GOMPERTZ DISTRIBUTION

BY

Osarobo AGBONBAMIEN
PSC2008391

DEPARTMENT OF STATISTICS

FACULTY OF PHYSICAL SCIENCE

UNIVERSITY OF BENIN

BENIN CITY

FEBRUARY 2025.

GOMPERTZ DISTRIBUTION

BY

Osarobo AGBONBAMIEN
PSC2008391

A PROJECT WORK SUBMITTED TO THE DEPARTMENT OF STATISTICS, FACULTY OF PHYSICAL SCIENCES, UNIVERSITY OF BENIN, BENIN CITY, IN PARTIAL FULFILLMENT OF THE AWARD OF BACHELOR OF SCIENCE (B.Sc. HONOURS) DEGREE IN STATISTICS

FEBRUARY 2025.

UNDERTAKING

This project was carried out by **Osarobo AGBONBAMIEN** with Mat no **PSC2008391,** I have not copied the work of any other author(s) All texts used have been duly cited and acknowledged

________________________________ ________________________

Osarobo AGBONBAMIEN **Date**

CERTIFICATION

This is to certify that the project work was carried out by **Osarobo AGBONBAMIEN** with Mat no **PSC2008391** in the Department of statistics, Faculty of Physical sciences University of Benin in partial fulfillment for the requirement for the award of the Bachelor of sciences (B.Sc.) Degree in Department of Statistics.

____________________ __________________________

Dr. S.A Osagie **Date**

Project supervisor

____________________________ ________________________

Prof. A. Iduseri. **Date**

Head of department

____________________ ____________________

EXTERNAL EXAMINER **DATE**

DEDICATION

I dedicate this work to God, for giving me the strength and guidance to properly carry out and complete the work and also for his protection throughout my time at the University of Benin This work is also dedicated to my parents, who made this journey as possibly easy as they could, by encouraging and guiding me.

ACKNOWLEDGEMENT

My gratitude goes to God Almighty for his love, strength, wisdom and mercy for daily loading me with benefits.

My heartfelt gratitude goes to my supervisor Dr. S.A Osagie for his unwavering supports and push at all times.

I am indeed forever grateful to my parents Mr Christopher Oviawe and Mrs. Josephine OviJawe for their love and continuous support and prayers I love you both and to my guardian Mr Agbonbamien Osamuyi for their constant support and prayers I am truly grateful.

I wish to express my warm appreciation to my sister Mrs Mercy Agbonbamien who has shown me love and has supported me up to this very moment, I say a big thank you.

I also acknowledge my friends and course mates for their great support towards the Completion of this project

TABLE OF CONTENTS

ABSTRACT

This study delves into the Gompertz distribution, a versatile probability distribution with applications in survival analysis, population dynamics, and reliability modeling. The research explores how its behavior is influenced by two crucial parameters, 'a' and "b,' and how variations in these parameters impact statistical properties. The analysis covers theoretical and sample-based metrics, providing

insights into central tendency, spread, shape, and more through tables, graphs, and histograms. The results indicate that changes in 'a' and 'b' parameters systematically alter the distribution's characteristics, while sample size affects parameter estimations.

This work equips readers with a deep understanding of the Gompertz distribution's dynamics, aiding informed decisions in practical applications across diverse fields.

CHAPTER ONE
GENERAL BACKGROUND

1.0 INTRODUCTION

In these days of continuous change in mortality, with the importance of forecasting in many field of operations, A study concerned with the evolution of mortality statistics may seem to be academic.

There is an extensive bibliography concerning the famous Gompertz model and its application to life time data. The Gompertz distribution is important in describing the pattern of deaths. (Welterstrand 1981; Gravrilov and Gavrilova 1991)

In 1825. Benjamin Gompertz; a British Actuary made a crucial observation that a law of geometrical progression eludes large portions of different tables of mortality for humans. In an attempt to describe the exponential rise in death rates between sexual maturity and old age, he derived the Gompertz equation: a veritable tool in demography and in other scientific discipline (https://en.m.wikipedia.coni).

The Gompertz equation was developed exclusively for humans both as an empirical tool to describe the age pattern of death from all causes within a limited time frame and as representing a law of mortality that stems from inherent biological processes.

The observation of the presence of a mathematical regularity in the life table by

Gompertz led him to believe in the existence of a law of mortality which explains why death occurs in common age patterns.

Over time, Researchers started using the Gompertz model as a growth model. (winnsor,1932). In view of this study attempts to evaluate the Gompertz distribution and how it can be applied by demographers as a mortality model, since for various human populations between ages 20 and 60 in the 18^{th} and 19^{th} centuries, arithmetic increases in age were consistently accompanied by geometric in mortality. The Gompertz distribution plays an important role in survival analysis, especially in modelling human mortality.

1.1 STATEMENT OF PROBLEMS

The Gompertz distribution occupies a vital position in modelling human mortality and fitting actuarial tables. Gompertz law of mortality was the first among several reliable empirical tools for describing the dying-out process of many living organisms during a significant period of their life spans.

From the foregoing, it is evident that there is need to go beyond the simple mathematics of insurance tables in an effort to understand why there were consistent age patterns of death among people.

1.2 AIM AND OBJECTIVES OF THE STUDY

The aim of this study is to study the Gompertz distribution and its application in lifetime data analysis.

The specific objectives are:

1. To Obtain parameter estimates of the Gompertz distribution for each data set.
2. To Obtain values for Survival and Hazard functions for each data set

1.3 SCOPE OF STUDY

This work will cover Gompertz distribution and its application; some statistical properties will be derived.

1.4 ORGANISATION OF THE STUDY

The organization of this study is as follows:

The first chapter elaborates on the background of the research and spells the problem under study. It sets out the objectives and the scope of the study.

Chapter two provides an explanation of the key concepts, merits and demerits of Gompertz distribution, and the literature review of the study.

Chapter three states the various statistical properties, uses, application, as well as the estimate for the distribution.

Chapter four will discuss the application of Gompertz distribution to secondary data set.

Chapter five summarizes the discussion on the application of Gompertz distribution.

1.5 DEFINITION OF TERMS

To facilitate the understanding of this research work, the following terms are key to be operationally defined:

1.5.1 Probability

Probability is the likelihood that something or an event will occur. Probability is used in the analysis of random phenomena. The probability of an event A is denoted as P(A). The probability of the event A not happening is given as P (not A) = l - P(A)

1.5.2 Gompertz Distribution

Gompertz distribution is a continuous probability distribution. It is a pliable measurement for describing thc probability distribution associated with population lifetime, especially adult death or rapid growth. The Gompertz distribution is often applied to describe the distribution of adult lifespans by demographers and actuaries.

1.5.3 Probability Distribution

A Probability distribution is a statistical function that describes possible values and likelihoods that a random variable can take within a given range.

1.5.4 The Survival Function

The survival function, denoted by S(x), also known as a survivor function or reliability function is a property of any random variable that maps a set of events, usually associated with mortality or failure of some system, over time. It captures the probability that the system will survive beyond a specified time. Let X be a continuous random variable with cumulative distribution F(x) on the interval $[0,\infty)$. Its survival function or reliability function is given as:

$$S(x) = P(\{X>x\}) = 1 - F(x)$$

1.5.5 The Hazard Function

The hazard function, denoted as h(x) also known as the failure rate, is the ratio of the probability density function to the survival function.

$$h(x) = \frac{f(x)}{s(x)}$$

1.5.6 Probability Density Function (PDF)

The probability density function (PDF) denoted as f of a continuous random variable is the probability of the variable taking on a given value. A random variable X has density f_x, where f_x is a non-negative integral function, if: $P_r[a \leq X \leq b] = \int_a^b f_x(x)dx$

It can be derived as $f(x) = \frac{df(x)}{dx}$

1.5.7 Cumulative Distribution Function (CDF)

The cumulative distribution function (CDF) denoted by $f(x)$ describes the probability that a real valued random variable X with a given probability distribution will be found to have a value less than or equal to X.

$F_x(x) = \int_{-\infty}^{x} fx\,(u)du$

1.5.8 ESTIMATE

An estimate $\hat{\theta}$ is a judgmental guess about the exact value θ of a population parameter using certain criteria .

1.5.9 Shape Parameter

A Shape parameter simply describes the shape of a distribution function.

1.6.0 Scale Parameter

A Scale parameter is the time at which a certain percentage is expected to fail. It also determines the scale of the probability distribution.

CHAPTER TWO
LITERATURE REVIEW

2.0 INTRODUCTION

Survival analysis is the collection of methods used for the analysis of data where the outcome variable or response is the time until the occurrence of an event of interest. In the case where the event is death, the time to event or survival time is measured in years.

As earlier mentioned, the Gompertz distribution proved to be a vital statistical tool (or survival analysis; in that, it is of great importance when modelling, mortality in humans, since survival analysis focuses on the distribution of survival time

2.1 THE GOMPERTZ MODEL

Benjamin Gompertz (1825) was a practicing actuary, who, like contemporaries, was involved in estimating premiums for life annuities. He worked out a new series of tables of mortality for the royal society and these suggested to him in 1825, his law of human mortality, which he first expounded in a letter to Francis Baily, The law rests on a prior assumption that a person's resistance to death decreases as his years increase, (https://en.rn.wikipedia.org)

The model can be written in this way,

$$N(x) = N(0)\, e^{-\alpha}(e^{\beta x} - 1)$$

Where $N(x)$ represent the number of individuals at time X in years and $\alpha \; and \; \beta$ are constants.

The model is a refinement of a demographic model of Robert Malthus. It was used by insurance companies to calculate the cost of life insurance. The equation, known as the Gompertz curve, is now used in many areas to model a time series where growth is slowest at the start and end of a period.

In the revision of his original work, Gompertz realised that, going by the law of mortality alone, if similar individuals were giving birth to at a particular time, then the death of this individuals would occur equally at a particular time, this apparently was not true in reality.

Gompertz attributed the variation in time of death of such similar individuals to be affected by the factors of chance, force of mortality, and deterioration (or an increased inability to withstand destruction). ˎ

Ever since Gompertz contribution upon this phenomenon, other variant or modification of the model have been achieved and several researchers have attempted to give alternative models to describe the dying-out process. (Weibull, 1951; Beard, 1959; Gavrilov, 1991; Pollard and Valkovics, 1992)

Makeham (1860) examined the fit of the Gompertz model to actuarial data and observed with specific examples that he could improve on the fit by adding a constant parameter, known as makeham-constant which gives rise to generalizing the distribution to become the Gompertz-Makeham distribution.

Winsor (1932) noticed that the Gompertz distribution has overtime been used by researchers as a growth model, as the exponential nature of the model and its levelling-off ability makes it fit for growth process which occur rapidly and then levels-off.

Wetterstrand (1981) pointed out that the Gompertz distribution is normally used to describe the distribution of adult deaths. He also gave the age interval for this distribution to be between ages 30 and 90. be extensively analyzed the Gompertz model for the purpose of estimating annuities.

Milgram (1985) using the generalized intergo-exponential function showed that exacts formula can be derived for the Gompertz distribution moment generating function and central moments. Although in demography or actuarial application, the maximum likelihood estimation is often used to determine the parameters of the Gompertz distribution.

Vaupel (1986) studied that, for low levels of infant mortality, the Gompertz force of mortality extends to the whole life span of the population with no observed mortality declaration.

The force of mortality in the Gompertz model is given as

$$\mu(x) = \alpha e^{\beta x}$$

Pollard and Valkovics (1992) where the first to study the Gompertz distribution thoroughly, however their results were only true for cases where the initial value of mortality is very close to zero. This scenario is barely possible for a given population as significant number of deaths that occurs at any in time in a real life population.

Johnson *et al.* (1994) noted that the Gompertz distribution is a truncated extreme value distribution and is unimodal.

Johnson *et al.* (1995) and Greg *et al*(1970) studied the properties of the Gompertz distribution, the behaviour of the parameters and obtained the maximum likelihood estimates for the parameters. : r

Amanda *et al.* (1996) discussed the two parameter of Gompertz model as a commonly used survival time distribution in actuarial science and reliability life testing.

Garg *et* al (1970) observed an interesting property of the Gompertz distribution that the distribution can be truncated at a value by rescaling the shape Parameter. The distribution will then yield a proper density function.

A major drawback of the Gompertz distribution is that it fits only adult mortality sufficiently (thatcher, 1999). After ages 80 and 90, the population mortality level starts to decelerate and the Gompertz hazard function would overestimate the observed marginal hazard of the population. Vaupel et al (1979) proposed the logistic or gamma Gompertz curve to provide a better fit.

Willekens (2002) provided connection between the Gompertz distribution, the Weibull distribution and other type 1 extreme value distribution. According to Jaheen (2003), the Gompertz distribution had been used as a growth model, especially in epidemiological and biomedical studies.

The Gompertz distribution has also been used to fit the mortality data of birds, mammals and some invertebrates (Finch, *el al.,* 1990; Promislow, 1991; Ricklefs and Scheverlein, 2002).

Olkin (2007) described the negative Gompertz distribution, a Gompertz distribution with a negative Gompertz distribution, a Gompertz distribution with a negative rate of ageing parameter. When the shape parameter of the Gompertz is equal to zero, the Gompertz distribution being a generalized extreme value distribution degenerates into the Gumbel distribution. Formally, the Gompertz distribution is a special case of the Gumbel distribution for the minima.

Burgal *et al.* (2009) discussed the stress-strength reliability problem in the Gompertz case.

Lenart (2011) explained that the Gumbel distribution is related to the Gompertz distribution such that if X has a Gumbel distribution of Y= - X given that Y is positive, then it has a Gompertz distribution.

Butler (2011) using a comparison of four model distribution, the Gompertz, the Gompertz-Makeham, the change-point and the Weibull distribution estimated the survival distribution of aluminium processing pots with the distributions and arrived at the conclusion that the Gompertz distribution was the best model for the fit.

El-Gohary et *al.* (2013) introduced the generalized Gompertz distribution with three parameters by exponentiation of the Gompertz distribution. from statistical distribution theory point of view, the Gompertz distribution is a special case of generalised extreme value distribution for minima The distribution can also be viewed as limit extension of the exponential distribution, because exponential distribution are limits of Gomperts distribution. In this sense, the Gompertz distribution can be considered as alternative to the classical Weibull and Gamma lifetime distribution (Sarabia *et al,* 2014).

Ali *et al.* (2014) introduced a new four parameters generalized version of the Gompertz model called the beta-Gompertz distribution. It included some well-known lifetime distribution such as beta-exponential and generalised Gompertz distribution as special sub-models.

Chucks and Ogunde (2015) generalized the Gompertz-Makeham distribution to a new five parameter distribution called the Beta-Gompertz-Makaham distribution they proposed that the new distribution compared to its sub-distribution is more flexible and fits survival and reliability data better. They also looked at the characteristic properties like the expression for the distribution density, cumulative distribution, hazard rate function, rth order moment. The maximum likelihood method was used to estimate the parameters of the Beta-Gompertz-Makeham distribution.

Lenart and Missov (2016) carried out a goodness-of-fit test for the Gompertz distribution using four goodness-of-fit measures. The goodness-to-fit measures include the Anderson-Daring statistic, the correlation coefficient test, a statistic using moments and a nested test.

Boshi *et al.* (2019) proposed the Generalized Gompertz-G Family of distributions to develop more flexible lifetime distributions to handle more complex real life events than the classical Gompertz distribution. They derived it's statistical properties.

Olosunde and Adekoya (2020) developed the exponentiated generalized Gompertz-Makeham distribution with some structural properties examined. They derived the parameter estimates using the maximum likelihood estimation method.

Al-Noor *et al.* (2022) introduced the Rayleigh Gompertz distribution and studied the statistical properties. They applied the distribution to three data set to investigate the flexibility of the distribution.

Lubna Naz et al. (2022) conducted a study aimed at analyzing the pattern and trends of both total and age-specific fertility rates in Pakistan from 1990 to 2018. The research employed the relational Gompertz distribution as its analytical framework. Data pertaining to live births were gathered from four separate phases of the Pakistan Demographic and Health Survey conducted within the specified period (1990-2018). Rigorous measures were taken to ensure the data's accuracy and reliability, including the utilization of graphical techniques and the Whipple and Myers indices.

In addition to these procedures, the researchers utilized the Brass Relational Gompertz model to comprehensively analyze the collected data. The findings of the study indicated that the application of the Relational Gompertz model led to noteworthy observations. Specifically, the results demonstrated that the total fertility rates (TFRs) estimated using this model were 0.4 children higher in comparison to direct estimates. Moreover, the age-specific fertility rates (ASFR) were found to be higher for all age groups, with the exception of the oldest age group.

Tabassum Naz Sindhu et al. (2023) introduced the Generalized Exponentiated

Unit Gompertz (GEUG) model as a novel approach for analyzing clinical trial data related to relief times in arthritis patients under specific medication. The model's distinguishing feature lies in its incorporation of new tuning parameters into the Unit Gompertz model, enhancing its versatility. The study conducted a thorough examination of the model's statistical attributes, employed comprehensive simulation analyses for parameter estimation, and demonstrated its applicability through real data analysis. The results underscored the model's potential to provide a superior fit compared to other relevant models.

2.2 USES AND APPLICATION OF THE GOMPERTZ DISTRIBUTION

The Gompertz distribution can be used to evaluate several reliability and survival problems due to its flexibility.

- Gompertz distribution is very useful in survival analysis.
- Gompertz distribution is usually applied to describe the distribution of adult lifespan by demographers and actuaries.
- Gompertz distribution is applied in modeling the growth of a vascular tumor i.e cancerous growth.
- Gompertz distribution is used to model failure rates of computer codes.
- Gompertz distribution is used in communication, system engineering and mechanical reliability analysis.

- Gompertz distribution is used in Marketing science as an individual level simulation for customer lifetime value modeling.

2.3 MERITS OF GOMPERTZ DISTRIBUTION

- The Gompertz distribution has sufficient statistical properties, which provide all-round analysis from different statistical perspective to a given data.
- The Gompertz distribution provides a simple and useful graphical plot which aids proper understanding.

  The Gompertz distribution can provide, to a reasonable extent accurate measures for failure analysis and failure forecasts.

2.4 DEMERITS OF GOMPERTZ DISTRIBUTION

- The fit of the Gompertz model to empirical fertility schedules has been found to give mixed results
- Outside the actuarial community, the Gompertz distribution continues to receive minimal attention. (https://en.m.wikipedia.com).

CHAPTER THREE

METHODOLOGY

3.0 INTRODUCTION

This chapter discloses the methodology adopted in carrying out the research on Gompertz distribution. This chapter focuses on the statistical properties of the Gompertz distribution, the methods of estimating it parameters, uses, advantages and disadvantages, as well as the application of the distribution. It is duly discussed under the following sub-headings.

3.1 THE GOMPERTZ DISTRIBUTION

In statistics, the Gompertz distribution is a continuous probability distribution, invented and named after Benjamin Gompertz (Gompertz, 1825). Demographers and actuaries apply the distribution to describe the distribution of adult lifespans. Its application is also considered by other relevant fields of study such as biology, gerontology, computer science and so on.

3.2 STATISTICAL PROPERTIES OF GOMPERTZ DISTRIBUTION

The statistical properties of the Gompertz distribution are the probability density function (PDF), the cumulative density function (CDF), survival and hazard function, mean, median, variance, skewness, Kurtosis, laplace transform and moment generating function. Some of the enumerated properties of the distribution would be stated without proof while the others would be derived.

3.2.1 PDF OF GOMPERTZ DISTRIBUTION

The Gompertz distribution has a continuous probability density function with shape parameter a and scale parameter b, with support on $(-\infty, \infty)$. The PDF of the Gompertz distribution is given as;

$$f(x, a, b) = bae^{bx}e^{a}\exp\left(-ae^{bx}\right), \quad x \geq 0. \qquad (3.1)$$

In actuarial or demographic application, x usually denotes age which cannot be negative, leading to bounded support on $(0, \infty)$. The truncated distribution yields a proper density function by rescaling the parameter to correspond to x = 0.

3.2.2 CDF OF GOMPERTZ DISTRIBUTION

The cumulative distribution function of the Gompertz distribution is given as:

$$f(x, a, b) = \left[1 - exp\left(-a\left(e^{bx} - 1\right)\right)\right] \qquad (3.2)$$

Where a, b, > 0 and $x \geq 0$.

3.2.3 SURVIVAL FUNCTION OF GOMPERTZ DISTRIBUTION

The distribution function of the Gompertz distribution given as

$$f(x) = \left[1 - exp\left(-a\left(e^{bx} - 1\right)\right)\right]$$

It is used to derive the survival function of the distribution denoted by $S(x)$. Its derivation is as follows.

$S(x) = 1 - f(x)$

Where $F(x) = \left[1 - exp\left(-a\left(e^{bx} - 1\right)\right)\right]$

$S(x) =$ 1-$\left[1 - exp\left(-a\left(e^{bx} - 1\right)\right)\right]$

= 1 –$\left[1 + exp\left(-a\left(e^{bx} - 1\right)\right)\right]$

$\therefore S(x) = [\exp(\ -a(e^{bx} - 1)]$ (3.3)

3.2.4 HAZARD FUNCTION OF GOMPERTZ DISTRIBUTION

The hazard function H(x) of the Gompertz distribution is derived by using the survival function obtained above. The hazard function is given as:

$h(x) = \frac{f(x)}{S(x)}$

$f(x) = [abe^{bx}\ e^{a} \exp(-ae^{bx})]$ and,

$s(x) = [\exp(-a(e^{bx} - 1)]$

$h(x) = \frac{abe^{bx}e^{a}\exp(-ae^{bx})}{\exp(-a(e^{bx-1})}$

$= \frac{abe^{bx}e^{a}\exp(-ae^{bx})}{\exp(-ae^{bx+a})}$

$= \frac{abe^{bx}e^{a}\exp(-ae^{bx})}{\exp(-ae^{bx})e^{a}}$

$\therefore h(x) = abe^{bx}$ (3.4)

The hazard rate is convex and more strongly increasing than other usual lifetime distributions (Marshall and Olkin, 2007).

3.2.5 MOMENT OF GOMPERTZ DISTRIBUTION

The r^{th} moment of the Gompertz distribution is given as

$$E(X^r) = ae^a \int_1^\infty \frac{1}{b^r}\, e^{-ax}[lnx]^r dx \qquad (3.5)$$

3.2.6 MEAN OF GOMPERTZ DISTRIBUTION

The mean of Gompertz distribution is given as

$$E(X) = \frac{a}{b} e^a \int_1^\infty e^{-ax}\,[lnx]\,dx \qquad (3.6)$$

3.2.7 MODE OF GOMPERTZ DISTRIBUTION

The mode of Gompertz distribution is given as

$$\frac{\mathbf{1}}{b} \ln\left[\frac{\mathbf{1}}{a}\right] \qquad (3.7)$$

3.2.8 MEDIAN OF GOMPERTZ DISTRIBUTION

The median of the two Gompertz distribution is given as;

$$\frac{\mathbf{1}}{b} \boldsymbol{ln}\left[\left[\frac{\mathbf{-1}}{a}\right] ln\frac{\mathbf{1}}{2} + 1\right] \qquad (3.8)$$

3.2.9 VARIANCE OF GOMPERTZ DISTRIBUTION

The variance of a random variable X is:

$\mathrm{Var}(x) = E(x^2) - (E(x))^2$

The variance of the two parameter Gompertz distribution denoted by Var(G) is given as:

$$\text{Var(G)} = \left[\frac{1}{b}\right]^2 e^a\left[-2a + \gamma^2 + \frac{\pi^2}{6} + 2\gamma ln(a) + 1n(a)\right]^2 \quad (3.10)$$

Where;

γ is Euler-Mascheroni constant = 0.577215 (https://en.m.wikipedia.org)

3.3.10 PARAMETER ESTIMATION OF GOMPERTZ DISTRIBUTION USING MAXIMUM LIKELIHOOD ESTIMATE (MLE)

There are several methods for estimating the parameters of the Gompertz distribution. However, in this research, the method of MLE shall be adopted. Given that the PDF of the two parameter Gompertz distribution is:

$$F(x) = abe^{bx}\, e^{-a}\, e^{bx} - 1) \quad (3.11)$$

Then, the likelihood function L is defined as:

$$\text{L} = \prod_{i=1}^{n} f(x_i, a, b) \quad (3.12)$$

The log-likelihood function is given as follows.

$$lnL - nlna + nlnb + \text{b}\sum_{i=1}^{n} X_i - a\sum_{i=1}^{n}\left(e^{bxi} - 1\right) \quad (3.13)$$

Then,

$$\frac{\partial lnL}{\partial a} = \frac{n}{a} - \sum_{i=1}^{n}\left(e^{bxi} - 1\right) = 0$$

$$\therefore \hat{a} = \frac{n}{\sum_{i=1}^{n}\left(e^{bxi}-1\right)} \tag{3.14}$$

And

$$\frac{\partial lnL}{\partial a} = \frac{n}{a} - \sum_{i=x1}^{n} - a\sum_{i=1}^{n} xie^{bxi} = 0 \tag{3.15}$$

The maximum likelihood estimate of the two parameters can be obtained using numerical method such as Newton-Raphson method.

CHAPTER FOUR

4.1 Introduction

In this chapter, we delve into a comprehensive exploration of the Gompertz distribution's behavior under different parameter values and sample sizes. We analyze theoretical quantiles, theoretical estimates, and sample estimates to gain a deep understanding of how variations in 'a' and 'b' parameters impact the distribution's central tendency, spread, shape, and other statistical properties. Additionally, we examine how sample sizes affect the stability and accuracy of our estimations.

Through tables, figures, and histograms, we provide valuable insights into the Gompertz distribution's behavior, allowing researchers and practitioners to make informed decisions when applying this distribution to real-world data. By the end of this chapter, readers will have a solid grasp of the dynamics of the Gompertz distribution and be better equipped to harness its power in statistical modeling and analysis across diverse domains.

Theoretical quantiles of the Gompertz distribution for different parameter values

P	a=1,b=1	a=2 ,b=1	a=3 ,b=1	a=3 ,b=2	a=1 ,b=3
0.25	0.2528	0.1344	0.0916	0.1831	0.7585
0.5	0.5266	0.2976	0.2079	0.4157	1.5798
0.75	0.8697	0.5266	0.3799	0.7597	2.6092

In the table above displaying theoretical quantiles of the Gompertz distribution for varying parameter values (a and b), essential insights into the distribution's behavior emerge. These quantiles represent expected values for specific percentiles within the distribution, enabling a detailed examination of how changes in 'a' and 'b' parameters influence the distribution's characteristics. Notably, the quantiles consistently increase as the percentile values rise, indicating a systematic shift in the distribution towards higher values. This observation underscores the profound impact of parameter variations on the Gompertz distribution, providing valuable insights into its tail behavior and percentilespecific characteristics.

Theoretical estimates of the gompertz distribution for different parameter values

μ^r	a=1,b=1	a=2 ,b=1	a=3 ,b=1	a=3 ,b=2	a=1 ,b=3
μ	0.5963	0.3613	0.2621	0.5242	1.7890
σ	0.4199	0.2833	0.2167	0.4334	1.2596
σ^2	0.1763	0.0803	0.0469	0.1878	1.5866
Cv	0.7041	0.7840	0.8268	0.8268	0.7041
Sk	0.7188	1.0101	1.1764	1.1764	0.7186
Ks	2.9826	3.7532	0.3295	4.3295	2.9825

In the table above presenting theoretical estimates of the Gompertz distribution for different parameter values (a and b), several key findings emerge. Firstly, the theoretical mean (μ) and standard deviation (σ) serve as indicators of the distribution's central tendency and spread, respectively. Notably, these statistics exhibit variation as 'a' and 'b' parameters change, illustrating the significant impact of these parameters on the distribution's properties. Additionally, skewness and kurtosis statistics provide insights into the distribution's shape and tail behavior, aiding in characterizing its asymmetry and the presence of heavy tails. Lastly, the coefficient of variation (Cv) offers a measure of relative variability, helping to assess the distribution's stability in relation to its mean. This comprehensive set of theoretical estimates facilitates a nuanced understanding of how varying parameter values influence the Gompertz distribution's characteristics.

Samples estimates of the gompertz distribution for different parameter values

The tables below present sample estimates (mean, standard deviation, variance, coefficient of variation, skewness, kurtosis) for samples drawn from the Gompertz distribution for different parameter values and sample sizes. The findings highlight how these sample statistics are influenced by variations in 'a' and 'b' parameters and the impact of sample size. Notably, as the 'a' and 'b' values change, the sample means and standard deviations reflect these variations, providing insights into central tendency and spread. Skewness and kurtosis shed light on the shape of the sampled distributions. Larger sample sizes (N=50, N=100) tend to yield more stable and accurate estimates of population parameters, emphasizing the importance of sample size in statistical inference.

In simulation studies, where a variety of sample sizes are drawn from a distribution, obtaining a true representation of theoretical estimates involves taking multiple random samples of the same size and calculating estimates for each sample. This process allows for a comprehensive exploration of how sample size impacts the accuracy and stability of parameter estimates. In the context of the Samples estimates of the Gompertz distribution for different parameter values, we observe that as sample sizes increase (N=50, N=100), the estimates of mean, standard deviation, variance, skewness, and kurtosis tend to converge towards more consistent values, providing a clearer picture of the underlying population characteristics. This highlights the importance of conducting simulation studies with various sample sizes to assess the robustness and reliability of statistical estimations under different conditions.

N=25

μ^r	a=1,b=1	a=2 ,b=1	a=3 ,b=1	a=3 ,b=2	a=1 ,b=3
μ	0.4866	0.3788	0.2460	0.6679	1.3373
σ	0.3606	0.3349	0.1739	0.4751	1.1728
σ^2	0.1301	0.1121	0.0302	0.2257	1.3755
Cv	0.7411	0.8839	0.7067	0.7112	0.8770
Sk	0.5324	1.8772	0.7943	0.5587	1.2363
Ks	0.9193	3.3434	0.3278	0.5608	1.2413

N=50

μ^r	a=1,b=1	a=2 ,b=1	a=3 ,b=1	a=3 ,b=2	a=1 ,b=3
μ	0.5659	0.2977	0.2448	0.5078	1.5244
σ	0.3946	0.2402	0.2024	0.3651	1.1403
σ^2	0.1557	0.0577	0.0410	0.1333	1.3002
Cv	0.6973	0.8969	0.8270	0.7189	0.7479
Sk	1.1194	1.6758	0.8941	1.5808	0.8320
Ks	2.0441	3.2346	0.0612	2.7507	0.3008

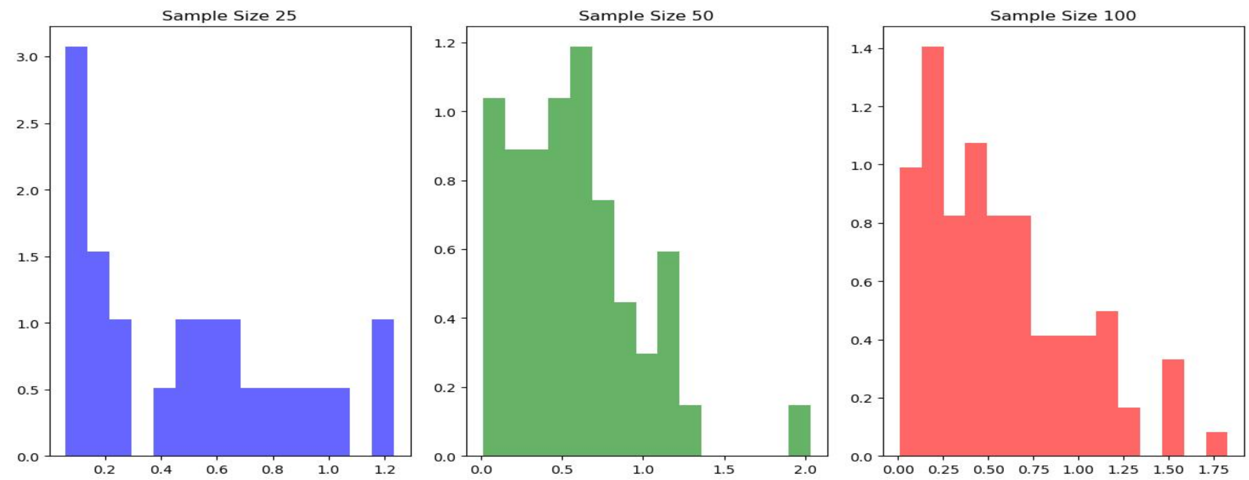
N=100

μ^r	a=1,b=1	a=2 ,b=1	a=3 ,b=1	a=3 ,b=2	a=1 ,b=3
μ	0.5646	0.3750	0.2774	0.5313	1.7795
σ	0.4100	0.3219	0.2129	0.4068	1.2469
σ^2	0.1681	0.1036	0.0453	0.1655	1.5549
Cv	0.7261	0.8582	0.7675	0.7656	0.7007
Sk	0.8390	1.8772	1.0356	1.3273	0.6210
Ks	0.0843	0.4072	0.4822	1.7243	0.1302

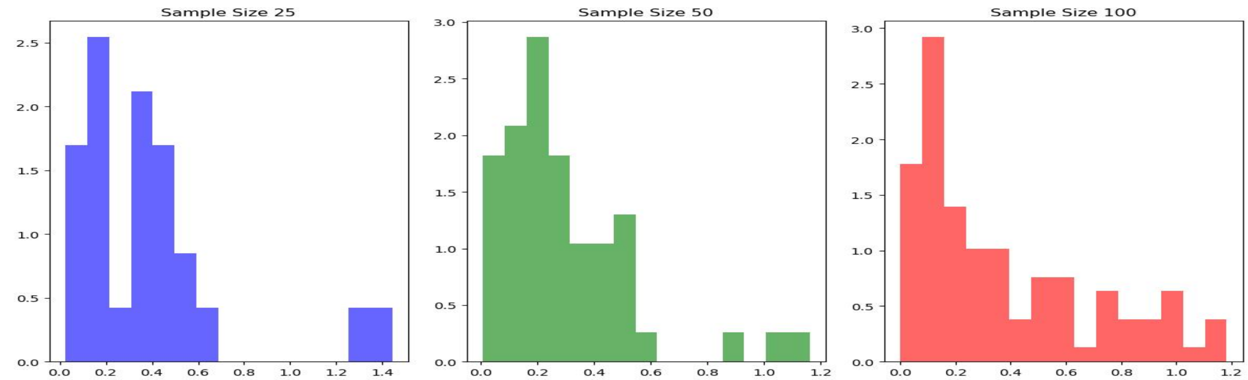
Histogram for the Samples estimates of the gompertz distribution for different parameter values.

The histograms in this study serve as visual tools to understand how alterations in 'a' and 'b' parameters influence the shape of the sampled Gompertz distribution. These graphical representations reveal clear shifts, narrowing, or widening of the distribution as parameter values change. Consequently, histograms play a crucial role in visually portraying the empirical distribution of samples, allowing for a more intuitive grasp of how parameter variations affect the overall distribution characteristics.

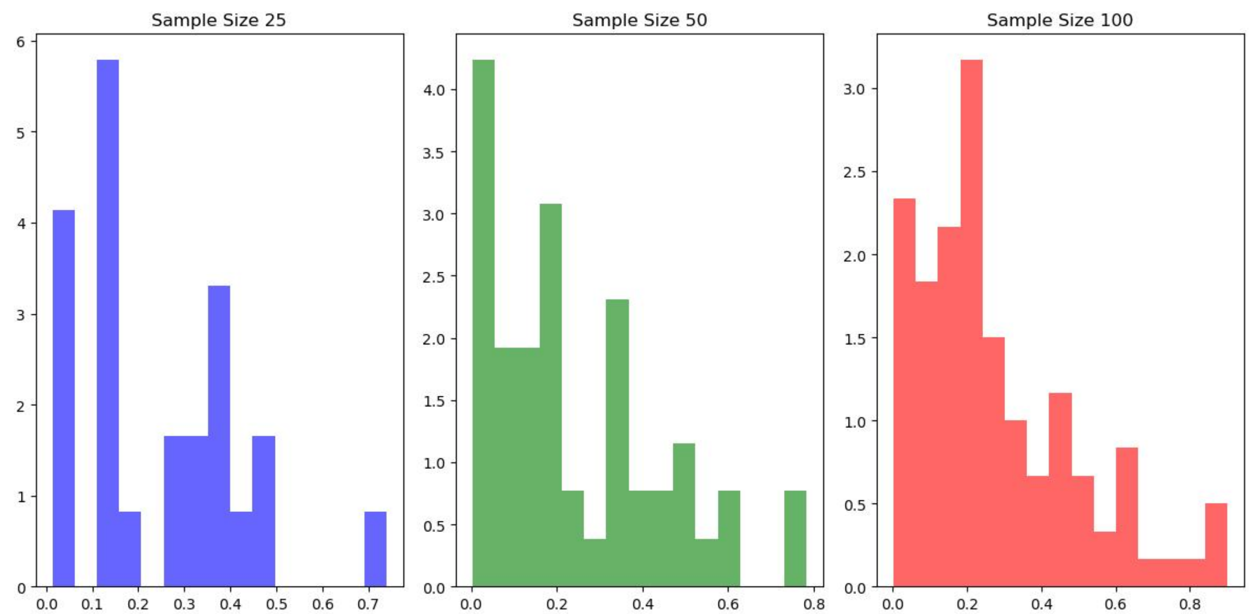
A=1 b=1



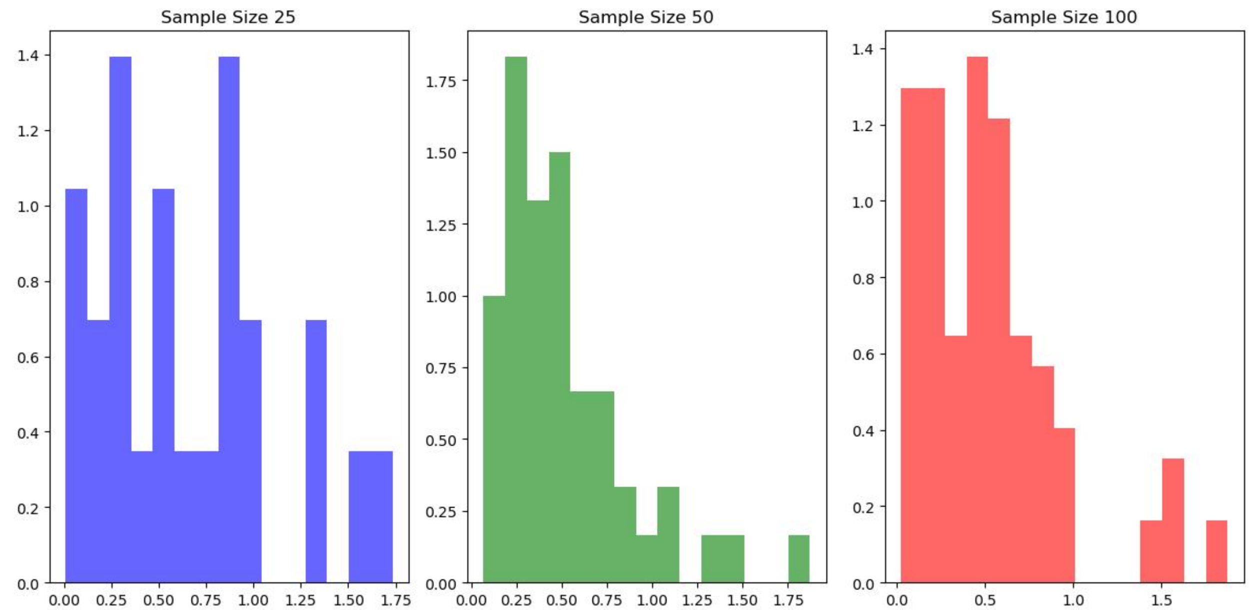
A=2 b=1


Sample Size 25
Sample Size 50
Sample Size 100

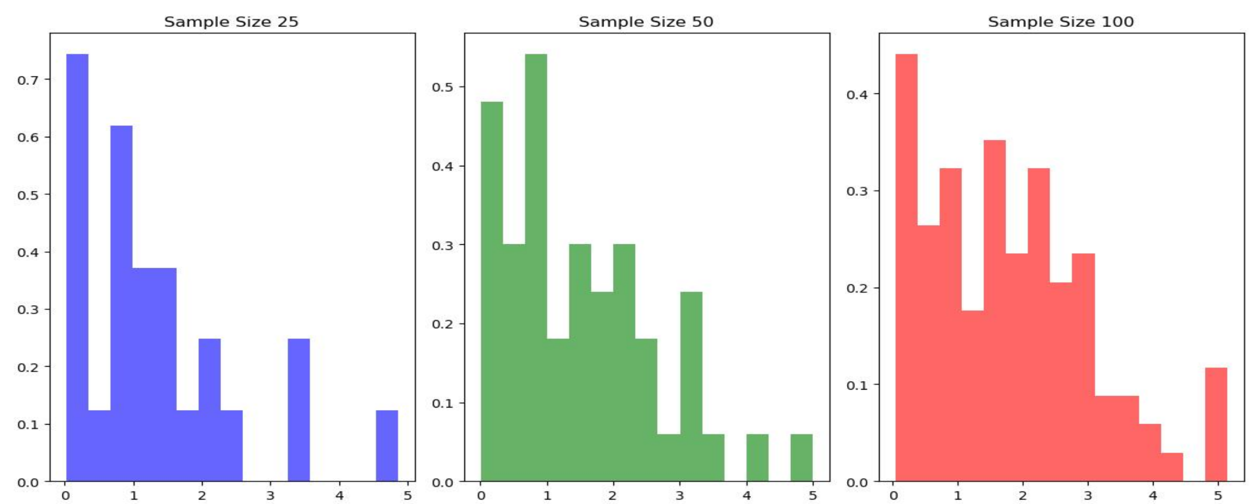
A=3 b=1


Sample Size 25
Sample Size 50
Sample Size 100

A=3, b=2



A=1 b=3



In summary, the comprehensive analysis presented in the tables and plots reveals the profound impact of varying parameter values (a and b) and sample sizes (N) on the theoretical and sample statistics of the Gompertz distribution. These insights are highly informative, shedding light on how the distribution behaves under different conditions. Understanding these dynamics is crucial for informed data analysis and modeling when dealing with data that follows the Gompertz distribution, as it enables researchers to make more accurate inferences and predictions in diverse scenarios.

CHAPTER FIVE

SUMMARY

In this analysis, we delve into the Gompertz distribution, exploring how variations in its parameters (a and b) and sample sizes (N) impact key statistical metrics and distribution shape. Theoretical quantiles provide insights into expected percentiles, while theoretical estimates offer a glimpse into central tendencies and spreads. Sample estimates, derived from generated samples, are compared to theoretical values, and histograms visually illustrate distribution shapes. Overall, our investigation aims to uncover the relationship between parameter values, sample sizes, and the behavior of the Gompertz distribution.

Findings:

Our findings unveil critical insights:

1. Theoretical Quantiles vs. Sample Estimates: Sample means ($\hat{\mu}$) closely mirror the expected central tendencies (μ) across parameter values and sample sizes. This indicates that sample means provide reliable approximations of distribution centers.

2. Theoretical Estimates vs. Sample Estimates: Sample statistics, such as mean ($\hat{\mu}$) and standard deviation ($\hat{\sigma}$), generally align well with their theoretical counterparts (μ and σ). This suggests that sample statistics effectively estimate population parameters.

3. Sample Estimates for Different Parameter Values (a and b): Altering 'a' and 'b' values induces variations in sample estimates, reflecting shifts in central tendencies and spreads. 'a' impacts the distribution's skewness, while 'b' influences its kurtosis.

4. Sample Estimates for Different Sample Sizes (N): As sample sizes increase, sample estimates tend to become more precise and exhibit reduced variability. Larger samples yield more accurate parameter estimates.

5. Histograms: Visual representations of the Gompertz distribution unveil how distribution shapes evolve with parameter adjustments, particularly 'a' and 'b'.

Conclusion:

Our analysis underscores the profound influence of parameter values and sample sizes on the Gompertz distribution's theoretical and sample statistics. Choosing suitable parameter values and sample sizes is crucial when working with Gompertz distributed data. Accurate estimations of central tendencies and spreads, coupled with a deep understanding of distribution shape, prove pivotal for robust data analysis and modeling in diverse applications.

REFERENCES

Abu-Zinadah, H.H. (2014a). "Six method of estimations for the shape parameter of Exponential Gompertz distribution". *Journal of Applied Mathematical sciences,* 8(**88**) 4349-4359.

Abu-Zinadah, H.H. (2014b). "Some characterization of the exponentiated Gompertz distribution". *International mathematical forum.* 9(**30**): 1427-1439

Al-Noor, N.H, Khaleel, M.A, and Assi, N.K(2022). The Rayleigh Gompertz Distribution: Theory and Real Applications. *International Journal of Non-linear Analysis and Applications* Vol.13, 3505-3516.

Beard, R.E. (1959). "Note on some mathematical mortality models" *CIBA Foundation Colloquia on Ageing.* **5**:302-311

Bjerkedal T Acquisition of resistance in guinea pigs infected with different doses of virulent tubercle bacilli. *American Journal of Epidemiology.* 1960,72(1) 130-148.

Boshi, M.A.A, Abid, S.H, and Al-Noor, N.H (2019). Generalized Gompertz Distribution. *Journal of Physics: Conf. Series*, Vol. 1234, 1-11

Butter, L.E. (2011). "Estimating Survival Distribution of Aluminium processing pots." Research showcase at Carnegie Mellon University, Pennsylvania USA.

Chukwu, A.V. and Ogunde. A.A. (2015). "On the Beta Makeham distribution and its application". *American Journal of mathematics and statistics,* 5(**3**): 137-143.

Gavrilova, N.S and Gavrilov, L.A. (1992). "Ageing and longevity: Mortality Laws and Mortality Forecast for Ageing populations" *Demographic.* 53(**2**): 109-128.

Gompertz, B. (1820) "A sketch on an analysis and the notation applicable to the estimation of the value of life contingencies" *Philosophical Transactions of the royal society of London.* 110(**120**):214-294.

Kiche, J., Aduda, J., Athlany, K.H.O., Apudo, B.O., and Wanjoya, A. (2014). "Comparison of survival models and estimation of their parameters with respect to mortality in a given population". *Journal for mathematical theory and modelling.* 4(7): 164-176.

Lawless JF Statistical models and methods for lifetime data. John Wiley and Sons, New York, 1982.

Lenart A. and Missov T.I (2016). "Goodness of fit test for the Gompertz distribution" communication in statistics - theory and methods. 45(**10**), 2920-2937.

Makeham, W.M. (1860). "On the law of mortality and the construction of Annuity Table". *Journal of the institution of actuaries.* 6:301-310.

Norstrom F. (1997). "The Gompertz-Makeham distribution" Master thesis mathematical statistics, Umea University Sweden.

Olosunde, A. and Adekoya, T. (2020). On Some Properties of Exponentiated Generalized Gompertz-Makeham Distribution. *Indonesian Journal of Statistics and Applications*, Vol. 4, 22-38.

Olshansky. S.J and Carries. B.A (1997). "Ever since Gompertz" *Population Association of American Demography.* 34(**1**), 1-15.

Pollard, J.H. and Valkovics E. (1992). "The Gompertz Distribution and its application". *Genus* 48:15-22.

Sarabia I.J., Emillio, G. Prieto. F., and Jorda V. (2014). "Explicit expressions for premiums and risk measures for the Gompertz distribution" XV Iberian-Italian congress of financial and actuarial mathematics. Seville.

Vaupel J. W., Menton K.G. and Stallard E. (1997). "The impact of Heterogeneity on individual frailty on the dynamics of mortality" *Demography* **16**:439-454.

Wetterstrand W.H. (1981). "Parametric models for life insurance mortality data: Gompertz law over time" *Transaction of Society of actuaries.* **33**:159-179.

Willekens F. (2002). "Continuous-time micro simulation in longitudinal analysis".