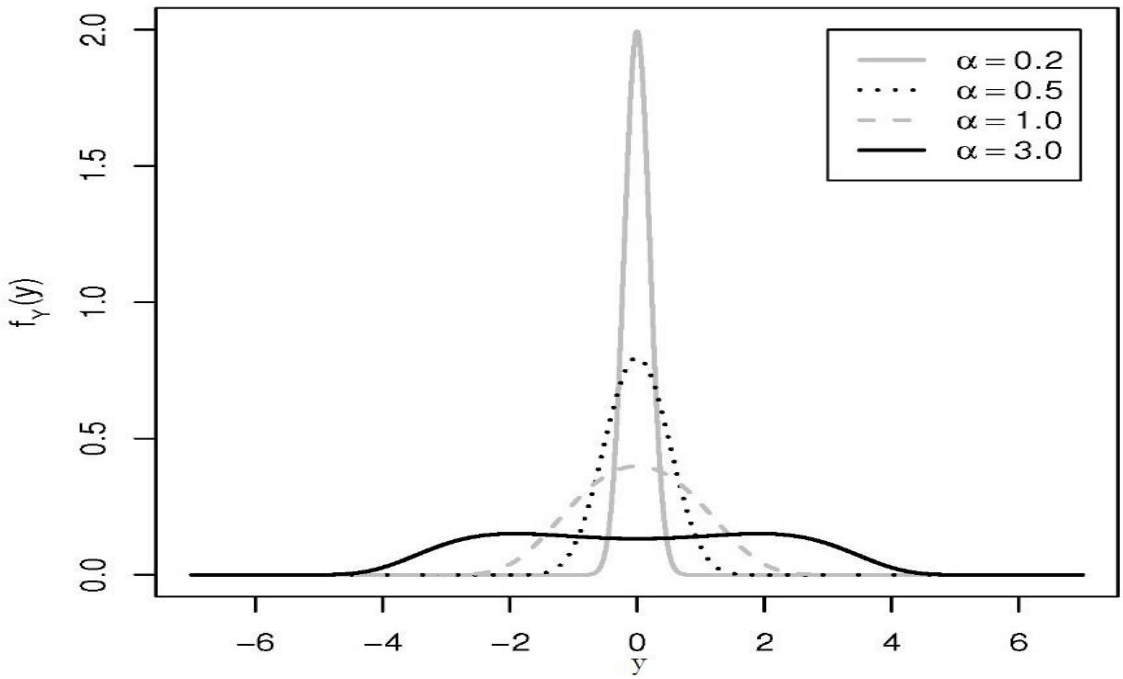

Figure 4.3. Plot of the log-Birnbaum–Saunders probability density function for the indicated values of $\alpha = 0.2, 0.5, 1.0, 3.0$ with $\eta = 0$.

4.6. Goodness of fit tests for the Birnbaum-Saunders Distribution

The identification of the distribution that best matches the data remains a major and complex problem of parametric statistical methods. However, the Kolmogorov-Smirnov test (KS), probability-probability (PP) and quantile-quantile (QQ) diagrams are general methods for the quality of the fit. Thus, for the selection of an appropriate statistical distribution, selection criteria are used. Life time data are described using probabilistic models with at least one of the following:

(1) a theoretical argument for the element failure mechanism

(2) a model that has already been used with a successful result

(3) a model that adapts well to data.

In the analysis of lifetime data, interest is sometimes focused on the lower or higher quantiles of the distribution. In the lifetime data, it has been verified that some families of lifetime distribution of two parameters, which have some flexibility, fit reasonably well in the central region, proven that a reasonable amount of data exists in the central part of the distribution. The lack of concentration of data at the extremities of an empirical distribution means that the adjustment of the lifetime distribution is often not significant. This thus leads to an observation of a considerable gap between certain life time distributions of similar characteristics. However, visual and adjustment methods

lead to disruption in identifying the appropriate lifetime distribution. In addition, the relevance of a life distribution can be based on life indicators. For more details, (Cox and Oates (1984)). Furthermore, these aspects are also to be considered:

- Study the failure rate of T, $h_T(t)$, or its logarithm, $\log(h_T(t))$, against t or $log(t)$.
- To analyze the cumulative failure rate, $H_T(t)$, or its logarithm, relative to t or $log(t)$.

In practice, when survival data are analyzed and a distribution is proposed to model them, a histogram is often developed. As is known, the histogram is an empirical estimate of the probabilistic density function of the data, which is simple and easy to develop. However, it is always practical to examine the data failure rate, as distributions with similar probability density functions may have different failure rates. This is the case, for example, of log-normal distributions, gamma distributions, inverse gauss distributions, Weibull and Birnbaum-Saunders. The problem is that it is not easy to empirically estimate the failure rate.

Therefore, one tool that makes this task easier is the Total Time on Test Graph (TTT). This graph allows us to detect the shape of the default rate of a random variable. The TTT function of the random variable T is given by

$$T(u) = \int_0^{F^{-1}(u)} [1 - F(x)]dx, u \in [0,1] \tag{46}$$

with, $F^{-1}(\cdot)$ is the inverse function of the cumulative distribution function of T. Moreover, the scaled TTT transform is given by the ratio

$$T_S = \frac{\int_0^{F^{-1}(u)} [1 - F(x)]dx}{\int_0^{F^{-1}(1)} [1 - F(x)]dx} \tag{47}$$

T_S can be approximate, allowing us to construct the TTT curve empirically by drawing the points $\left(k/n\,, T_s(k/n)\right)$, with

$$T_s\left(\frac{k}{n}\right) = \frac{\sum_{i=1}^{k} t_{i:n} + (n-k)t_{k:n}}{\sum_{i=1}^{n} t_{i:n}}, \qquad k = 1, \dots, n$$

where $t_{i:n}$ is the expressed value of the statistics of ith order, for $i = 1, \dots, n$. Consequently, a convex (or concave) TTT function corresponds to the class of an increasing failure

rate function (IFR) or the class of a decreasing failure rate function (DFR). In addition, the bathtub failure rate (BT) or bathtub inverse failure rate (IBT) has been defined as any concave (convex) TTT function to the left of the change point and convex (concave) to the right of the change point. A line graph of the TTT is a consequence of an exponential distribution due to the distribution underlying the constant failure rate. Despite all that has been done earlier in this part, one cannot simply consider the use of these theoretical arguments and lifetime indicators. Therefore, the choice of at least one probability model as a candidate for the life distribution to be used to describe the data must be logical and supported by either visual tools and statistical tests of data suitability or model selection criteria. Consequently, a concave (or convex) TTT function corresponds to the graph below.

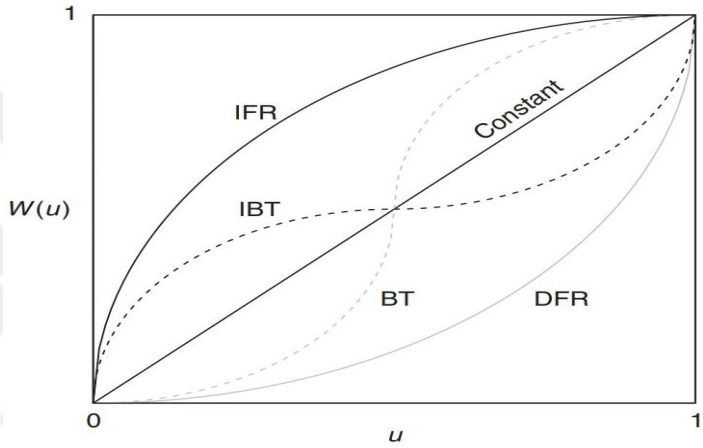

Figure 4.4. Theoretical TTT diagrams

The particular importance due to the comparison of data obtained from a number of sampling points is the usefulness of graphical methods as a moment-based adjustment. In addition, ratios without functionally independent dimension of the scale parameter occur through the particular utility of these methods in distributions within the scale family. Birnbaum – Saunders, gamma and log-normal distributions can be placed in support.

4.7. Two-Prameter Birnbaum Saunders Distribution

The two-parameter Birnbaum-Saunders (TBS) distribution was introduced by Kundu et al(2010). The two-variable random vector (T_1, T_2) follows a (TBS) distribution of parameters α_1, β_1, α_2, β_2, ρ if the cumulative distribution function of (T_1, T_2) is defined by

$$F(t_1, t_2) = \bar{\Phi}[\varepsilon_1(\bar{u}_1), \varepsilon_2(\bar{u}_2)], \tag{48}$$

here,

$$\begin{cases} \varepsilon_1(\bar{u}_1) = \frac{1}{\alpha_1}\left[\left(\frac{t_1}{\beta_1}\right)^{\frac{1}{2}} - \left(\frac{t_1}{\beta_1}\right)^{-\frac{1}{2}}\right] \\ \varepsilon_2(\bar{u}_1) = \frac{1}{\alpha_2}\left[\left(\frac{t_2}{\beta_2}\right)^{\frac{1}{2}} - \left(\frac{t_2}{\beta_2}\right) \right] \end{cases}$$

with, $(t_1, t_2) > (0,0)$, $(\alpha_1, \beta_1) > (0,0)$, $(\alpha_2, \beta_2) > (0,0)$ and $-1 < \tau < 1$. We will remember that $\bar{\Phi}(u, v; \tau)$ is the (cdf) of standard bivariate normal vector (Z_1, Z_2) with correlation coefficient ρ. However, the corresponding probability density function (pdf) is given by

$$f(t_1, t_2) = \frac{1}{8\pi\alpha_1\alpha_2\beta_1\beta_2}\Sigma(t)exp\left\{-\frac{1}{2(1-\tau^2)}\left[\left(\varepsilon(\bar{u}_1)\right)^2 + \left(\varepsilon(\bar{u}_2)\right)^2 - 2\tau\varepsilon(\bar{u}_1)\varepsilon(\bar{u}_2)\right]\right\}, \quad (49)$$

for, $(t_1, t_2) > (0,0)$, $(\alpha_1, \beta_1) > (0,0)$, $(\alpha_2, \beta_2) > (0,0)$ and $-1 < \rho < 1$ and,

$$\Sigma(t) = \left[\left(\frac{t_1}{\beta_1}\right)^{-\frac{1}{2}} + \left(\frac{\beta_1}{t_1}\right)^{\frac{3}{2}}\right]\left[\left(\frac{t_2}{\beta_2}\right)^{-\frac{1}{2}} + \left(\frac{\beta_2}{t_2}\right)^{\frac{3}{2}}\right].$$

5. ESTIMATION METHODS FOR THE PARAMETERS OF THE BIRNBAUM SAUNDERS

In practice, depending on the type of distribution, it is not as easy to estimate the parameters. This is exactly the case with the Birnbaum Saunders distribution. However, we discuss the maximum likelihood estimators, the moments and a modification of the moment estimators of a two-parameter Birnbaum–Saunders distribution. A simple graphical technique is used to estimate parameters and verify the quality of the adjustment of failure times following a Birnbaum-Saunders distribution. A bias-reduced estimation method is also presented. For a small sample size, the jackknife technique is considered another suitable method. We bring a synthesis of an EM algorithm.

5.1. Maximum Likelihood Estimators

Let T be a random variable having for random sample T_1, T_2...,T_n of size n and probability density function $f_T(\cdot)$. The T_1, T_2,....T_n are independent of each other. In addition, the likelihood function for a parameter θ is defined by

$$\mathcal{L}(\theta) = \prod_{i=1}^{n} \mathcal{L}_i(t_i, \theta), \tag{50}$$

with $\mathcal{L}_i(t_i, \theta) = f_T(t_i; \theta)$ the individual contribution of each observation to the likelihood function. However, the log-likelihood function for θ is given

$$\ell(\theta) = \sum_{i=1}^{n} \ell_i(\theta), \tag{51}$$

where $\ell_i(\theta) = \mathcal{L}_i(t_i, \theta)$ is the individual contribution of each observation to the log-likelihood function. Now let T_1, T_2,..., T_n be the random sample of size n of a population with CDF given by (24). Based on a random sample, the MLEs of the unknown parameters can be obtained by maximizing the log-likelihood function defined by equation (51), with

$$\sum_{i=1}^{n} \ell_i(\theta) = c + \frac{n}{\alpha^2} - \frac{1}{2\alpha^2} \sum_{i=1}^{n} \left(\frac{t_i}{\beta} + \frac{\beta}{t_i} \right) - n\log(\beta) + \sum_{i=1}^{n} \log(t_i + \beta), \tag{52}$$

and c a constant independent of θ. Using the first derivative of the log-likelihood function with respect to α and β leads to maximize it.

In this way,

$$\begin{aligned}\frac{\partial \ell}{\partial \alpha} &= \frac{1}{\alpha^3}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}\right)-\frac{n}{\alpha}-\frac{2n}{\alpha^3}\\ \frac{\partial \ell}{\partial \alpha} &= \frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta^2}-\frac{1}{t_i}\right)+\sum_{i=1}^{n}\frac{1}{t_i+\beta}-\frac{n}{2\beta}\end{aligned}, \tag{53}$$

Moreover, to obtain the equations of log-likelihood, we equal to zero the equations defined by the relation (53) to get,

$$\hat{\alpha} = \left(\frac{1}{n}\sum_{i=1}^{n}\left(\frac{t_i}{\hat{\beta}}+\frac{\hat{\beta}}{t_i}\right)\right)^{\frac{1}{2}}, \tag{54}$$

$$\hat{\beta} = \sqrt{\frac{\frac{1}{2\hat{\alpha}^2}\sum_{i=1}^{n} t_i}{\frac{1}{2\hat{\alpha}^2}\sum_{i=1}^{n}\frac{1}{t_i}+\frac{n}{2\hat{\beta}}-\sum_{i=1}^{n}\frac{1}{t_i+\hat{\beta}}}}. \tag{55}$$

A practical method for obtaining MLE joint $\hat{\alpha}$ and $\hat{\beta}$ is to use the arithmetic and harmonic means defined by

$$\mathcal{s} = \frac{1}{n}\sum_{i=1}^{n} t_i \text{ and } \mathcal{r} = \left[\frac{1}{n}\sum_{i=1}^{n} t_i^{-1}\right]^{-1}, \tag{56}$$

and consider the harmonic mean function defined by

$$k(y;\boldsymbol{t}) = \left[\frac{1}{n}\sum_{i=1}^{n}(y+t_i)^{-1}\right]^{-1}, y>0, \boldsymbol{t} = \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{pmatrix}^{\boldsymbol{t}}.$$

To obtain the maximum likelihood joint, it is necessary and sufficient to find the value of $\hat{\beta}$ as the solution of equation $g(y;t)=0$.

With $g(y;t) = y^2 - y[2\mathcal{r} + k(y;\boldsymbol{t})] + r[\mathcal{s} + k(y;\boldsymbol{t})]$ and calculate $\hat{\alpha} = \sqrt{\mathcal{s}/\hat{\beta} + \hat{\beta}/\mathcal{r} - 2}$.

Moreover, we note that $\mathcal{r} < \hat{\beta} < \mathcal{s}$. However, to solve $g(y;\boldsymbol{t}) = 0$, Two iterative methods are presented by Birnbaum and Saunders (1969b). Since the distribution of $\hat{\alpha}$ does not

depend on β, it is therefore possible to construct hypothesis tests and confidence intervals for α from $\hat{\alpha}$, where β is an unknown nuisance parameter. This result is observed for any $C_1 > 0$, and t, $g(C_1 y;\ C_1 t) = C_1^2 g(y;t)$, which implies that $\hat{\beta}$ is equivalent in the scale, and therefore $\hat{\alpha}$ is an invariant function in the scale of t. More precisely, $g(\hat{\beta}/\beta\,;\boldsymbol{t}/\beta) = 0$, and since the distribution of t/β does not depend on β, this is also true of the distribution of $\hat{\beta}/\beta$. Since $\hat{\alpha}$ is a function of $s/\hat{\beta}$ and $r/\hat{\beta}$; it follows that the distribution of $\hat{\alpha}$ does not depend on β. This result also follows from a general property of MLEs analogous to that given by J. Eastman and L. J. Bain (1973). Therefore, inference procedures on α with unknown β are possible.

5.1.1. Maximum Likelihood Method for Discrete Variable

The maximum likelihood method is a parametric estimation method that owes its popularity to

- the simplicity of its approach,
- its ability to adapt to complex modelling, i.e. the rule $\mathcal{L}(\theta)$ where θ symbolizes a multitude of unknown parameters : $\theta = (\theta_1, \theta_2, \theta_3 \dots)^t$,
- the digital aspect accessible through the application of known optimization methods.

It allows to:

- build efficient estimators,
- Build accurate confidence intervals
- implement powerful statistical tests

Moreover, the maximum likelihood method was initially called the absolute criterion (in 1912), a heavy-meaning expression, even a century later.

Now consider T a discrete variable according to the $\mathcal{L}(\theta)$ law with θ an unknown parameter. It is up to us to remember that we want to estimate $t_1, \dots, t_n$ from the data. The vector of the $(t_1, \dots, t_n)$ data is a realization of a n-sample $(T_1, \dots, T_n)$ of T. Thus, the maximum likelihood method is based on the following idea:

- the fact of having observed the $t_1, \dots, t_n$ value is not surprising, either:
- the assumption of observing $t_1, \dots, t_n$ values rather than others was most likelihood.

Thus, θ is considered a real variable and the values of θ, which

- make the observation of $t_1, \ldots, t_n$ values as likely as possible, that is
- maximize the chances of achieving the event $\{(T_1, \ldots, T_n) = (t_1, \ldots, t_n)\}$, that is:
- maximize the probability $\mathrm{P}_0\big((T_1, \ldots, T_n) = (t_1, \ldots, t_n)\big)$

Such a value is therefore a solution to an optimization problem aimed at maximizing a function of θ characterizing the likelihood of having obtained $t_1, \ldots, t_n$. Thus, θ^* is a point estimator of unknown parameter θ called the maximum likelihood estimator. So, we specify that we can define the likelihood of $t_1, \ldots, t_n$ data by the function of θ:

$$L(t_1, \ldots, t_i, \ldots, n; \theta) = \mathrm{P}_0\big((T_1, \ldots, T_n) = (t_1, \ldots, t_n)\big). \tag{57}$$

As $(T_1, \ldots, T_n)$ is a n-sample, by the independence and the identical distribution (iid) of the $T_1, \ldots, T_n$ variables, we can also write:

$$L(t_1, \ldots, t_i, \ldots, n; \theta) = \mathrm{P}_0\left(\bigcap_{i=1}^{n}\{T_i = t_i\}\right) = \prod_{i=1}^{n}\mathrm{P}_0(T_i = t_i). \tag{58}$$

5.1.2. Maximum Likelihood Method for Continuous Variables

Maximum likelihood estimation is a common statistical method used to derive the parameters of the probabilistic distribution of a given sample. This method was developed by statistician and geneticist Ronald Fisher from 1912 to 1922. The maximum likelihood estimator may exist and be unique, not unique or non-existent. If T is a density variable according to the $\mathcal{L}(\theta)$ law of $f_\theta(t)$ density, the reasoning developed above still holds. However, the likelihood of the data can no longer be measured by the function defined by equation (57) because it is now null. Now, instead of $\{(T_1, \ldots, T_n) = (t_1, \ldots, t_n)\}$, an idea is to introduce a not insignificant near:

$\{(T_1, \ldots, T_n) \in [t_1, t_1 + o_1[\times \ldots \times [t_n, t_n + o_n[\}$ with $o_1, \ldots, o_n$ rather small. This leads us

to measure the likelihood of the data by the function of θ:

$$L^{(o_1,\ldots,o_n)}(\theta) = \mathbb{p}_0\big((T_1, \ldots, T_n) \in [t_1, t_1 + o_1[\times \ldots \times [t_n, t_n + o_n[\big). \tag{59}$$

Thus, by noting $f_\theta(t)$ a density of $(T_1, \ldots, T_n)$, we have

$$L^{(o_1,\ldots,o_n)}(\theta) = \int_{x_1}^{x_1+o_1} \ldots \int_{x_i}^{x_i+o_i} \ldots \int_{x_n}^{x_n+o_n} f_\theta(t_1, t_2, \ldots, t_n; \theta)\, dt_1 \ldots dt_n.$$

However, determining a θ that maximizes $L^{(o_1,\ldots,o_n)}(\theta)$ is not an easy ting:

- the analytical expression of an integral function such as $L^{(o_1,\ldots,o_n)}(\theta)$ does not always exist,
- it depends on $(o_1, \ldots, o_n)$ whose smallness remains subjective.

Therefore, one solution is to consider an idealized version of $L^{(o_1,\ldots,o_n)}(\theta)$ that is

$$L^{(o_1,\ldots,o_n)}(\theta) = f_\theta(t_1, t_2, \ldots, t_n). \tag{60}$$

The Equation (60) is a consequence of the result which, by noting $F_\theta(\cdot)$ the distribution function of $(T_1, \ldots, T_n)$, it comes

$$\lim_{\mathrm{O}\to\mathbb{O}} \frac{L^{(o_1,\ldots,o_n)}(\theta)}{o_1 \times \ldots \times o_n} = \frac{\partial^n}{\partial o_1 \ldots \partial o_n} F_\theta(t_1, t_2, \ldots, t_n) = f_\theta(t_1, t_2, \ldots, t_n), \tag{61}$$

where, $\mathrm{O} = \begin{pmatrix} o_1 \\ \vdots \\ o_n \end{pmatrix}^t$ et $\mathbb{O} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}^t$.

Indeed, we call the likelihood of θ given the observations $(t_1, \ldots, t_i, \ldots, t_n)$ of an independently and identically distributed n-sample (iid) according to distribution $\mathcal{L}(\theta)$, the number:

$$L(t_1, \ldots, t_i, \ldots, n; \theta) = f(t_1; \theta) \times f(t_2; \theta) \times \ldots \times f(t_n; \theta) = \prod_{i=1}^{n} f(t_i; \theta). \tag{62}$$

We try to find the maximum of this likelihood so that the probabilities of the observed realizations are also maximum. This is a matter of optimization. As a general rule, we use the fact that if $L(t_1, \ldots, t_i, \ldots, n; \theta)$ is differentiable and accepts an overall maximum in $\theta = \theta^*$ value, then the first derivative cancels out in $\mathcal{L}(\theta)$ and the second derivative remains negative.

Conversely, if this takes place, then $\theta = \theta^*$ is a local (and not global) maximum of $L(t_1, \ldots, t_i, \ldots, n; \theta)$. We must ensure that it is indeed a global maximum. Since likelihood is

positive and logarithm is an increasing function, it is equivalent and often easier to maximize the Nepean logarithm of likelihood. Moreover, We call the estimator of the maximum likelihood of θ for $(t_1, \ldots, t_i, \ldots, t_n)$ a real θ^* that maximizes the likelihood function defined by (62) in θ, i.e. for any θ,

$$L(t_1, \ldots, t_i, \ldots, n; \theta) \leq L(t_1, \ldots, t_i, \ldots, n; \theta^*).$$

Subsequently an alternative expression is:

$$\theta^* \in \underset{\theta}{\operatorname{argmax}}\, L(t_1, \ldots, t_i, \ldots, t_n; \theta)$$

where $\operatorname{argmax}(\cdot)$ denotes the maximum argument which is the set of points in which an expression reaches its maximum value. Since it depends on $t_1, \ldots, t_i, \ldots, t_n$, θ^* is a point estimate of θ.
Moreover, we talk about the local maximum if the sufficient condition is met at the critical point $\theta = \theta^*$.

i.e.

$$\begin{cases} \dfrac{\partial^2 L(t_1, \ldots, t_i, \ldots, n; \theta)}{\partial \theta^2} < 0 \\ \dfrac{\partial^2 \ln L(t_1, \ldots, t_i, \ldots, n; \theta)}{\partial \theta^2} < 0 \end{cases}. \qquad (63)$$

The estimator obtained using the maximum likelihood method is considered to be:

- convergent,
- asymptotically efficient, it reaches the Cramér-Rao bound,
- asymptotically distributed according to a normal distribution,

However, it can be biased into a finite sample.

Since the maximum likelihood estimator is asymptotically normal, a confidence interval Ic_n can be constructed so that it contains the true parameter with a probability $1 - \alpha$ where

$$Ic_n(T_1, \ldots, T_n) = \left[\theta^* - z_\alpha \frac{1}{[I_n(\theta^*)]^{1/2}}, \theta^* + z_\alpha \frac{1}{[I_n(\theta^*)]^{1/2}}\right].$$

5.2. Method of Moments Estimation

Introduced by Pearson (1894), the technique of the parameters estimation of a statistical model that is used to find the values of the parameters causing in match between the sample moments and population moments is called the "method of moments. Its advantage is to adjust a simple mixing model. Having superior theoretical properties in many contexts, the moment method is considered a poor relation of the maximum likelihood estimate. Because of their conceptual and computational simplicity, the moment method and its generalizations continue to be useful in practice for some types of estimation problems. For the usual moment estimators in a two-parameter case, the moment estimators of α and β are established by equating the sample mean and variance, respectively, with $E(T)$ and $Var(T)$, defined to relations (38) and (39).

Further, if the mean and variance of the sample are referred to as s and r respectively, the resolution of the equations below gives the moment estimators of α and β.

i.e.

$$s = \beta\left(1 + \frac{\alpha^2}{2}\right) \quad \text{and} \quad r = (\alpha\beta)^2\left(1 + \frac{5}{4}\alpha^2\right). \tag{64}$$

From the nonlinear equation

$$\alpha^4(5s^2 - r) + 4\alpha^2(s^2 - 4) - 4 = 0, \tag{65}$$

one can be obtained easily the moment estimator of α as the root of the equation (65). It is clear to us that the moment estimate of α is judged by the coefficient of variation of the sample about the value $\sqrt{5}$.

However, when the sample variation coefficient is greater than $\sqrt{5}$, the moment estimator of α may not exist. Thus Ng et al. (2003) proposed Modified Moments (MMEs) estimators of parameters α and β using the fact that if $T \sim BS(\alpha, \beta)$, then $T^{-1} \sim BS(\alpha, \beta^{-1})$.

Thus, the assimilation of arithmetic and harmonic means respectively s and r of the sample to the corresponding population versions is

$$\tilde{\alpha} = \left\{2\left[\left(\frac{s}{r}\right)^{1/2} - 1\right]\right\}^{\frac{1}{2}}, \tag{66}$$

$$\tilde{\beta} = \sqrt{sr}.$$

5.3. Modified Moments Estimates Method

When the parameters of the Birnbaum–Saunders distribution are estimated using the standard moment method, the moment estimators cannot exist or be unique. Another way to estimate the parameters of the Birnbaum – Saunders distribution is to use the modified moments method, which provides estimators that exist always and are unique; (Ng et al. (2003)). The modified estimation method is a special case of the generalized timing estimation method. For more details on the generalized moment estimation method, (Mátyás (1999), Santos-Neto et al. (2014)) for the application of this method to a reparameterized version of the Birnbaum–Saunders distribution proposed by Santos-Neto et al. (2012); (Leiva et al. (2014c)). Consider the random sample defined in subsection (5.1.2) and designate by $\tilde{\alpha}$ and $\tilde{\beta}$ the modified moment estimates of α and β given respectively by

$$\tilde{\alpha} = \sqrt{2\left(\sqrt{\frac{s}{r}} - 1\right)} \quad \text{and} \quad \tilde{\beta} = \sqrt{sr}\,. \tag{67}$$

With s and r are defined by equation (56). This can be defined during the bivariate normality of the asymptotic joint distribution of the modified moment estimators $\tilde{\alpha}$ and $\tilde{\beta}$ by

$$\sqrt{n}\left[\begin{pmatrix}\tilde{\alpha}\\ \tilde{\beta}\end{pmatrix} - \begin{pmatrix}\alpha\\ \beta\end{pmatrix}\right] \sim \mathcal{N}\left(\begin{pmatrix}0\\ 0\end{pmatrix}, \begin{pmatrix}\frac{\alpha^2}{2} & 0\\ 0 & (\alpha\beta)^2\frac{4+3\alpha^2}{(2+\alpha^2)^2}\end{pmatrix}\right).$$

However, the asymptotic marginal distributions of $\tilde{\alpha}$ and $\tilde{\beta}$, when $n \to \infty$, are defined respectively by

$$\begin{aligned}\sqrt{n}(\tilde{\alpha} - \alpha) &\sim \mathcal{N}\left(0, \frac{\alpha^2}{2}\right),\\ \sqrt{n}(\tilde{\beta} - \beta) &\sim \mathcal{N}\left(0, (\alpha\beta)^2\frac{4+3\alpha^2}{(2+\alpha^2)^2}\right),\end{aligned} \tag{68}$$

We can therefore construct confidence intervals and test hypotheses for the α and β parameters of the Birnbaum–Saunders distribution from the asymptotic results given in equation (68).

5.4. Graphic Estimate Method

For the distribution of Birnbaum–Saunders, a simple graphical method analogous to probabilistic diagrams was developed by Chang and Tang (1994). Such a graphical method serves as an adequate visual tool for adjustment and can also be used as a method of estimating the parameters of the Birnbaum – Saunders distribution. Initial values of the necessary iterative procedure in the maximum likelihood method can be found by this method. Using a simple linear regression method, the slope and ordinate at the origin of the line are estimated. However, having considered the cumulative distribution function $F_T(\cdot)$ of a random variable T with the Birnbaum–Saunders distribution defined by equation (21), we have

$$\begin{aligned} t &= \alpha\sqrt{\beta}\sqrt{t}\Psi^{-1}[F_T(t)] + \beta \\ &= \sqrt{\beta}\sqrt{t}\varphi\left(\frac{\beta}{t}\right) + \beta. \end{aligned} \tag{69}$$

For the probabilistic plot, it is not easy to derive the linear function about t from equation (69). This is how it can be described

$$\mathrm{p} = \sqrt{t}\,\varphi\left(\frac{\beta}{t}\right), a = \sqrt{\beta} \text{ et } b = \beta, \tag{70}$$

with α the ordinate at the origin and β the slope of the linear function defined by

$$y = ax + b, \tag{71}$$

where the x-axis is defined by $x = p$, while that of the y-axis is presented by $y = t$.

Now consider $T_1, T_2, \cdots, T_n$ a sample of size n from $T \sim BS(\alpha, \beta)$ with $t_1, t_2, \cdots, t_n$. observations. When data are from a Birnbaum–Saunders distribution, the t_i depending on $\dot{p}_i$ plot displays a result similar to a straight line.

Where,

$$\dot{p}_i = \sqrt{t}\Psi^{-1}\left(\dot{F}_T(t_i)\right) \text{ et } \dot{F}_T(t_i) = \frac{i - 0.3}{n + 0.4}, i = 1,2, \cdots, n,$$

being the Average rank. For more details (Chang and Tang (1994)). The coefficient of determination R^2 can be used to analyze and visually study the quality of the Birnbaum–Saunders distribution fit.

5.5. Bias-Reduced Estimation Method

A multiple analysis focused on the Monte Carlo simulation study, allowed us to conclude that the MM and ML estimators have almost the same appearance in terms of bias and mean square error. This is distinguished by small values of α. In addition, we observe according to the study the model of bias of MLE and MME This one

$$\text{Bias}(\hat{\alpha}) \approx \text{Bias}(\tilde{\alpha}) \approx -\frac{\alpha}{n}, \tag{72}$$

$$\text{Bias}(\hat{\beta}) \approx \text{Bias}(\tilde{\beta}) \approx \frac{\alpha}{4n}.$$

Now, almost unbiased maximum likelihood estimators (UMLE noted by $\hat{\tilde{\alpha}}\ et\ \hat{\tilde{\beta}}$) and almost unbiased modified moment estimators (UMME designated by $\hat{\tilde{\alpha}}\ et\ \hat{\tilde{\beta}}$) are constructed through a standard bias reduction method.

Thus, let us designate by $\hat{\alpha}$ and $\hat{\beta}$ respectively the estimators cited above to define the bias-reduced estimators by

$$\hat{\tilde{\alpha}} = \left(1 + \frac{1}{n-1}\right)\hat{\alpha} \ \text{ et } \ \hat{\tilde{\beta}} = \left(1 + \frac{\alpha^2}{4n}\right)^{-1}\hat{\beta}\,, \tag{73}$$

$$\hat{\tilde{\alpha}} = \left(1 + \frac{1}{n-1}\right)\tilde{\alpha} \ \text{ et } \ \hat{\tilde{\beta}} = \left(1 + \frac{\alpha^2}{4n}\right)^{-1}\tilde{\beta}. \tag{74}$$

From equation (71) we deduce the asymptotic joint distribution of $\hat{\tilde{\alpha}}$ and $\hat{\tilde{\beta}}$ which is bivariate normal and defined by

$$\begin{pmatrix}\hat{\tilde{\alpha}}\\ \hat{\tilde{\beta}}\end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix}\alpha\\ \beta\end{pmatrix}, \begin{pmatrix}\frac{n\alpha^2}{2(n-1)^2} & 0\\ 0 & \frac{16n(\alpha\beta)^2(4+3\alpha^2)}{(4n+\alpha^2)^2(2+\alpha^2)^2}\end{pmatrix}\right). \tag{75}$$

5.6. Bias-Corrected Estimate

In this section, we illustrate the principle of correcting bias estimation in the first-order dynamic panel data model. Here, the dependent variable y_{it} is determined by the offset value of a period of the dependent variable $y_{i,t-1}$, the additional regression x_{it}, the individual-specific effect not observed ε_i and a general disturbance term O_{it}.

$$y_{it} = \delta y_{i,t-1} + \beta x_{it} + \varepsilon_i + O_{it}, \qquad i = 1, \dots, N \; ; \; t = 1, \dots, T, \tag{76}$$

where, $O_{it} \sim \mathcal{N}(0, \sigma^2)$.

The regression x_{it} can be correlated with the specific individual ε_i effect, but we assume that it is strictly exogenous at the end of the general error O_{it}.

Given the start-up observations of y_{i0}, we assume that they are not correlated with the error terms O_{it}. In other words, it is unnecessary to assume that the model (76) is dynamically stable. The introduction of each variable on the average of its specificity leads to the elimination of unknown individual effects.

Let $\ddot{y}_{it} = y_{it} - \bar{y}_i, \ddot{y}_{i,t-1} = y_{i,t-1} - \bar{y}_{i,t-1}, \ddot{x}_{it} = x_{it} - \bar{x}_i$ et $\ddot{O}_{it} = O_{it} - \bar{O}_i$. it will follow that, equation (76) becomes

$$\ddot{y}_{it} = \delta \ddot{y}_{i,t-1} + \lambda \ddot{x}_{it} + \ddot{O}_{it}, \qquad t = 1, \dots, T; i = 1, \dots, N. \tag{77}$$

Using ordinary least squares (OLS), the calculation of the dummy least squares variable (DLSV) estimator leads us to

$$\hat{\delta}_{dlsv} = \frac{\sum\sum \ddot{x}_{it}^2 \sum\sum \ddot{y}_{i,t-1}\ddot{y}_{it} + \sum\sum \ddot{y}_{i,t-1}^2 \sum\sum \ddot{x}_{it}\ddot{y}_{it}}{\sum\sum \ddot{x}_{it}^2 \sum\sum \ddot{y}_{i,t-1}^2 - \left(\sum\sum \ddot{x}_{it}\ddot{y}_{t,t-1}\right)^2}, \tag{78}$$

and

$$\hat{\lambda}_{dlsv} = \frac{\sum\sum \ddot{y}_{i,t-1}^2 \sum\sum \ddot{x}_{it}\ddot{y}_{it} - \sum\sum \ddot{x}_{it}\ddot{y}_{i,t-1} \sum\sum \ddot{y}_{i,t-1}\ddot{y}_{it}}{\sum\sum \ddot{x}_{it}^2 \sum\sum \ddot{y}_{i,t-1}^2 - \left(\sum\sum \ddot{x}_{it}\ddot{y}_{t,t-1}\right)^2}, \tag{79}$$

where the double summations are for $i = 1, \dots, N$ et $t = 1, \dots, T$.

Due to the correlation between δ and λ, the DLSV estimators of $\ddot{y}_{t,t-1}$ and $\ddot{O}_{it}$ are biased and inconsistent for fixed T. The extent of the inconsistency can therefore be calculated as follows. Equations (78) and (79) become

$$\hat{\delta}_{dlsv} = \delta + \frac{\sum\sum \tilde{x}_{it}^2 \sum\sum \tilde{y}_{i,t-1}\ddot{O}_{it} - \sum\sum \ddot{x}_{it}\ddot{y}_{i,t-1} \sum\sum \ddot{x}_{it}\ddot{O}_{it}}{\sum\sum \ddot{x}_{it}^2 \sum\sum \ddot{y}_{i,t-1}^2 - \left(\sum\sum \ddot{x}_{it}\ddot{y}_{t,t-1}\right)^2} \tag{80}$$

and

$$\hat{\lambda}_{\text{dlsv}} = \lambda - \frac{\sum\sum \tilde{x}_{it}\tilde{y}_{i,t-1} \sum\sum \tilde{y}_{i,t-1}\ddot{O}_{it} - \sum\sum \ddot{y}^2_{i,t-1} \sum\sum \ddot{x}_{it}\ddot{O}_{it}}{\sum\sum \ddot{x}^2_{it} \sum\sum \ddot{y}^2_{i,t-1} - \left(\sum\sum \ddot{x}_{it}\ddot{y}_{i,t-1}\right)^2}. \tag{81}$$

By using continuous substitution and under equation (77), we obtain

$$y_{it} = \delta^t y_{i0} + \lambda\xi + \frac{1-\delta^t}{1-\delta}\varepsilon_i + \gamma, \tag{82}$$

where $\xi = x_{it} + \delta x_{i,t-1} + \cdots + \delta^{t-1}x_{i1}$ et $\gamma = O_{it} + \delta O_{i,t-1} + \cdots + \delta^{t-1}O_{it}$.

5.7. EM Algorithm

Along this section, we fix $\Theta \subset \mathbb{R}^d$ the parameter space. That is, $\{Q(.,\theta'), \theta' \in \Theta\}$ family of auxiliary functions which is the focal point of the EM algorithm. By notation abuse, we use T both to refer to this restriction and the subspace of unobservable variables. Let $Y = (Y_1, \ldots, Y_i, \ldots, Y_n)$ a n-sample generated by a statistical model of likelihood function $g(Y, \theta)$ and dependent on unobservable variables noted also $T = (T_1, \ldots, T_i, \ldots, T_q)$. Consideration is given to the likelihood of complete data (T, Y) rated $g_c(T, Y; \theta)$. The intermediate quantity or Q-function of the EM algorithm is the function defined by:

$$Q(\theta, \theta') = E\left[\log\left(\mathrm{g_c}(T, Y; \theta)\right)|Y; \theta'\right], \tag{83}$$

where the expectation is evaluated under the conditional density of T knowing Y, at θ' fixed. Generally speaking, EM breaks down into two stages:

a. Expectation, calculate $Q\left(\theta, \theta^{(k)}\right) = E\left[\log \mathrm{g_c}(X, Y; \theta)\middle|Y; \theta^k\right]$
b. Maximization $\theta^{(k+1)} = \arg\max_{\theta\in\Theta} Q\left(\theta, \theta^{(k)}\right)$.

Thus, we build the sequence $\{\theta^{(k)}\}_{k\geq 1}$ with the data of $\theta^{(0)}$ such that the incomplete likelihood is improved at each iteration. Furthermore, for any $Q\left(\theta^{(k+1)}, \theta^{(k)}\right) \geq Q\left(\theta^{(k)}, \theta^{(k)}\right)$ which leads

$$L\left(\theta^{(k+1)}\right) \geq L\left(\theta^k\right). \tag{84}$$

This property of monotony sums up in itself a first attraction that can arouse the EM. Replace a quasi-impossible maximization with a series of model parameter estimators that increase the likelihood of the model to a given a priori precision.

Algorithm 1: Expectation-maximization

1 : Choose an initial parameter $\theta^{(0)}$

2 : For $k = 0,2, \ldots$ do

- Step E : Evaluate $Q\left(\theta, \theta^{(k)}\right) = E\left[\log(\mathrm{g_c}(X, Y; \theta)\middle| Y; \theta^k)\right]$
- Step M : Find $\theta^{(k+1)} = \arg\max_{\theta \in \Theta} Q\left(\theta, \theta^{(k)}\right)$

3 : End for.

In the interest of conservatism, we make some assumptions to establish the basis for the above and to support the rest of the presentation. However,

a. The parameter space Θ is an open set of $\mathbb{R}^d$ with $d \in \mathbb{N}^*$;
b. For any $\theta \in \Theta, L(\theta) \in]0, +\infty[$;
c. $\theta \mapsto L(\theta)$ is continuously differentiable on Θ ;
d. For any $(\theta, \theta') \in \Theta \times \Theta$,

$$\int_X |\nabla_\theta \log(P(t|y,\theta)|P(t|y,\theta'))dt < +\infty, \tag{85}$$

with ∇_θ is the gradient operator applied to the point $\theta = \theta'$.

e. For any $\theta' \in \Theta, \theta' \mapsto \mathcal{H}(\theta, \theta')$ is continuously differentiable on Θ.

where $\mathcal{H}(\theta, \theta') = E[\log(P(X|Y,\theta)[Y,\theta')]$. Now for any $\theta \in \Theta, \theta \mapsto Q\left(\theta, \theta^{(k)}\right)$ is continuously differentiable and we have:

$$\nabla_\theta Q\left(\theta, \theta^{(k)}\right)\Big|_{\theta=\theta^{(k)}} = \nabla_\theta /(\theta)|_{\theta=\theta^{(\theta)}}, \text{ with } \theta^{(k)} \text{ is fixed.} \tag{86}$$

Proof :

Note now that the equation defined by (64) is written $\ell(\theta) = Q\left(\theta, \theta^{(k)}\right) - \mathcal{H}\left(\theta, \theta^k\right)$ and by combining the points (c) and (e) assumptions (63) function $\theta \mapsto Q\left(\theta, \theta^{(k)}\right)$ is differentiable as the difference between two differentiable functions and we have

$$\nabla_\theta /(\theta)\rfloor_{\theta=\theta^k} = \nabla_\theta Q\left(\theta, \theta^{(k)}\right)\Big|_{\theta=\theta^{(k)}} - \nabla_\theta \mathcal{H}\left(\theta, \theta^{(k)}\right)\Big|_{\theta=\theta^{(k)}}. \tag{87}$$

It is then sufficient that the second right-hand member of the equality (87) is null to obtain the desired equality. Using Jensen inequality, the $\theta \mapsto \mathcal{H}\left(\theta, \theta^{(k)}\right)$ function is maximum in $\theta = \theta^{(k)}$.

Consequently,

$$\nabla_\theta \mathcal{H}\left(\theta, \theta^{(k)}\right)\Big|_{\theta=\theta^{(k)}} = 0.$$

This second property characterizes the stability points of the EM algorithm as stationary points, i.e. of $\theta^{(*)}$'s such as

$$\nabla_\theta /(\theta)|_{\theta=\theta^{(*)}} = 0.$$

Note that some additional difficulties may remain in phase E or rather in phase M. At phase M level, it may be difficult to find the parameter $\theta^{(\theta+1)}$ that achieves the maximum of the objective function $Q(\theta, \theta^{(k)})$.

This is the reason why Wu (1983) proposed to choose $\theta^{(k+1)}$ so that the following relation is satisfied:

$$Q\left(\theta^{(k+1)}, \theta^{(k)}\right) \geq Q\left(\theta^{(k)}, \theta^{(k)}\right). \quad (88)$$

This gives rise to the generalized EM algorithm. Indeed, Inequality (88) is a sufficient condition to satisfy the monotony of the EM. In other words, the likelihood does not decrease after a generalized EM iteration. Thus, instead of a local or global maximum, we use only this criterion improving the objective function. The summary of this algorithm is given in the following table

---

Algorithm 1: Generalized Expectation-Maximization

1 : Choose an initial parameter $\theta^{(0)}$

2 : For $k = 0,1,2,\ldots$ do

- Step E : Evaluate $Q\left(\theta, \theta^{(k)}\right) = E\left[\log g_c(T, Y; \theta)\middle| Y; \theta^k\right]$
- Step M : Find $\theta^{(k+1)}$ such as $Q(\theta^{(k+1)}, \theta^{(k)}) \geq Q(\theta^{(k)}, \theta^{(k)})$

3 : And for.

---

of distributions which is frequent in classification. Moreover, it is very often for step E to be very difficult in terms of the form of full likelihood. A simple way around this problem is to replace theoretical expectation with its empirical counterpart. At the moment k, is supposed to have a n-sample $(T_1, T_2, \ldots, T_n)$ from the density of the missing data $p(\cdot\,|Y\,, \theta^{(k-1)})$.

Wei et Tanner (1990) recommend replacing the evaluation of $Q(\theta, \theta^{(k-1)})$ with its Monte Carlo approximation:

$$\hat{Q}_n\left(\theta, \theta^{(k-1)}\right) = \frac{1}{n}\sum_{i=1}^{n} \log[g_c(T_i, Y, \theta)] \tag{89}$$

then maximize it in step M. This approximation admits convergence properties in particular under certain hypotheses including that which makes that $\{T_i\}$ forms a Markov chain of invariant law the conditional distribution of T knowing $T = t$

$$\hat{Q}_n\left(\theta, \theta^{(k-1)}\right) \to Q\left(\theta, \theta^{(k-1)}\right) \tag{90}$$

almost certainly, as $n \to +\infty$. Note the particular case of $n = 1$ called Stochastic EM (SEM). The latter approach in the direction of minimizing the risk of falling on local optimum with an additional step Stochastic for the basic EM. For a detailed presentation, see Celeux and Diebolt (1992). In addition, it should be noted that there are several ways to simulate sample $(T_1, T_2, \ldots, T_n)$, notably the Gibbs sampler (Chan and Ledolter (1995)) or a sequential version of the Monte Carlo methods.

5.8. Jaknife Estimation Method

Based on the sequential suppression of a t_i sampling point, the jackknife is a resampling method. It is requested to recalculate the MLE and MME from the sample deduced from size $n - 1$.

Moreover, the jackknife technique was introduced by Quenouille and Maurice H between 1949 and 1956. Thus, this technique took its name in 1958 proposed by Tukey, John W. Now we remove the

point t_m from the data set, then recalculate r and s along with the harmonic mean function which we denote by K.

$$s_{(m)} = \frac{1}{n-1}\sum_{i=1}^{n} t_i = \frac{ns - t_m}{n-1}, \qquad \text{pour tout } i \neq j$$

$$r_{(m)} = \left[\frac{1}{n-1}\sum_{i=1}^{n} t_i^{-1}\right]^{-1} = \frac{nr - t_m^{-1}}{n-1},$$

$$K_{(m)}(\mathrm{v}) = \left[\frac{1}{n-1}\sum_{i=1}^{n} (\mathrm{v} - t_i)^{-1}\right] = \frac{nK(\mathrm{v}) - (\mathrm{v} - t_i)^{-1}}{n-1}.$$

Now let $G(\beta)$ be the function defined by

$$G(\beta) = \beta^2 - \left(2r_{(m)} + K_{(m)}(\beta)\right)\beta + \left(s_{(m)} + K_{(m)}(\beta)\right)r_{(m)}.$$

Thus, solving the equation

$$\beta^2 - \left(2r_{(m)} + K_{(m)}(\beta)\right)\beta + \left(s_{(m)} + K_{(m)}(\beta)\right)r_{(m)} = 0,$$

we will have $\hat{\beta}_{(m)} = \sqrt{s_{(m)}r_{(m)}}$ as the only positive root and

$$\hat{\alpha}_{(m)} = \left(\frac{s_{(m)}}{\hat{\beta}_{(m)}} + \frac{\hat{\beta}_{(m)}}{r_{(m)}} - 2\right)^{\frac{1}{2}}.$$

Therefore, the Jacknife estimators are defined as follows

$$\hat{\tilde{\alpha}}_{(m)} = \left\{2\left[\left(\frac{s_{(m)}}{r_{(m)}} - 1\right)^{\frac{1}{2}}\right]\right\} \text{ et } \hat{\tilde{\beta}}_{(m)} = \sqrt{s_{(m)}r_{(m)}}$$

6. GENETIC ALGORITHM FOR PARAMETER ESTIMATION OF TWO-PARAMETER BIRBAUM-SAUNDERS DISTRIBUTION

The use of genetic algorithms (GA) to solve problems is the origin of research by John Holland (1960) and his colleagues and students at the University of Michigan who have proposed genetic algorithms. This new research group's novelty was the addition of the cross-over operator to their model of mutation. This operator is often used to find the optimum of a function by combining the genes of different individuals in the population. The first result of this research was the publication of "Adaptation in Natural and Artificial Systems" (Holland, 1975). GAs are often used in machine learning and optimization practices. The GA is designed to find the best solution among many possible ones. Overall, we start with a population of strings of characters, each corresponding to a chromosome. In genetic algorithms, the genes stored in a given population are best suited to the environmental needs of the species. Indeed, certain gene variations give individuals who possess them a competitive advantage over the rest of the population. This competitive advantage results in better reproduction of these individuals, which can be passed on to the entire population in successive generations. The principle of genetic algorithms is to try different operations to see which ones produce the best results. Genetic algorithms can be used to solve problems effectively. However, the use of these methods must be conditioned by specific characteristics of the problem. The computational time of the fitness function for the evaluation should be reasonably short. It is rated many times.

A suitable optimization method should serve two opposite purposes: explore the solution domain (ergodicity property) and use the best solutions found. The Genetic Algorithm (GA), which is based on the theory of evolution, is commonly used in stochastic optimization applications. The algorithm is run on a population of N individuals (also known as "chromosomes"), each characterized by a gene (a component of the vector). Identifying the most functionally productive individual is a complex task that requires careful consideration of a variety of factors. Generations of individuals are followed by each other, with the most productive members continuing to produce until a limit is reached. The transition from one generation to the next is carried out through "operators,"

In this section, information, values and their interpretation should be given regularly, indicating to what extent the objectives stated in the last part of the introduction have been achieved. To emphasize enough and to ensure clarity in the expression, it would be appropriate to give the results in items. Suggestions can be made to new researchers for future studies on the subject. which are divided into several categories.

Recent studies have found that the GA method is a useful tool for estimating parameters in statistical models. Lee et al. studied the effects of a new diet on people's health. They found that the diet was effective in reducing the risk of various diseases. (2006) used a GA approach for two parameters in a mesoscale meteorological model.

A genetic algorithm is used along with Bayes' information criteria to identify the best periodic autoregressive models with one exogenous variable. Tekeli and Yüksel (2020) used a genetic algorithm to estimate the parameters of a twofold Weibull mixture model. In the same year, Akay et al.(2020) proposed optimal clustering of units using the GA approach.

6.1. Numerical Example

The psi31 data used in this section was taken from Birnbaum and Saunders's (1969) article. The dataset has been used in the past in the nucleic acid research study. This data illustrates the proposed method.

Table 6.1. Psi31 data.

Uncensored				Censored		
70	112	124	132	141	151	168
90	112	124	132	141	152	170
96	113	124	133	142	155	174
97	114	128	134	142	156	196
99	114	128	134	142	157	212
100	114	129	134	142	157	
103	116	129	134	142	157	
104	119	130	134	142	157	
104	120	130	136	144	158	
105	120	130	136	144	159	
107	120	131	137	145	162	
108	121	131	137	146	163	

In Table 6.1, there are 101 observations, with 80 of them describing the uncensored data and 21 of them describing the censored data. We use the Logrank test (a nonparametric test for skewed and censored data) to compare the hazard function of the two groups at each time point of the observed event.

The test is built by counting the number of observed and expected events in one of the groups at each time of the observed event, then calculating an overall summary at all the times when there is an event. The Logrank test is based on the same assumptions as the Kaplan-Meier survival curve, which states that the survival probabilities are the same for subjects recruited early

and late in the year. The study was conducted at the specified times. Deviations from these assumptions are significant if they are fulfilled differently in the comparison groups, for example, if censorship is more likely in one group than in another.

We also used the two-sample Kolmogorov-Smirnov test, which is one of the most useful and general nonparametric methods for comparing two samples, since it is sensitive to differences in the position and shape of the accumulated empirical data of the distribution functions of the two samples. In Table 2, the test statistics are given by

Table 6.2. Genetic Algorithm Phases

Structure of Chromosome	$[\alpha, \beta]$
Population	30
Iteration Count	30
Mutation Rate	0.1
Fitness Function	max(log-likelihood)

The model's performance was compared to that of the MLE method, and the results are shown in Table 3. Cross-validation helps to create several sets of data that can be used to estimate the accuracy of a model more accurately, with less bias and variation. With k-block cross-validation, we divide the original sample into k −samples and choose one of the samples as the validation set while the other $k - 1$ samples form the training set After training, the validation efficiency can be calculated. The validation process is repeated by selecting another validation sample from among the predefined blocks. At the end of the procedure, we obtained k performance scores for each block. The mean and standard deviation of the k performance scores can be used to estimate the bias and variance of validation performance. To ensure the validity of the estimate and to increase the reliability of the model performance comparison, a 5-fold cross-validation is used in the MLE and GA methods. In the cross-validation procedure, the data is divided into five parts, and each part is used to validate the accuracy of the model. Four of the data points were used to develop models, and the other data point was used to assess the performance of the models. Five different parts were used to perform the test data five times. The fitness averages of the models in each iteration with the training and test data are statistically analyzed using the Log-rank test.

After confirming the method's performance, the last model was created using all the data, omitting any models that appeared during iterations. Therefore, model parameters for the Birnbaum-Saunders distribution were estimated by the MLE and GA methods for the psi31 data.

We aimed to maximize the likelihood function as a fitness function in the genetic algorithm method. Log-rank and Kolmogorov-Smirnov were used as performance measures to assess the statistical significance of differences. Although the Kolmogorov-Smirnov value was generally the same in both methods, the logarithmic rank value turned out to be effective in the GA method for

some nuclei and the MLE method for some nuclei. Table 3 lists the ranges of GA's success. However, the models obtained by the GA method are more suited for the psi31 data set.

Table 6.3. Results of MLE and GA methods using different kernels for psi31 data

Kernel	Method	Model		Log-rank	Kolmogorov-Smirnov
		α	β		
T_2	MLE	0.1303	131.3480	**0.0493**	0.2183
	GA	0.1312	132.4502	0.0664	0.2183
T_4	MLE	0.1570	133.2545	0.0904	0.2083
	GA	0.1632	133.0659	**0.0888**	**0.1988**
T_8	MLE	0.1789	133.7577	0.1334	0.1988
	GA	0.1753	134.1157	**0.1312**	0.1988
T_{15}	MLE	0.1850	133.8760	0.1654	0.1988
	GA	0.1835	134.9488	**0.1620**	0.1988
T_{50}	MLE	0.1982	134.3694	**0.2087**	0.1988
	GA	0.1924	135.4835	0.2118	0.1988
T_{100}	MLE	0.2006	134.4090	0.2104	0.1888
	GA	0.1936	135.5795	**0.1898**	0.1888

6.2. Simulation Study

In this study, we simulate the performance of GA and the current method, ML, to evaluate their relative merits. Several simulations were run for different sample sizes. In this simulation study, the sample size (n) takes values of 50, 100, 250, and 500, the censoring rate (CR) is 10%, 20%, 30%, and 40%, and the value of ν is 5 and 20. The model parameters are α = 0.3, β = 60, and 120. Thus, four different α and β values are investigated. For each case, 100 different synthetic data sets were generated, and bias and MSE values for the parameters were calculated. The estimated mean square error (EMSE) and estimated bias (EBIAS) for each case are calculated according to the Equations $(91, 92, 93, 94)$.

$$EMSE_{\alpha} = \frac{1}{m}\sum_{i=1}^{m}(\hat{\alpha}_i - \alpha_i)^2 \tag{91}$$

$$EMSE_{\beta} = \frac{1}{m}\sum_{i=1}^{m}\left(\hat{\beta}_i - \beta_i\right)^2 \tag{92}$$

$$EBIAS_{\alpha} = \frac{1}{m}\sum_{i=1}^{m}\hat{\alpha}_i - \alpha \tag{93}$$

$$EBIAS_{\beta} = \frac{1}{m}\sum_{i=1}^{m}\hat{\beta}_i - \beta \tag{94}$$

with $\theta = (\alpha, \beta)$, m is the number of tests in the simulation. The results are shown in Figures $(5, 6, 7, 8)$. According to simulation values, the value of ν does not affect the results. As the sample size (n) increased, although the GA method estimated the β parameter better than the ML method, it was seen that there was no difference between the two methods in the estimation of the α parameter. In other words, the larger the sample size, the better the GA method can predict.

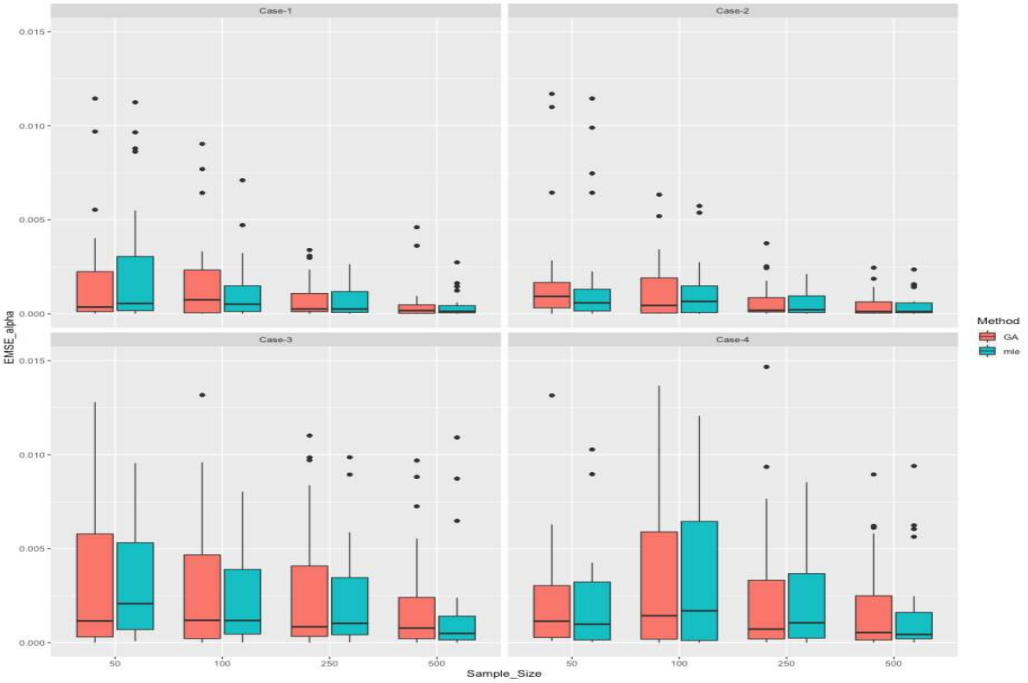

Figure 6.1. Box plots of $EMSE_{\alpha}$ according to sample size for *CR=10%*

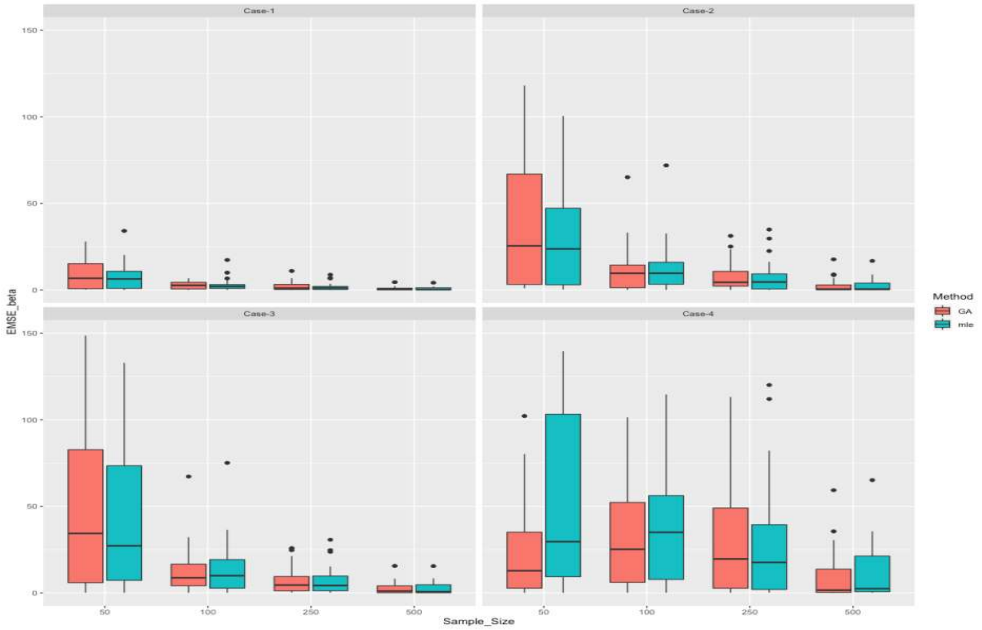

Figure 6.2. Box plots of $EMSE_{\beta}$ according to sample size for *CR=10%*

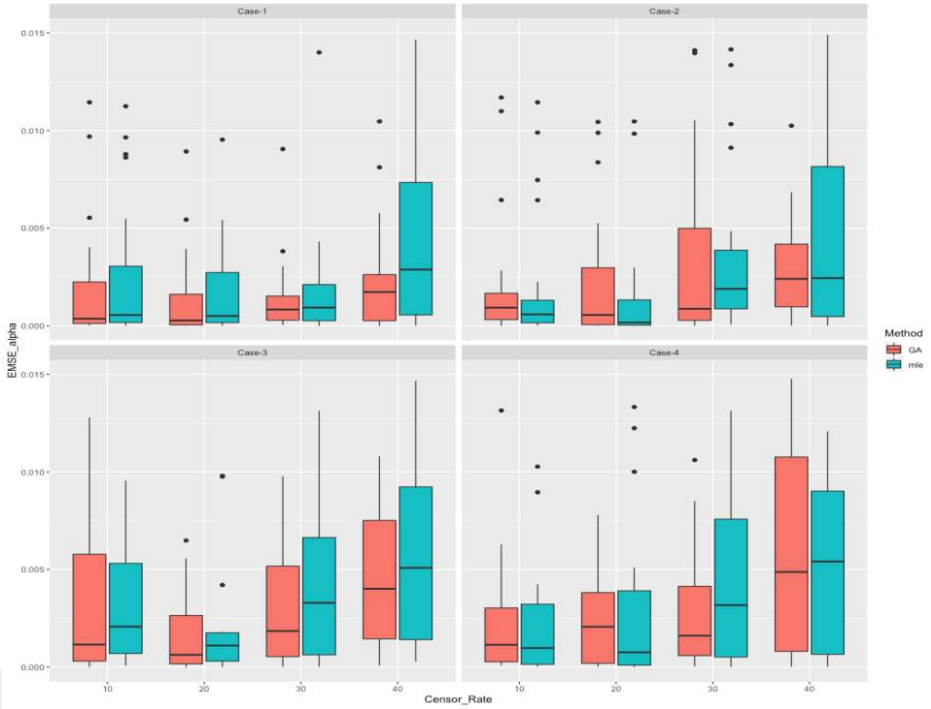

Figure 6.3. Box plots of $EMSE_{\alpha}$ according to censor rates.

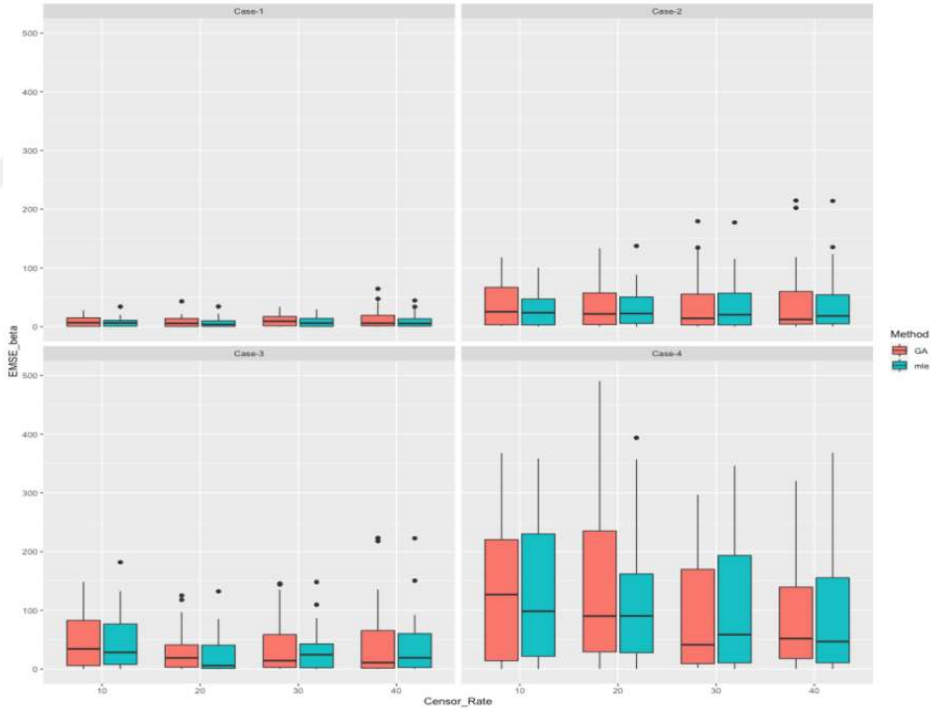

Figure 6.4. Box plots of $EMSE_{\beta}$ according to censor rates.

According to simulation results, for small samples, we can say that the $EMSE_{\alpha}$ value calculated by the GA method gives more minor results than $EMSE_{\alpha}$ values obtained by the ML method when $n = 50$, censor rates for 0.1 and 0.2, and $\beta = 60$ except fort he fist line of the simulation. Again, when $CR = 0.3$ and 0.4 are taken, we can say that the $EMSE_{\alpha}$ values obtained by the GA method are smaller at each α value of $\beta = 120$. As the sample size increases, we can

say that the GA method for $EMSE_\alpha$ does not differ significantly except in some cases. It can be said that it is recommended to use the MLE method as n increases. If we look at the $EMSE_\beta$ analysis for $n = 50$, it can be said that the MLE estimation gives better results than GA for each censorship rate and each value of α and β.

We conclude that in all four cases, α and β for $n = 100$ and $CR = 0.1$, the $EMSE_\beta$ value calculated with GA is smaller than the $EMSE_\beta$ calculated with MLE. However, as the censor ratio increases, we can see that the MLE method has a smaller $EMSE_\beta$ value, except for some cases. When $n = 250$ is taken, the $EMSE_\beta$ value of the GA method calculated with $\alpha = 0.3$ and $\beta = 120$ and $\alpha = 0.6$ and $\beta = 60$ values for $CR = 0.2$ is smaller than the $EMSE_\beta$ values of the MLE method. In other cases, we can say that the MLE method gives better results.

When the sample size increases, it can be seen that for $n = 500$, for the values of $CR = 0.1$ and 0.3, the $EMSE_\beta$ values obtained with MLE are more significant than the $EMSE_\beta$ values calculated with GA in most of the α and β values. However, this situation is reversed as the censorship rate increases to 0.2 and 0.4. Excepte for $CR = 0.1, \alpha = 0.6, \beta = 60$ and $CR = 0.4, \alpha = 0.6$ and $\beta = 120$, we not that for the small samples $(n = 50)$, the $BIAS_\alpha$ value calculated with the GA method is greater than the $BIAS_\alpha$ calculated by the MLE method.

For $n = 100$, the value of the $BIAS_\alpha$ value calculated by the MLE method is smaller than the $BIAS_\alpha$ value calculated by the GA method, except for $CR = 0.1, and\ \alpha = = 0.6, \beta = 60{,}120$. The $BIAS_\alpha$ calculated by the GA method is smaller than that calculated by the MLE method, for $n = 250, CR = 0.2, \alpha = 0.6, and\ \beta = 60{,}120$. In other cases, the $BIAS_\alpha$ value obtaıned by the GA method is greter than that obtaıned by the MLE method. In all cases for large samples $(n = 500)$, the $BIAS_\alpha$ calculated wıth the MLE method gives smaller resuls than the $BIAS_\alpha$ calculated by GA method.

7. CONCLUDING REMARKS

In this thesis, we considered the two-parameter Birnbaum-Saunders distribution. The Birnbaum-Saunders model has received considerable attention for various reasons. The two-parameter Birnbaum-Saunders distribution has a shape and a scale parameter. Due to the presence of the shape parameter, the probability density function (pdf) of a Birnbaum-Saunders distribution can take different forms. The MLE method used to estimate the model parameter based on type II censored data was compared to that of GA that we proposed. We performed a simulation study, based on type II right-censored data to estimate the model parameters, and it turns out that the ML method is worse in front of a large sample size.

The psi31 data set and simulation study were used to estimate the parameters of the two-parameter Birnbaum-Saunders model. For parameter estimation, we compared the genetic algorithm method proposed by us and the maximum likelihood method used in the literature. The results show that GA provides a better estimate when the sample size is large. In addition, the log-rank fit test result was found to be the smallest value of the test phase when the entire data set was applied with the GA method. For this reason, it can be said that the model that was established with the GA method performs the best in predicting outcomes. Besides, a Monte Carlo simulation study was carried out in order to examine the results better. The model's compatibility with the simulation study has been tested for different sample sizes, and it has been found that the models found with the proposed GA method fit the data better when the sample size increases.

In future studies, large sample characteristics of estimators using genetic algorithms can be examined for censored data. It can also be suggested to use genetic algorithms to estimate the parameters of the Generalized Birnbaum Saunders distribution.

REFERENCES

Aalen, O.O. (1978) Nonparametric Estimation of Partial Transition Probabilities in Multiple Decrement Models. The Annals of Statistics, 6,534-545.

Ahmed, S.E.; Castro-Kuriss, C.; Flores, E.; Leiva, V. and Sanhueza, A. (2010). A truncated version of the Birnbaum–Saunders distribution with an application in financial risk, Pak. J. Statistist., 26, 293-311.

Akay, Ö., Tekeli, E. & Yüksel, G. (2020) Genetic Algorithm with New Fitness Function for Clustering. Iran. J. Sci. Technol. Trans Sci. , 44, 865-874.

Alemdar, N. and Özyildirim, S. (1998), "A genetic game of trade, growth and externalities", Journal of Economic Dynamics and Control 22(6), 811-832.

Axeldrod, R. (1987). The evolution of strategies in the iterated Prisoner's Dilemma. In L. D. Davis, ed., Genetic Algorithms and Simulated Anmealing. Morgan Kaufinann

Azevedo, C., Leiva, V., Athayde,E.,Balakrishnan, N.,2012. Shape and change point analyses of the Birnbaum- Saunders-t hazard rate and associated estimation. Comput. Stat. Data Anal. 56, 3887-3897.

Bhatti, C.R. (2010). The Birnbaum–Saunders autoregressive conditional duration model, Math. Comp. Simul., 80, 2062–2078.

Balakrishnan, N., Cohen, A. C. (1991) Order Statistics and Inference: Estimation Methods, Academic Press, San Diego.

Balakrishnan, N., Leiva, V., and Lòpez, J. (2007). Acceptance sampling plans from truncated life tests based on the generalized Birnbaum-Saunders distribution. Communications in Statistics: Simulayion and Computation, 36: 643–656.

Balakrishnan, N., Leiva, V., Sanhueza, A., Vilca, F., 2009b. Estimation in the Birnbaum–Saunders distribution based on scale-mixture of normals and the EM-algorithm. Stat. Oper. Res. Trans. 33, 171-192.

Balakrishnan, N., Gupta, R., Kundu, D., Leiva, V., Sanhueza, A., 2011. On some mixture modesls based on the Birnbaum–Saunders distribution and associated inference. J. Stat. Plan. Inference 141, 2175–2190.

Barros, M., Leiva, V., Ospina, R., Tsuyuguchi, A. (2014), "Goodness-of-fit tests for the Birnbaum-Saunders distribution with censored reliability data", IEEE Transactions on Reliability,vol.63, 543-554.

Barros, M., Paula, G., Leiva, V., 2008. A new class of survival regression models with heavy-tailed errors: robustness and diagnostics. Lifetime Data Anal. 14, 316–332.

Barros, M.,Paula, G.A.,Leiva,V.(2009), "An R implementation for generalized Birnbaum-Saunders distributions", Computational Statistics and Data Analysis, vol. 53, 1511 – 1528.

Birnbaum-Saunders (1958). A statistical model for life-length of materials. Journal of the American Statistical Association, vol, 53. pp. 151-160.

Birnbaum, Z.W. Esary. J.D. Marshall . A. W.(1966). A stochastic characterization of wear-out for components and systems . Annals of Mathematical Statistics . 37 . 4. 816–825.

Birnbaum, Z.W.Saunders, S.C. (1969) Estimation for a family of life distributions with applications to fatigue. *J. Appl. Probab.*, 6(2), 328-347. https://doi.org/10.2307/3212004.

Birnbaum,Z.W,Saunders, SC(1969a):A new family of life distributions. J.Appl. Probab. 6,319-327.

Birnbaum, Z.W. and Saunders, S.C.(1969b). Estimation for a family of life distributions with applications to fatigue, J. Appl. Prob., 6, 328–347.

Bland JM, Altman DG *(*1998). Survival probabilities (the Kaplan-Meier method) *BMJ.*;317:1572.

Breslow, N. E. (1972), "Discussion of Professor Cox's Paper," Journal of the Royal Statistical soc, Series B, 34, 216–217.

Breslow, N. (1974). Covariance analysis of censored survival data, Biometrics, 30: 89-99.

Castillo, N.O.,Gomez, H.W.,Bolfarine, H.(2011), "Epsilon Birnbaum-Saunders distribution family: properties and inference", Statistical Papers, vol. 52, 871 – 883.

Celeux, G,Diebolt,J.,1992. A stochastic approximation type EM algorithm for the mixture problem. Stochastics and Stochastics Reports 41, 127–146.

Chan, K. S. and Ledolter, J. (1995). Monte Carlo EM estimation for time series involving counts. Journal of the American Statistical Association 90 242–252.

Chang, D., Tang, L., (1993). Reliability bounds and critical time for the Birnbaum–Saunders distribution. IEEE Trans. Device Mater. Reliab. 42, 464–469.

Chang, D.S. and Tang, L.C. (1994). Graphical analysis for Birnbaum–Saunders distribution, Microelect. Rel., 34, 17–22.

Cramér, H., (1946). Mathemati cal Methods of Statistics. Princeton University Press, Princeton.

Cordeiro, M., Lemonte, J., 2011. The β-Birnbaum–Saunders distribution: an improved distribution for fatigue life modelling. Comput. Stat. Data Anal.55, 1445–1461.

Cox, DR, Oates, D. (1984) : Analysis of Survival Data. Chapman and Hall, New York.

David Cox, D. R. (1972). Regression models and life tables (with discussion). Journal of the Royal Society, Series B, 34:187–220.

De Jong, K. (1980). Adaptive System Design: A Genetic Approach. IEEE Trans. On Systems, Man, and Cybernetics, **10**, 556-574.

Desmond, A.F. (1985). Stochastic models of failure in random environments. *Canadian Journal of Statistics, **13**, 171-183.*

Díaz-García, J., Leiva, V., 2005. A new family of life distributions based on elliptically contoured distributions. J. Stat. Plan. Inference 128, 445–457.

Efron, B. (1977). The efficiency of Cox's likelihood function for censored data, The Journal of the American Statistical Association, 72: 557-565.

Esary, J., Marshall, A.W., Proschan, F., 1973. Shock models and wear processes. Ann. Probab. 1, 627-649.

Ferreira, M., Gomes, M. I., Leiva, V. (2012), "On an extreme value version of the Birnbaum-Saunder distribution", REVSTAT-Statistical Journal, vol. 10, 181 – 210.

Freudenthal, A., Shinozuka, M., 1961. Structural Safety Under Conditions of Ultimate Load failure and Fatigue. Technical Report WADD-TR-61-77, Wright Air Development Division.

Galea, M., Leiva, V., Paula, G., 2004. Influence diagnostics in log-Birnbaum–Saunders regression mdels. J. Appl. Stat. 31, 1049–1064.

Gill, R. D.(1980). Censoring and stochastic integrals. Mathematical Centre Tracts 124, Amsterdam.

Goldberg, D. (1989). Genetic Algorithms in Search, Optimization, and Machine Learning Addison-Wesley Publishing.

Gómez, H., Olivares-Pacheco, J., Bolfarine, H., 2009. An extension of the generalized Birnbaum–Saunders distribution. Stat. Probab. Lett. 79, 331–338.

Guiraud, P., Leiva, V., Fierro, R., 2009. A non-central version of the Birnbaum–Saunders sistribution for reliability analysis. IEEE Trans. Reliab. 58,152–160.

Holland, J.H. (1975). Adaptation in natural and artificial systems. University of Michigan press, Michigan.

Holland, J. H. (1960). On Iterative Circuit Computers Constructed of Microelectronic Components and Systems. In Proceedings of the Western Joint Computer Conference (WJCC) – Session on the Design, Programming, and Sociological Implications of Microelectronics, National Joint Computer Commite, IEEE.

J. Eastman and L. J. Bain (1973), "A Property Of Maximum Likelihood Estimators In The presence of Location-Scale Nuisance Parameters," *Communications in Statistics*, vol. 2, no. 1, 23 – 28.

Johnson, N.L., Kotz, S., Balakrishnan. N. (1995) *Continuous Univariate Distributions*. Vol. 2, 2nd Edition, New York: John Wiley.

Kao, J. (1959) A Graphical Estimation of Mixed Weibull Parameters in Life-Testing of Electron tubes. Technometrics, 1, 389-407. http://dx.doi.org/10.1080/00401706.1959.10489870.

Kotz, S., Leiva, V., Sanhueza, A., 2010. Two new mixture models related to the inverse Gaussian distribution Methodol. Comput. Appl. Probab. 12, 199–212.

Krishnakumar K, Goldberg DE (1992). Control-System Optimization Using Genetic Algorithms. J Guid Control Dyn ;**15**:735-740.

Kundu, D., Kannan, N., Balakrishnan, N. (2008) On the hazard function of Birnbaum–Saunders distribution and associated inference. *Comput. Statist. Data Anal.*, 52(5), 2692-2702.

Leiva, V., Barros, M., Paula, G., Galea, M., 2007. Influence diagnostics in log-Birnbaum–Saunders regression models with censored data. Comput. Stat. Data Anal. 51, 5694–5707.

Leiva, V., Barros, M., Paula, G., Sanhueza, A., 2008a. Generalized Birnbaum–Saunders distribution applied to air pollutant concentration. Environmetrics 19, 235–249.

Leiva, Víctor, Edgardo Rojas, Manuel Galea, and Antonio Sanhueza. 2014. Diagnostics in Birnbaum–Saunders accelerated life models with an application to fatigue data. Applied Stochastic Models in Business and Industry 30: 115–31.

Leiva, V., Santos-Neto, M.; Cysneiros, F.J.A. and Barros, M. (2014c). Birnbaum–Saunders statistical modelling: a new approach, Stat. Mod., 14, 21–48.

Leiva, V., Riquelme, M., Balakrishnan, N., Sanhueza, A., 2008c. Lifetime analysis based on the generalized Birnbaum–Saunders distribution. Comput. Stat. Data Anal. 52, 2079–2097.

Lee, Y.H., Park, S.K. and Chang D.-E. (2006) Parameter estimation using the genetic algorithm and its impact on quantitative precipitation forecast. *Ann. Geophys.*, 24, 3185–3189.

Lemonte, A., 2013. A new extension of the Birnbaum Saunders distribution. Braz. J. Probab. Stat. 27, 133–149.

Lieblein, J., Zelen, M., 1956. Statistical investigation of the fatigue life of deep ball bearing. J. Res. Nat. Bur. Stand. 57, 273–316.

Marchant, C.; Bertin, K.; Leiva, V. and Saulo, H. (2013a). Generalized Birnbaum–Saunders kernel density estimators and an analysis of financial data, Comp. Stat. Data Anal., 63, 1–15.

Marchant, C.; Leiva, V.; Cavieres, M.F. and Sanhueza, A. (2013b). Air contaminant statistical distributions with application to PM10 in Santiago, Chile, Rev. Environ. Contam. Tox., 223, 1-31.

Marco Tomassini, Olivier V. Pictet, Michel M. Dacorogna, Bastien Chopard, Mouloud Oussaidene and Roberto Schirru (1995). Using genetic algorithms for robust optimization in financial applications.Neural Network World.

Marshall, A., Olkin, I., (2007). Life Distributions. Springer, New York.

Martínez-Flórez G., Bolfarine H. and Gómez H.W. (2014), "An alpha-power extension for the Birnbaum- Saunders distribution", Statistics 48, No. 4, 896–912.

Mátyás, L., 1999. Generalized Method of Moments Estimation. Cambridge University Press, New York.

Miner, MA (1945) : Cumulative Damage in Fatigue. J. Appl. Mech. Trans. ASME. 67, 159 – 164.

Naderi, M., Arabpour, A., Lin, T.I., Jamalizadeh, A. (2017) Nonlinear regression models based on the normal mean-variance mixture of Birnbaum- Saunders distribution, *J. Korean Stat. Soc.,* 46 (3),476–485.https://doi.org/10.1016/j.jkss.2017.02.002.

Ng, H.K.T., Kundu, D. , Balakrishnan, N. (2003) Modified moment estimation for the two parameter Birnbaum-Saunders parameter Birnbaum-Saunders distribution. *Comput. Statist. Data Anal.*, 43(3), 283-298. https://doi.org/10.1016/S0167-9473(02)00254-2.

Ng, H. K. T., Kundu, D., Balakrishnan, N. (2006) Point and interval estimation for the two-parameter Birnbaum–Saunders distribution based on Type-II censored samples. *Comput. Statist. Data Anal*., 50(11), 3222-3242. https://doi.org/10.1016/j.csda.2005.06.002.

Owen, W.J. (2006). A new three-parameter extension to the Birnbaum–Saunders distribution. IEEE Transactions on Reliability 55, 475–479.

Özyildirim, S. (1996), "Three country trade relations : A discrete dynamic game approach", *Computer and Mathematics with Applications* **32**, 43-56.

Özyildirim, S. (1997), "Computing open-loop noncooperative solution in discrete dynamic games", *Evolutionary Economics* **7** (1), 23-40.

Pearson K (1894) Contribution to the mathematical theory of evolution. Philos Tr R Soc S-A 185:71-110.

Podlaski, R. (2008). Characterization of diameter distribution data in nearnatural forests using the Birnbaum-Saunders distribution, Can. J. Forest Res., 18, 518–526.

Quenouille, Maurice H. (1949). *"Problems in Plane Sampling"*. The Annals of Mathematical Statistics **20** (3): 355–375.

Quenouille, Maurice H.(1956). "Notes on Bias in Estimation".Biometrika.**43** (3–4): 353-360. *doi*:*10.1093/biomet/43.3-4.353*.

Rieck, JR, Nedleman, J (1991): A Log-Linear Model for the Birnbaum-Saunders Distribution. Technometrics. 33, 51-60.

Rieck, J R (1999): A Moment-Generating Function with Application to the Birnbaum-Saunders distribution. Communications in Statistics - Theory and Methods. 28, 2213–2222.

Robert Pereira (2000), Genetic Algorithm Optimization for Fiance and Investments, *MPRA* Paper, **8610**, University Library of Munich, Germany.

Santos-Neto, M., Cysneiros, F. J. A., Leiva, V., and Ahmed, S. (2012). On new parameterizations of the Birnbaum-Saunders distribution. Pakistan Journal of Statistics, 28:1–26.

Santos-Neto, M., Cysneiros, F. J. A., Leiva, V., Barros, M. (2014), "On a reparameterized Birnbaum- Saunders distribution and its moments, estimation and applications", REVSTAT- Statistical Journal, vol. 12, 247 – 272.

Saulo, H, Leão, J, Bourguignon, M. (2012). The Kumaraswamy Birnbaum–Saunders distribution. Journal of Statistical Theory and Practice 6, 745-759.

Saunders, S.C. (2007). Reliability, Life Testing and Prediction of Services Lives, Springer, New York.

Shih W (2002) . Problems in dealing with missing data and informative censoring in clinical trials. *Curr Control Trials Cardiovasc Med*; 3:4.

Tekeli, E., Yüksel, G. (2020) Estimating the parameters of twofold Weibull mixture model in right-censored reliability data by using genetic algorithm. *Commun. Stat. B: Simul Comput.* https://doi.org/10.1080/03610918.2020.1808681.

Therneau, Terry M.,et Patricia M.Grambsch.2000.Modelling survival data:extending the Cox model. *Springer.*

Tsionas, Efthymios G.2001.Inférence bayésienne dans la régression de Birnbaum-Saunders.*Communication in Statistics: Theory and Methods* 30: 179-193.

Tukey, Tukey, John W. (1958). "Bias and confidence in not quite large samples (abstract)". *The Annals of Mathematical Statistics*. **29** (2): 614. doi:10.1214/aoms/1177706647.

Vilca, F.;Sanhueza, A.;Leiva, V.and Christakos, G. (2010). An extended Birnbaum–Saundersmodel and its application in the study of environmental quality in Santiago, Chile,Stoch. Environ. Res. Risk Assess., 24, 771–782.

Vilca, F.; Santana, L.; Leiva, V. and Balakrishnan, N. (2011). Estimation of extreme percentiles in Birnbaum-Saunders distributions, Comp. Statist. Data Anal., 55, 1665–1678.

Villegas, Cristian, Gilberto A.Paula, and Víctor Leiva. 2011.Birnbaum–Saunders mixed models for censored reliability data analysis. IEEE Transactions on Reliability 60: 748–58.

Volodin, I.and Dzhungurova, O. (2000). On limit distribution emerging in the generalized Birnbaum-Saunders model, Journal of Mathematical Science, 99, 1348–1366.

Wu, C.F.Jeff (1983). On the Convergence Properties of the EM Algorithm, The Annals of Statistics, 11, 95-103.

Xie, F.C.and Wei, B.C.(2007).Diagnostics analysis for log-Birnbaum-Saunders regression models", computational Statistics & Data Analysis, vol. 51, 4692-4706.

Zelen, M . and Dannemiller, M.C . (1961). The robustness of life testing procedures derived from the exponential distribution. Technometrics, (3), 29-40.

CURRICULUM VITAE

Education

Degree	**Institution**	**Graduation Year**
Master Es- Mathematics Sciences	Gamal A.Nasser university of Conakry	2013 - 2015
Bachelor	Gamal A.Nasser university of Conakry	2009 - 2012
General certificate of education	Muigni Baraka School in the Comoros	2006 - 2009

Experience

Year	**Work Place**	**Title**
2019 – 2023	Çukurova University	Phd Studient
2017 - 2018	Çukurova University Turkish Language Teaching application and Research Center	Turkish Language
2016 - 2017	University of Comoros	Mathematics Teacher
2012 – 2015	Guinea Conakry	National exam corrector
2012 – 2015	Guinea Conakry	Mathematics Teacher

Publications

“Estimation of Two Parameter Birnbaum-Saunders Distribution Based on Type-II Right Censored Reliability Data Using Genetic Algorithm” accepted to Publication in “ Communications in Mathematics an Applications”.

APPENDICES

Appendix I. EBIAS and EMSE for $n = 50$ in simulation. **($v = 20$)**

CR	α	β	MLE				GA			
			$BIAS_\alpha$	$BIAS_\beta$	$EMSE_\alpha$	$EMSE_\beta$	$BIAS_\alpha$	$BIAS_\beta$	$EMSE_\alpha$	$EMSE_\beta$
0.1	0,3	60	0,0205	0,5649	0,0030	7,8339	0,0230	0,9758	0,0034	8,5844
	0,3	120	0,0115	0,5760	0,0019	29,7584	0,0153	1,1121	0,0027	34,3756
	0,6	60	0,0364	3,1430	0,0099	44,3298	0,0225	3,2260	0,0070	48,9484
	0,6	120	0,0295	2,3289	0,0096	137,2190	0,0431	2,0523	0,0111	143,9680
0.2	0,3	60	0,0081	0,4267	0,0026	7,2874	0,0114	0,8424	0,0020	8,5350
	0,3	120	0,0038	0,8494	0,0015	31,6288	0,0152	0,6878	0,0030	32,7400
	0,6	60	0,0036	0,6684	0,0079	27,1224	0,0142	1,3854	0,0068	34,1012
	0,6	120	0,0016	2,0533	0,0070	135,2965	0,0290	1,3314	0,0114	144,1401
0.3	0,3	60	0,0117	0,4495	0,0025	8,5015	0,0334	1,6938	0,0048	10,6719
	0,3	120	0,0229	1,3087	0,0039	36,6987	0,0272	2,2748	0,0034	38,2205
	0,6	60	0,0133	0,5037	0,0104	31,9327	0,0529	2,4761	0,0160	36,8076
	0,6	120	0,0491	2,2820	0,0144	139,8633	0,0527	4,6486	0,0126	155,6379
0.4	0,3	60	0,0077	0,0529	0,0050	10,0570	0,0364	1,8824	0,0056	13,4630
	0,3	120	0,0086	0,8835	0,0055	40,0596	0,0103	1,8883	0,0034	41,9987
	0,6	60	0,0332	0,7100	0,0205	37,8934	0,0527	2,3496	0,0224	40,8812
	0,6	120	0,0346	2,2388	0,0190	193,2440	0,0336	3,9176	0,0142	190,3549

Appendix II. EBIAS and EMSE for $n = 100$ in simulation. **($v = 20$)**

CR			MLE				GA			
	α	β	$BIAS_\alpha$	$BIAS_\beta$	$EMSE_\alpha$	$EMSE_\beta$	$BIAS_\alpha$	$BIAS_\beta$	$EMSE_\alpha$	$EMSE_\beta$
0.1	0,3	60	0,0174	-0,1593	0,0012	3,0987	0,0230	0,4329	0,0018	2,7314
	0,3	120	0,0171	-0,1446	0,0011	13,0703	0,0194	0,3689	0,0012	12,2567
	0,6	60	0,0415	0,1043	0,0058	13,6092	0,0406	0,6550	0,0050	12,7474
	0,6	120	0,0383	-0,4488	0,0052	48,5017	0,0365	0,1640	0,0051	42,3841
0.2	0,3	60	0,0160	0,0168	0,0014	3,1662	0,0200	0,4719	0,0018	3,5926
	0,3	120	0,0163	-0,1502	0,0014	14,1960	0,0199	0,2412	0,0016	12,4760
	0,6	60	0,0312	0,1326	0,0050	12,5726	0,0332	0,4805	0,0050	12,1026
	0,6	120	0,0275	0,0347	0,0049	49,1496	0,0399	0,8854	0,0062	51,4403
0.3	0,3	60	0,0187	0,1640	0,0016	3,8223	0,0348	0,9936	0,0026	5,5119
	0,3	120	0,0170	0,2077	0,0016	13,2496	0,0211	0,5122	0,0021	12,8348
	0,6	60	0,0396	0,4210	0,0075	15,1177	0,0534	1,2340	0,0087	16,3896
	0,6	120	0,0265	-0,2713	0,0069	50,5689	0,0372	1,6085	0,0086	50,2440
0.4	0,3	60	0,0066	-0,2682	0,0017	3,8734	0,0278	1,1994	0,0026	6,8881
	0,3	120	0,0086	-0,1901	0,0024	11,9040	0,0171	0,7374	0,0026	13,2442
	0,6	60	0,0177	-0,3373	0,0083	14,3330	0,0419	0,9243	0,0108	14,7966
	0,6	120	0,0160	-0,8434	0,0096	50,7946	0,0328	1,2795	0,0094	51,9959

Appendix III. EBIAS and EMSE for $n = 250$ in simulation. **($\nu = 20$)**

CR	α	β	MLE				GA			
			$BIAS_\alpha$	$BIAS_\beta$	$EMSE_\alpha$	$EMSE_\beta$	$BIAS_\alpha$	$BIAS_\beta$	$EMSE_\alpha$	$EMSE_\beta$
0.1	0,3	60	0,0173	-0,1012	0,0006	1,8500	0,0194	0,2962	0,0008	2,1694
	0,3	120	0,0177	-0,2621	0,0006	7,6492	0,0197	-0,4029	0,0007	7,9614
	0,6	60	0,0344	-0,1140	0,0022	7,0492	0,0356	-0,0156	0,0026	7,1125
	0,6	120	0,0364	-0,4429	0,0025	28,9340	0,0375	-0,1432	0,0026	30,3716
0.2	0,3	60	0,0190	-0,1674	0,0006	1,8631	0,0219	0,5951	0,0008	2,7524
	0,3	120	0,0190	-0,3355	0,0007	7,6374	0,0194	-0,1744	0,0007	7,5077
	0,6	60	0,0390	-0,2406	0,0027	7,2035	0,0382	0,0082	0,0029	6,9840
	0,6	120	0,0392	-0,5107	0,0028	27,0372	0,0380	-0,2564	0,0026	31,0835
0.3	0,3	60	0,0159	-0,1310	0,0005	1,7773	0,0320	0,8263	0,0017	4,5721
	0,3	120	0,0170	-0,1513	0,0006	7,0226	0,0209	0,2097	0,0007	8,7711
	0,6	60	0,0381	-0,1999	0,0028	7,1546	0,0462	0,4153	0,0038	8,1626
	0,6	120	0,0301	-0,7312	0,0022	28,1480	0,0405	0,4490	0,0029	29,4637
0.4	0,3	60	0,0156	-0,2170	0,0008	1,7828	0,0246	0,7618	0,0012	3,5249
	0,3	120	0,0143	-0,3787	0,0007	8,1642	0,0238	0,5055	0,0011	8,8527
	0,6	60	0,0328	-0,4308	0,0031	7,2640	0,0387	0,3307	0,0038	6,6391
	0,6	120	0,0323	-0,7351	0,0029	33,6439	0,0437	0,4144	0,0038	31,5758

Appendix IV. EBIAS and EMSE for $n = 500$ in simulation. **($v = 20$)**

CR	α	β	MLE				GA			
			$BIAS_\alpha$	$BIAS_\beta$	$EMSE_\alpha$	$EMSE_\beta$	$BIAS_\alpha$	$BIAS_\beta$	$EMSE_\alpha$	$EMSE_\beta$
0.1	0,3	60	0,0147	-0,2046	0,0004	0,6617	0,0163	0,2274	0,0006	0,7095
	0,3	120	0,0133	-0,4043	0,0004	2,5768	0,0153	-0,4148	0,0005	2,5546
	0,6	60	0,0279	-0,4013	0,0017	2,6272	0,0328	-0,2688	0,0020	2,4320
	0,6	120	0,0257	-0,9642	0,0016	11,2665	0,0289	-0,8576	0,0018	10,1097
0.2	0,3	60	0,0105	-0,2250	0,0003	0,6526	0,0151	0,4214	0,0005	1,7869
	0,3	120	0,0103	-0,5558	0,0003	2,9039	0,0126	-0,2633	0,0004	3,0530
	0,6	60	0,0232	-0,5494	0,0015	2,8302	0,0278	-0,2527	0,0018	2,8893
	0,6	120	0,0209	-1,1914	0,0013	11,2624	0,0237	-0,8918	0,0015	11,0077
0.3	0,3	60	0,0144	-0,1133	0,0006	0,7231	0,0184	0,6179	0,0009	2,6941
	0,3	120	0,0143	-0,3146	0,0006	2,5978	0,0163	-0,0847	0,0006	2,5345
	0,6	60	0,0279	-0,2426	0,0023	2,5354	0,0366	-0,0018	0,0025	2,3388
	0,6	120	0,0277	-0,7001	0,0022	10,5295	0,0333	-0,2857	0,0027	10,0050
0.4	0,3	60	0,0144	-0,1260	0,0005	1,5825	0,0265	0,7647	0,0011	3,6343
	0,3	120	0,0118	-0,2868	0,0004	2,6385	0,0195	0,2615	0,0006	3,3184
	0,6	60	0,0244	-0,4495	0,0022	2,5008	0,0320	-0,0646	0,0028	2,5958
	0,6	120	0,0184	-1,1371	0,0029	11,3287	0,0329	-0,5660	0,0031	9,9236