

THE ZERO-TRUNCATED POISSON-WEIGHTED EXPONENTIAL DISTRIBUTION WITH APPLICATIONS

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

MASTER OF SCIENCE

IN

STATISTICS

UNIVERSITY OF REGINA

By

Jin Qin

Regina, Saskatchewan

October 2021

© Copyright 2021: Jin Qin

UNIVERSITY OF REGINA
FACULTY OF GRADUATE STUDIES AND RESEARCH
SUPERVISORY AND EXAMINING COMMITTEE

Jin Qin, candidate for the degree of Master of Science in Statistics, has presented a thesis titled, ***The Zero-Truncated Poisson-Weighted Exponential Distribution with Applications***, in an oral examination held on October 7, 2021. The following committee members have found the thesis acceptable in form and content, and that the candidate demonstrated satisfactory knowledge of the subject material.

External Examiner:	*Dr. Yiyu Yao, Department of Computer Science
Supervisor:	*Dr. Andrei Volodin, Department of Mathematics & Statistics
Committee Member:	*Dr. DianLiang Deng, Department of Mathematics & Statistics
Chair of Defense:	*Dr. Na Jia, Faculty of Engineering & Applied Science

*via ZOOM Conferencing

Abstract

This research proposes a new distribution for non-zero count data, namely the zero-truncated Poisson-weighted exponential distribution (ZTPWE). The Poisson-weighted exponential distribution (P-WE) has been proved to be a flexible two-parameter distribution; therefore, Zero-truncated models can be used to investigate data without zero counts. The combination of two such methods will be discussed in two parts. In the first part (the theoretical part), the probability mass function is derived from two methods. Then theoretical properties of the zero-truncated Poisson weighted exponential distribution are discussed: such as probability generating function, moment generating function, characteristic function, and moments. Furthermore, the method of maximum likelihood estimation is applied to estimate the parameters. In the second part, software simulations and fittings of two real data sets are discussed. The performance of the new model will be compared with other proposed zero-truncated models.

Acknowledgements

Many people have helped, and the following I would like to mention.

I want to express my sincere appreciation to my supervisor, Dr. Andrei Volodin, for guiding me during my program.

I am grateful to study at the University of Regina. The staff, classmates, and professors in Mathematics & Statistics department are beyond perfect.

I would like to thank my former roommates, Guanyi Wang and Meng Wang, who have accompanied me to get through all those dark and cold Saskatchewan mornings and nights.

Finally, I want to thank my parents for their continuous encouragement and support, and especially my grandma, who inspired me since the very beginning.

Notations

In the following, we will use these notations:

WE weighted-exponential distribution;

$P - WE$ Poisson-weighted exponential distribution;

ZTP zero-truncated Poisson distribution;

$ZTPL$ zero-truncated Poisson-Lindley distribution;

$ZTTPPL$ zero-truncated two-parameter Poisson-Lindley distribution;

$ZTPWE$ zero-truncated Poisson-weighted Exponential distribution;

$Poi(x)$ the Poisson distribution with parameter x ;

pmf the probability mass function;

pgf the probability generating function;

mgf the moment generating function;

$\psi(x)$ the moment generating function of variable x ;

$\varphi(x)$ the characteristic function of variable x ;

$p(i)$ the probability of i th event

$f(x)$ the probability mass function (pmf) of the distribution;

ImX image of set X ;

$h(x)$ the probability density function (pdf) of the distribution;

$F_X(x)$ the cumulative distribution function with parameter x ;

cdf the cumulative distribution function;

$G(x)$ the probability generating function with parameter x ;

$E(x)$ the expected value of variable x ;

$Var(x)$ the variance of variable x ;

$L(\alpha, \beta; x)$ the likelihood function;

MLE maximum likelihood estimator;

MSE the mean squared error of an estimator of a parameter;

AIC the akaike information criterion;

AD the Anderson-Darling test goodness of fit.

Contents

Abstract	i
Acknowledgements	ii
Notations	iii
Table of Contents	v
1 INTRODUCTION	1
1.1 Truncated Data and Distributions	1
1.2 Research Outline	2
1.3 Research Objectives	3
1.4 Research Advantages	3
2 THEORETICAL BACKGROUND AND REVIEW OF LITERATURE	4
2.1 Theoretical Background	4

2.1.1	Zero-Truncated Distribution	4
2.1.2	Poisson-Weighted Exponential Distribution	6
2.1.3	Conditional Random Variable Sampling	9
2.1.4	Cumulative Functions, Unimodal and Logconcave	10
2.2	The Theory of Point Estimation	11
2.2.1	Moment and Moment Generating Function	11
2.2.2	Maximum Likelihood Estimation	13
2.2.3	Mean Squared Error and Bias	15
2.2.4	AIC, AD Test and p-value	16
2.3	Review of the Related Literature	16
3	THEORETICAL PART	19
3.1	ZTPWE model	19
3.1.1	Plots of pmf's	22
3.1.2	Cumulative Distribution Function	30
3.2	Theoretical Probabilistic Properties of the ZTPWE Distribution . . .	31
3.2.1	The Probability Generating Function of ZTPWE Distribution	31
3.2.2	The Moment Generating Function of ZTPWE Distribution . .	32
3.2.3	The characteristic Function of ZTPWE Distribution	33
3.2.4	Methods of Moments	33
3.2.5	Parameter Estimation	36

4	SIMULATION and APPLICATION	38
4.1	Simulation Result	39
4.2	Application	40
4.2.1	Plots of Observed Values and Models	45
5	CONCLUSIONS AND DISCUSSIONS	46
5.1	Theoretical Part	46
5.1.1	Advantage for using ZTPWE distribution	46
5.1.2	Progress for Construction of ZTPWE Model	47
5.2	Response of Simulation	48
5.3	Interpretation of Application	48
5.4	Future Study and Expectations	49
	Bibliography	50
	Appendix. R Source Code	54

Chapter 1

INTRODUCTION

1.1 Truncated Data and Distributions

In a period of time, certain events that have been recorded form a count data. For example, the number of road hazards a driver has in a month intuitively contains many zeros, and such data could not be observed very well with ordinary distributions; such data sets follow zero-inflated distributions. Another situation might give us the opposite result. Consider the number of people in a driving car. Intuitively, such data sets contain no 0s at all; to model such data sets, zero omission distributions are proper to be applied. Such distributions are called zero-truncated or positive distribution. The purpose of this study is to show that the modified distributions are recommended so that better performance is guaranteed. Some modified zero-truncated models that have been proposed already, and they all appear to have advantages in data modeling. Gupta and Kundu have demonstrated that the Poisson-weighted exponential(P-WE)

distribution is a better fit than some other distributions for certain dispersed data sets. This research is to establish ZTPWE distribution and determine its performance for positive dispersed data sets.

1.2 Research Outline

In the following chapter, we go over a few theoretical probabilistic concepts that will be used to establish ZTPWE distribution.

In Chapter 3, we first derive the pmf of ZTPWE distribution by two methods, plot it, then obtain other properties such as cumulative distribution function(cdf), moment generating function(mgf) and moments.

In Chapter 4, we simulate 1000 times for two parameters to obtain MSE and bias. The sample size of the simulation will be $n = 20, 50, 100, 200, 500$. The BFGS method in optimx function (Nash and Varadhan) of R language [11] is applied to estimate the parameters of the ZTPWE distribution. Compare average MSE and average bias for parameters α and β , and we consider the most common confidence level 95%. R-programming(version 3.6.3) is used to run simulations. Then two real data sets are considered to fit with different zero-truncated distributions. The criteria information such as AIC and p -values are compared.

In the last chapter, we discuss the significance of our new model and the future study expectations.

1.3 Research Objectives

1. Construct the ZTPWE distribution, derive theoretical properties such as pmf, mean, variance, etc.
2. Simulate random samples and look at Mean Square Error(MSE) and Maximum Likelihood Estimator (MLE) of parameters.
3. Apply the ZTPWE distribution to real data sets, and compare it to other zero-truncated distributions. Analyze the result.

1.4 Research Advantages

This study has the following advantages:

1. Since P-WE distribution has been derived, the ZTPWE distribution can be derived directly.
2. Several zero-truncated distributions are proposed by other researchers. Thus it is relatively easy to make comparisons among zero-truncated distributions.

Chapter 2

THEORETICAL BACKGROUND AND REVIEW OF LITERATURE

2.1 Theoretical Background

2.1.1 Zero-Truncated Distribution

It is possible that in a data set, zero counts appear frequently or none at all. Traditional distributions such as Poisson or binomial distributions would have biased parameters to model such data sets. Zero-inflated distributions can be applied to model such data set that contain lots of zeros. For data sets that contain no zeros, zero-truncated distributions are proper to model such data sets. There are several distributions that have been studied for none-zero data. The zero-truncated Poisson

(ZTP) distribution was developed in 1952 (David and Johnson, p 275-285). The zero-truncated Poisson-Amarendra distribution was introduced in 2017 (Shanker, pp. 82-92), the zero-truncated Poisson-Akash distribution was introduced in 2016 (Shanker, pp. 227-236), and the zero-truncated two-parameter Poisson-Lindley distribution was introduced in 2017 (Shanker and Mishra, pp. 85-95).

The first way to derive the zero-truncated pmf is: by definition, the probability that a point chosen at random lies in a subregion A of Ω is the ratio $\text{measure}(A)/\text{measure}(\Omega)$ (Rohatgi, Saleh, p. 2), say p_i , for $i = 1, 2, 3, \dots$. Assume there is a data set contains count variables that include zeros and other integers, and suppose p_i is the probability of the i th count, for $i = 1, 2, 3, \dots$, then by definition, the collection of numbers p_i is the pmf of the original distribution, satisfying for all i and $\sum_{i=1}^{\infty} p_i = 1$, then the pmf of zero-truncated counts is

$$\begin{aligned}
 f(x) &= \sum p_i(x_i \neq 0) \\
 &= \sum \frac{p_i}{1 - p(x=0)} \\
 &= \frac{\sum p_i}{1 - f_0(0)} \\
 &= \frac{f_0(x)}{1 - f_0(0)} \tag{1}
 \end{aligned}$$

2.1.2 Poisson-Weighted Exponential Distribution

Thus from equation (1), the pmf of zero-truncated distributions can be derived from the original pmf of models. Now consider the *weighted exponential distribution*(WE) with shape parameter $\alpha > 0$ and scale parameter $\beta > 0$, it is introduced in 2009 (Gupta, Kundu, pp. 624), the pdf of X is

$$g_0(\lambda; \alpha, \beta) = \frac{\alpha + 1}{\alpha} \beta e^{-\beta \lambda} (1 - e^{-\alpha \beta \lambda}); \lambda > 0. \quad (2)$$

The Poisson-Weighted Exponential distribution(P-WE) is proposed in 2014 (Zamani, Ismail, pp. 148). The P-WE distribution is a mixture distribution of conditional Poisson distribution and weighted exponential distribution. Suppose that random variable $x|\lambda$ is Poisson distributed, the pmf of x is

$$f(x|\lambda) = \frac{e^{-\lambda} \lambda^x}{x!}, x = 0, 1, 2, \dots$$

, and the random variable λ is distributed as WE with pdf:

$$g(\lambda) = \frac{\alpha + 1}{\alpha} \beta e^{-\beta \lambda} (1 - e^{-\alpha \beta \lambda}), \lambda, \alpha, \beta > 0,$$

the mixture pmf of the Poisson and WE distribution is then can be derived by:

$$\begin{aligned} f_0(x) &= \int f(x|\lambda) g(\lambda) d\lambda \\ &= \frac{(\alpha + 1)\beta}{\alpha} [(\beta + 1)^{-(x+1)} - (\alpha\beta + \beta + 1)^{-(x+1)}] \end{aligned} \quad (3)$$

for $x = 0, 1, 2, \dots$ and $\alpha, \beta > 0$.

In the research of Zamani, by applying this new model to the US NMES data sets and compare to other models such as negative binomial via log likelihood, AIC, and BIC, it is confident to say that P-WE distribution performs well and can be an alternative model to fit count data sets. The pmf of the P-WE distribution has a closed-form expression. Thus this new P-WE distribution could be a potentially better model for zero-truncated data sets.

Once the pmf of P-WE distribution and substitute it into equation (1), the pmf of ZTPWE distribution can be derived directly. There is another way to derive the pmf of ZTPWE distribution. It is a rather new class of distributions and is going to be enlightening for future research and study of its theoretical characteristics.

Consider a random size-biased Poisson distribution with parameter λ from the textbook (Gut, pp. 38-42), let the ZTPWE distribution as a combination of a conditional variable with a parameter that follows a new class of distribution. Thus the unconditional distribution can be derived under certain circumstances.

Probability Generating Function of the ZTPWE Distribution

The probability generating function (pgf) is the sequence p_0, p_1, \dots such that for function

$$G_X(s) = p_0 + p_1s + p_2s^2 + \dots,$$

for all values of s for which the right-hand side converges absolutely (Grimmett, 52).

Since $\sum p_i = 1$, it is not hard to prove that

$$G_X(0) = p(0) \text{ and } G_X(1) = 1$$

By the Law of the unconscious statistician (Grimmett, 31), If X is a discrete random variable and $g : \mathbb{R} \rightarrow \mathbb{R}$, then

$$\mathbb{E}(g(X)) = \sum_{x \in \text{Im} X} g(x) \mathbb{P}(X = x)$$

, for $|s| < 1$, indicates that

$$G_X(s) = \mathbb{E}(s^X).$$

The importance of the pgf can be shown by the following theorem: *Suppose X and Y have probability generating functions G_X and G_Y , respectively. Then*

$$G_X(s) = G_Y(s) \text{ for all } s$$

if and only if

$$\mathbb{P}(X = k) = \mathbb{P}(Y = k) \text{ for } k = 0, 1, 2, \dots$$

The above theorem actually indicates that if the pgf of two integer-valued random variables are the same, then their pmf's are the same.

2.1.3 Conditional Random Variable Sampling

Suppose a random variable X has parameter λ , then X is denoted as a *size-biased Poisson distribution*, if its pmf is:

$$f(x|\lambda) = e^{-\lambda} \frac{\lambda^{x-1}}{(x-1)!}, x = 1, 2, \dots$$

There are two ways to derive this distribution. The first method is to apply the size-biased transformation to the original Poisson distribution. Assume that the original pmf of Poisson distribution is $f_0(x), x = 0, 1, 2, \dots$, then the pmf of *size-biased* distribution is defined as

$$f(x) = f_0(x-1), x = 1, 2, \dots,$$

remind that the original pmf of Poisson distribution is:

$$f_0(x) = e^{-\lambda} \frac{\lambda^x}{x!}$$

simply substitute $f_0(x)$ into the size-biased Poisson transformation, obtain:

$$f(x) = f_0(x-1) = e^{-\lambda} \frac{\lambda^{x-1}}{(x-1)!}$$

For the second method, it is natural that the same result is obtained when apply the *length-biased* transformation. Similarly, assume that the original pmf of Poisson distribution is $f_0(x), x = 0, 1, 2, \dots$, then the pmf of *length-biased* distribution is defined as

$$f_X(x) = \frac{x f_0(x)}{E(X)}, x = 1, 2, \dots$$

again substitute $f_0(x)$ into the length-biased Poisson transformation, obtain:

$$f_X(x) = \frac{xe^{-\lambda} \frac{\lambda^x}{x!}}{\lambda} = e^{-\lambda} \frac{\lambda^{x-1}}{(x-1)!}$$

2.1.4 Cumulative Functions, Unimodal and Logconcave

A random variable X is a function mapping a sample space into the real number set, \mathbb{R} , and for each element in the domain, there exists one and only one $P(X = x)$ that indicates the probability of x . Similarly, there exists a unique $P(X \leq x)$ for all x . Such function $P(X \leq x)$ is defined as the *cumulative distribution function*, also denoted as $F_X(x)$, such that $0 \leq F_X(x) \leq 1$, and $\lim_{x \rightarrow \infty} F_X(x) = 1$. A discrete distribution is always a step function with limit amount of steps. For a specific step, $P(X = x) = f(x)$, where $f(x)$ is the probability mass function. Notice that $\lim_{x \rightarrow \infty} F_X(x) = \sum_i f(x_i) = 1$, for $i = 0, 1, 2, \dots$. The set $\{x_i\}$ is known as the discrete random variables. For a discrete distribuion, if $f(x)f(x+2)/f^2(x+1) > 1$, it is said to be *logconvex*, if $f(x)f(x+2)/f^2(x+1) < 1$, it is said to be *logconcave*. In this research, discrete data sets are observed, but ZTPWE distribution model is a continuous distribution. For a continuous pdf, if $df_X(x)/dx = 0$ and $d^2f_X(x)/dx^2 < 0$, the distritbution is a *mode*. The location of mode is at $X = x$ if

$$f_X(x - c_1 - 1) < f_X(x - c_1) \leq \dots \leq f_X(x)$$

and

$$f_X(x) \geq \dots \geq f_X(x + c_2) > f_X(x + c_2 + 1)$$

, where $0 \leq c_1, 0 \leq c_2$. In ordinary words, a distribution is a mode if it has peak values. If it has only one peak value, then it is *unimodal*; otherwise, it is *multimodal*. (Johnson, Kemp, pp. 43-51).

2.2 The Theory of Point Estimation

In the estimation section, the Method of Moments and the Method of Maximum Likelihood Estimation are used to test the performance of the ZTPWE model.

2.2.1 Moment and Moment Generating Function

In a physics way to describe the idea of the moment of a distribution, it is pretty much like an average "distance" of all data sets to a certain point that is interested in. By definition (Casella, p. 59),

For each integer n , the n th *moment* of X (or $F_X(x)$), u'_n , is

$$u'_n = EX^n.$$

The n th *central moment* of X , u_n , is

$$u_n = E(X - u)^n,$$

where $u = u' = EX$. In a physics way to understand this concept, assume $n = 1$, then the *first moment* of X is simply the average distance between data sets to 0.

The first *central moment* is the average distance between data sets to the position of

mean. If $n = 2$, then it is the average squared distance that is interested in, and so on, notice that the second *central moment* is usually called *variance*. The moment generating function indicates that it can generate moments of a distribution, though two different distributions can have the same moments(Shao, pp.32-33). However, the main purpose of mgf is to determine the properties such as mean and variance rather than derive the moments. By Definition(Casella, p. 62),

Let X be a random variable with cdf $F_X(x)$. The *moment generating function*(mgf) of X , denoted by $\psi(t)$ is

$$\psi(t) = Ee^{tX}$$

, if the expectation does not exist, then the mgf does not exist. The mgf of X can be derived by:

$$\psi(t) = \sum_x e^{tx} P(X = x) \quad \text{if } X \text{ is discrete,}$$

or

$$\psi(t) = \int_{-\infty}^{\infty} e^{tx} f_X(x) dx \quad \text{if } X \text{ is continuous.}$$

The mgf has the following property with $t = 0$:

$$EX^n = \psi^{(n)}(0)$$

and defined that:

$$\psi^{(n)}(0) = \left. \frac{d^n}{dt^n} \psi(t) \right|_{t=0}$$

That is, in other words, the n th moment is the n th derivative of mgf at $t = 0$. Simply look at the derivative:

$$\begin{aligned}\frac{d}{dt}\psi(t) &= \frac{d}{dt} \int e^{tx} f_X(x) dx \\ &= \int \left(\frac{d}{dt} e^{tx} \right) f_X(x) dx \\ &= \int (x e^{tx}) f_X(x) dx \\ &= EX e^{tX}\end{aligned}$$

If $t = 0$, then it can be obtained that:

$$\left. \frac{d}{dt} \psi(t) \right|_{t=0} = EX e^{tX} \Big|_{t=0} = EX.$$

and

$$\left. \frac{d^2}{dt^2} \psi(t) \right|_{t=0} = EX^2 e^{tX} \Big|_{t=0} = EX^2.$$

and so on.

2.2.2 Maximum Likelihood Estimation

The likelihood function is a compelling method to analyze data sets. It is defined as (Casella, p. 290, p. 310):

Definition 2.1 *Let $f(\mathbf{x}|\theta)$ denote the joint pdf or pmf of the sample $\mathbf{X} = (X_1, X_2, \dots, X_n)$, and let the pdf or pmf has parameter θ , then:*

$$L(\theta|\mathbf{x}) = f(\mathbf{x}|\theta)$$

where $\mathbf{X} = \mathbf{x}$ that is observed.

The maximum likelihood estimation is a powerful method to estimate the parameters of a probability distribution.

Definition 2.2 Recall that if $f(\mathbf{x}|\theta_1, \dots, \theta_k)$ denote the pdf or pmf of the sample $\mathbf{X} = (X_1, X_2, \dots, X_n)$, that from a population, the likelihood function is defined as:

$$L(\theta_1, \dots, \theta_k|\mathbf{x}) = \prod_{i=1}^n f(x_i|\theta_1, \dots, \theta_k)$$

then for each sample point \mathbf{x} , let $\hat{\theta}$ be a parameter value where $L(\theta|\mathbf{x})$ gains its maximum as a function of θ , with \mathbf{x} remain fixed. A maximum likelihood estimator(MLE) of the parameter θ based on a sample \mathbf{X} is $\hat{\theta}(\mathbf{X})$.

MLE provides an inspection of the most likely value of a parameter could be given the count data set that has been observed. If the likelihood function is differentiable over θ_i , then one of the potential values for the MLE is:

$$\frac{\partial}{\partial \theta_i} L(\theta|\mathbf{x}) = 0, \quad i = 1, \dots, k.$$

The solution x s to the above equation are possible candidates for the MLE since the derivative being 0 is a sign for maximum, but it is not a sufficient condition. In many cases, it is much easier to apply the natural logarithm of $L(\theta|x)$, say $\log L(\theta|x)$ (known as the log-likelihood) instead.

2.2.3 Mean Squared Error and Bias

Definition 2.3 *The mean squared error(MSE) of an estimator W of a parameter θ is the function of θ defined as $E_{\theta}(W - \theta)^2$. (Casella, p. 330)*

By definition, we can tell that the MSE is essentially the squared "distance" or difference between the estimator W and the parameter θ . In fact, there are many ways to describe this "distance", for instance, the absolute value of $W - \theta$ would be a good measure of the performance of the estimator, the mean absolute error, $E_{\theta}|W - \theta|$. The reason that we use this MSE is that it is connected to another concept about the "distance" between the W and θ . The definition of the MSE is:

$$\begin{aligned} E_{\theta}(W - \theta)^2 &= E_{\theta}(W^2) + E_{\theta}(\theta^2) - 2E_{\theta}(W\theta) \\ &= E_{\theta}(W^2) - (E_{\theta}(W))^2 + (E_{\theta}(W))^2 + E_{\theta}(\theta^2) - 2E_{\theta}(W\theta) \\ &= \text{Var}_{\theta}(W) + (E_{\theta}(W) - \theta)^2. \end{aligned}$$

Thus, we define the bias of an estimator as follow:

Definition 2.4 *The bias of a point estimator W of a parameter θ is the difference between the expected value of W and θ .*

In mathematical words, $\text{Bias}_{\theta}W = E_{\theta}W - \theta$. If the bias is 0, then we say that the estimator is unbiased, $E_{\theta}W = \theta$, for all θ . Thus, we can tell that a good MSE should equal to variance, which means there is 0 bias.

2.2.4 AIC, AD Test and p-value

The AIC is defined by (Konishi and Kitagawa, pp. 60-61):

$$\text{AIC} = -2(\text{maximum log-likelihood}) + 2(\text{number of free parameters})$$

The number of free parameters in a distribution model is related to the dimensions of a parameter θ contained in the $f(x|\theta)$.

The AIC basically describes that the number of free parameters in the model plays an important role as a new "bias". Thus a good model should have a relatively small AIC value.

The Anderson-Darling(AD) statistic describes how close the points are to the fitted line estimated in a probability graphic (Jantschi, Bolboaca, p. 2). The maximum p -value is based on the discrete AD test, and the large p -value indicates that there is strong evidence to say the data and model are related (Anderson, Darling, p. 193).

2.3 Review of the Related Literature

In 1952, David and Johnson noticed that in the Poisson model, the zero group remains unobserved. They found that the Poisson model does not perform well when dealing with a certain number of 0 counts. Thus a truncated distribution should be considered when dealing with such data sets.

In 2007, Gupta and Kundu introduced a new version of weighted exponential distribution. A new parameter is introduced to gain this new distribution. Two different methods are applied to obtain the pdf of random variables. The first way is to take pdf of weighted exponential distribution as a special case of Jones' model (Jones, pp. 41-43). The other way is to take the skew-normal distribution (Arnold, Beaver, pp. 23-25) as a hidden truncation model. The weighted exponential distribution is similar to some other distributions. By evaluating the MLE, it seems the result does not perform better even with a large sample size. Fitting weighted exponential model to real data sets, the survival times of guinea pigs with different amounts of tubercle bacilli injection (Bjerkedal, 1960), compare with Weibull, gamma, and generalized exponential distributions, they seem perform similarly. The second data sets are 450 students' performance in Joint Entrance Examination from the Indian Institute of Technology Kanpur. Interestingly, this time the weighted exponential distribution performs better than the Weibull and all other models.

In 2014, Zamani and Noriszura proposed the research about P-WE distribution, it is a two-parameter model. Their method is to take $X|\lambda$ as a conditional Poisson distribution and let λ be distributed as weighted exponential distribution. Thus the combination can be obtained, namely P-WE distribution. Furthermore, the properties of P-WE distribution are derived. The shape of pmf of P-WE implies it can be used as an alternative for count data sets. The model is fitted with two real data sets.

The first one is Belgium insurance count data in 1993(Denuit, 1997). The second data set is US national Medical Expenditure Survey 1987/88(NMES) data from Deb and Trivedi(1997). The result is compared with Poisson, generalized Poisson and negative binomial distributions, investigate the criteria information, log-likelihood, AIC that is generated in R language. It seems that the P-WE performs better than other models.

The related pmf's are:

ZTP:

$$f(x; \theta) = \frac{(e^\theta - 1)^{-1} \theta^x}{x!}$$

for $x = 1, 2, 3, \dots$.

ZTPL:

$$f(x; \theta) = \frac{\theta^3}{\theta^2 + 3\theta + 1} \frac{x + \theta + 2}{(\theta + 1)^x}$$

for $x = 1, 2, 3, \dots, \theta > 0$.

ZTTPL:

$$f(x; \theta, \alpha) = \frac{\theta^2}{\theta^2 + 2\theta\alpha + \theta + \alpha} \frac{\alpha x + (\theta + \alpha + 1)}{(\theta + 1)^x}$$

for $x = 1, 2, 3, \dots, \theta > 0, \theta^2 + 2\theta\alpha + \theta + \alpha > 0$.

Chapter 3

THEORETICAL PART

3.1 ZTPWE model

In the last chapter, the two different ways to obtain the ZTPWE distribution is discussed, and the first way is substitute equation (3) into (2), obtain:

$$\begin{aligned} f(x) &= \frac{f_0(x)}{1 - f_0(0)} \\ &= \frac{\frac{(\alpha+1)\beta}{\alpha} [(\beta+1)^{-(x+1)} - (\alpha\beta + \beta + 1)^{-(x+1)}]}{1 - \frac{(\alpha+1)\beta}{\alpha} [(\beta+1)^{-1} - (\alpha\beta + \beta + 1)^{-1}]} \\ &= \frac{(\alpha+1)\beta [(\beta+1)^{-(x+1)} - (\alpha\beta + \beta + 1)^{-(x+1)}]}{\alpha - (\alpha+1)\beta [(\beta+1)^{-1} - (\alpha\beta + \beta + 1)^{-1}]} \\ &= \frac{(\alpha+1)\beta [(\beta+1)^{-(x+1)} - (\alpha\beta + \beta + 1)^{-(x+1)}] ((\beta+1)(\alpha\beta + \beta + 1))}{\alpha(\beta+1)(\alpha\beta + \beta + 1) - \alpha(\alpha+1)\beta^2} \\ &= \frac{(\alpha+1)\beta [(\beta+1)^{-(x+1)} - (\alpha\beta + \beta + 1)^{-(x+1)}] ((\beta+1)(\alpha\beta + \beta + 1))}{2\alpha\beta + \alpha^2\beta + \alpha} \\ &= \frac{(\alpha+1)\beta}{\alpha((\alpha+2)\beta + 1)} (\beta+1)^{-x} (\alpha\beta + \beta + 1)^{-x} ((\alpha\beta + \beta + 1)^{x+1} - (\beta+1)^{x+1}) \end{aligned}$$

for $x = 1, 2, 3, \dots$ and $\alpha, \beta > 0$.

The other way to gain the pmf of ZTPWE distribution is rather new. Now introduce a new class of distributions that have not been discussed in the literature yet. Assume that a random variable follows a *partly-exponential distribution* with three parameters $0 \leq \gamma \leq 1$, $\delta > 0$, and $0 < \tau < \delta(1 + \gamma)$, if its density function is:

$$h(t) = \begin{cases} c(1 - \gamma e^{-t/\delta})e^{-t/\tau} & \text{if } t \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

where the normalizing constant $c = \frac{\delta + \tau}{\tau[\delta(1 - \gamma) + \tau]}$. The fact that the above model is a valid density function is a particular case $\theta = 1$ of the following result.

Lemma 3.1 *For positive δ, τ , and θ , and any real γ :*

$$\int_0^\infty t^{\theta-1} (1 - \gamma e^{-t/\delta}) e^{-t/\tau} dt = \Gamma(\theta) \tau^\theta \left[1 - \gamma \left(\frac{\delta}{\delta + \tau} \right)^\theta \right].$$

If $\theta = 1$, rewrite the above equation to:

$$\begin{aligned} \int_0^\infty t^{1-1} (1 - \gamma e^{-t/\delta}) e^{-t/\tau} dt &= \Gamma(1) \tau^1 \left[1 - \gamma \left(\frac{\delta}{\delta + \tau} \right)^1 \right] \\ \int_0^\infty (1 - \gamma e^{-t/\delta}) e^{-t/\tau} dt &= \tau \left[1 - \gamma \left(\frac{\delta}{\delta + \tau} \right) \right] \\ \int_0^\infty \frac{\delta + \tau}{\tau[\delta(1 - \gamma) + \tau]} (1 - \gamma e^{-t/\delta}) e^{-t/\tau} dt &= 1. \end{aligned}$$

where the normalizing constant $c = \frac{\delta + \tau}{\tau[\delta(1 - \gamma) + \tau]}$, this result indicates that $h(t)$ has intergral equals to 1, thus it is a valid density function.

Proof. The statement is a simple corollary of the well known formula:

$$\int_0^\infty t^{\theta-1} e^{-t/a} dt = \Gamma(\theta) a^\theta$$

been substituted twice. For the left hand side of equation (4), it can be rewritten as:

$$\begin{aligned} \int_0^\infty t^{\theta-1} (1 - \gamma e^{-t/\delta}) e^{-t/\tau} dt &= \int_0^\infty t^{\theta-1} e^{-t/\tau} dt - \int_0^\infty t^{\theta-1} \gamma e^{-t(\frac{1}{\delta} + \frac{1}{\tau})} dt \\ &= \int_0^\infty t^{\theta-1} e^{-t/\tau} dt - \int_0^\infty t^{\theta-1} \gamma e^{-\frac{t}{\frac{\delta\tau}{\delta+\tau}}} dt \\ &= \Gamma(\theta) \tau^\theta - \gamma \Gamma(\theta) \left(\frac{\delta\tau}{\delta+\tau} \right)^\theta \\ &= \Gamma(\theta) \tau^\theta \left[1 - \gamma \left(\frac{\delta}{\delta+\tau} \right)^\theta \right]. \end{aligned}$$

Now suppose that there is a random variable X follows the conditional distribution with parameter $\lambda > 0$ size-biased Poisson distribution and that the parameter λ itself follows the partly-exponential distribution with parameter $\gamma = \frac{\beta+1}{\alpha\beta+\beta+1}$, $\delta = \frac{1}{\alpha\beta}$, and $\tau = \frac{1}{\beta}$; thus the pdf of λ is

$$\begin{aligned} h(\lambda) &= \frac{\delta + \tau}{\tau[\delta(1 - \gamma) + \tau]} (1 - \gamma e^{-t/\delta}) e^{-t/\tau} \\ &= \frac{\frac{1}{\alpha\beta} + \frac{1}{\beta}}{\frac{1}{\beta} \left[\frac{1}{\alpha\beta} \left(1 - \frac{\beta+1}{\alpha\beta+\beta+1} \right) + \frac{1}{\beta} \right]} \left(1 - \frac{\beta+1}{\alpha\beta+\beta+1} e^{-t/\frac{1}{\alpha\beta}} \right) e^{-t/\frac{1}{\beta}} \\ &= \frac{\alpha\beta + \beta}{\left(1 - \frac{\beta+1}{\alpha\beta+\beta+1} \right) + \alpha} \left(1 - \frac{\beta+1}{\alpha\beta+\beta+1} e^{-t/\frac{1}{\alpha\beta}} \right) e^{-t/\frac{1}{\beta}} \\ &= \frac{(\alpha+1)\beta(\alpha\beta + \beta + 1)}{(\alpha\beta + \beta + 1) - (\beta + 1) + \alpha(\alpha\beta + \beta + 1)} \left(1 - \frac{\beta+1}{\alpha\beta+\beta+1} e^{-t/\frac{1}{\alpha\beta}} \right) e^{-t/\frac{1}{\beta}} \\ &= \frac{(\alpha+1)\beta}{\alpha((\alpha+2)\beta + 1)} \left[(\alpha\beta + \beta + 1) - (\beta + 1)e^{-\alpha\beta\lambda} \right] e^{-\beta\lambda}, \lambda > 0. \end{aligned}$$

Now it is clear that a unconditional distribution $f(X)$ with parameter λ can be obtained by:

$$\begin{aligned}
f(x) &= \int_0^\infty f(x|\lambda)h(\lambda)d\lambda \\
&= \int_0^\infty e^{-\lambda} \frac{\lambda^{x-1}}{(x-1)!} \frac{(\alpha+1)\beta}{\alpha((\alpha+2)\beta+1)} [(\alpha\beta+\beta+1) - (\beta+1)e^{-\alpha\beta\lambda}] e^{-\beta\lambda} d\lambda \\
&= \frac{(\alpha+1)\beta}{\alpha((\alpha+2)\beta+1)} \int_0^\infty e^{-\lambda} \frac{\lambda^{x-1}}{(x-1)!} [(\alpha\beta+\beta+1) - (\beta+1)e^{-\alpha\beta\lambda}] e^{-\beta\lambda} d\lambda \\
&= \frac{(\alpha+1)\beta(\alpha\beta+\beta+1)}{\alpha((\alpha+2)\beta+1)\Gamma(x)} \int_0^\infty \lambda^{x-1} \left[1 - \frac{\beta+1}{\alpha\beta+\beta+1} e^{-\alpha\beta\lambda} \right] e^{-\beta\lambda} d\lambda \\
&\text{by Lemma 3.1, let } t = \lambda, \theta = x, \gamma = \frac{\beta+1}{\alpha\beta+\beta+1}, \delta = \frac{1}{\alpha\beta}, \text{ and } \tau = \frac{1}{\beta+1} \\
&= \frac{(\alpha+1)\beta(\alpha\beta+\beta+1)}{\alpha((\alpha+2)\beta+1)\Gamma(x)} \left\{ \Gamma(x) \left(\frac{1}{\beta+1} \right)^x \left[1 - \frac{\beta+1}{\alpha\beta+\beta+1} \left(\frac{\frac{1}{\alpha\beta}}{\frac{1}{\alpha\beta} + \frac{1}{\beta+1}} \right)^x \right] \right\} \\
&= \frac{(\alpha+1)\beta(\alpha\beta+\beta+1)}{\alpha((\alpha+2)\beta+1)\Gamma(x)} \left\{ \Gamma(x) \left(\frac{1}{\beta+1} \right)^x \left[1 - \left(\frac{\beta+1}{\alpha\beta+\beta+1} \right)^{x+1} \right] \right\} \\
&= \frac{(\alpha+1)\beta}{\alpha((\alpha+2)\beta+1)} \left[\frac{\alpha\beta+\beta+1}{(\beta+1)^x} - \frac{\beta+1}{(\alpha\beta+\beta+1)^x} \right] \\
&= \frac{(\alpha+1)\beta}{\alpha((\alpha+2)\beta+1)} \left[\frac{(\alpha\beta+\beta+1)^{x+1} - (\beta+1)^{x+1}}{(\beta+1)^x(\alpha\beta+\beta+1)^x} \right],
\end{aligned}$$

for $\alpha, \beta > 0$. And check with the pmf of ZTPWE distribution by the first method, the two pmf are identical, which is as expected.

3.1.1 Plots of pmf's

One of the most intuitive ways to understand a pmf function is to plot it, here use R language [11] to plot pmf with different set of parameters α and β .

The pmf plots with different parameter pairs are shown above. They all skewed

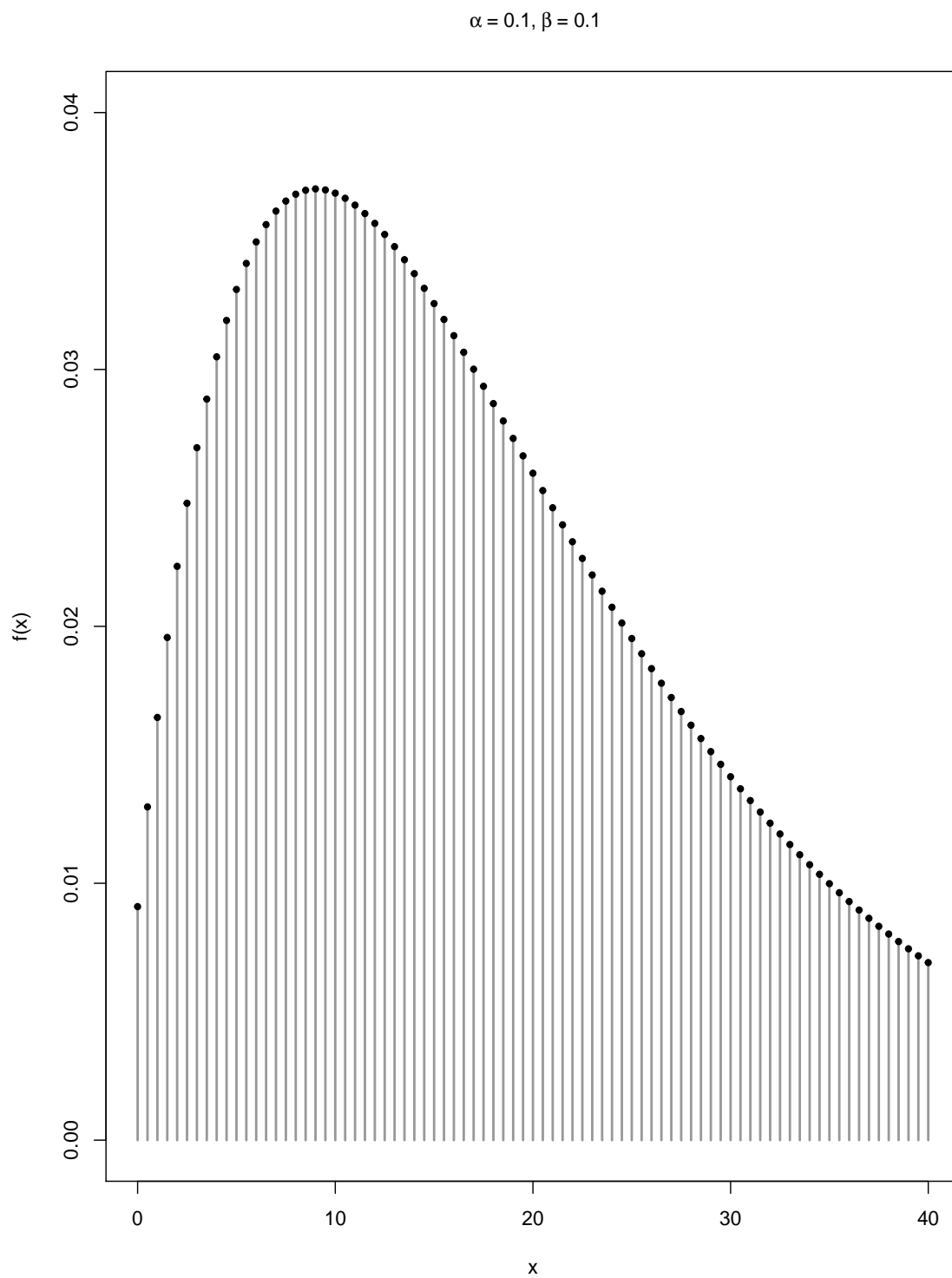


Figure 3.1: Pmf plots of the ZTPWE distribution with parameter α and β

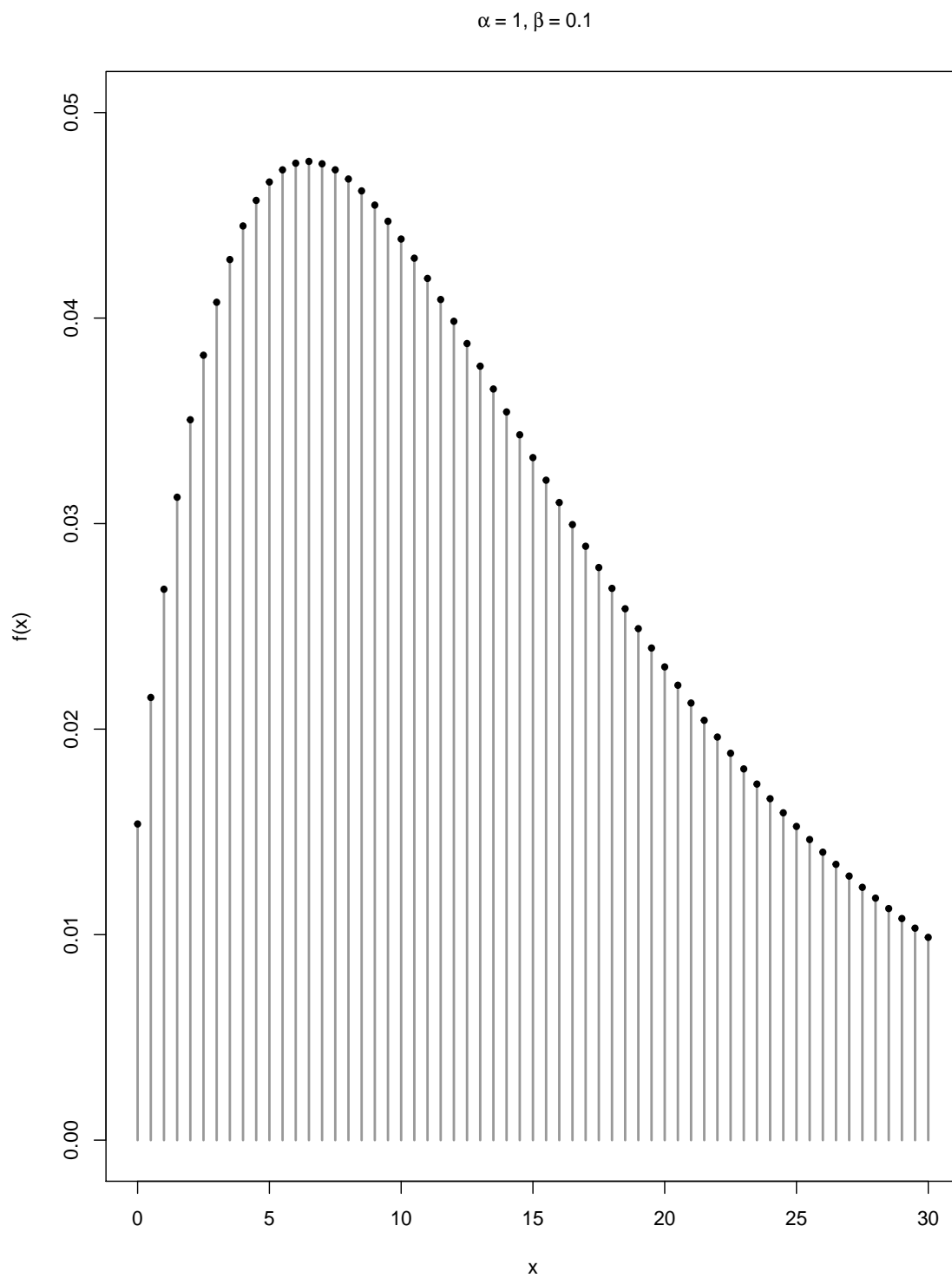


Figure 3.2: Pmf plots of the ZTPWE distribution with parameter α and β

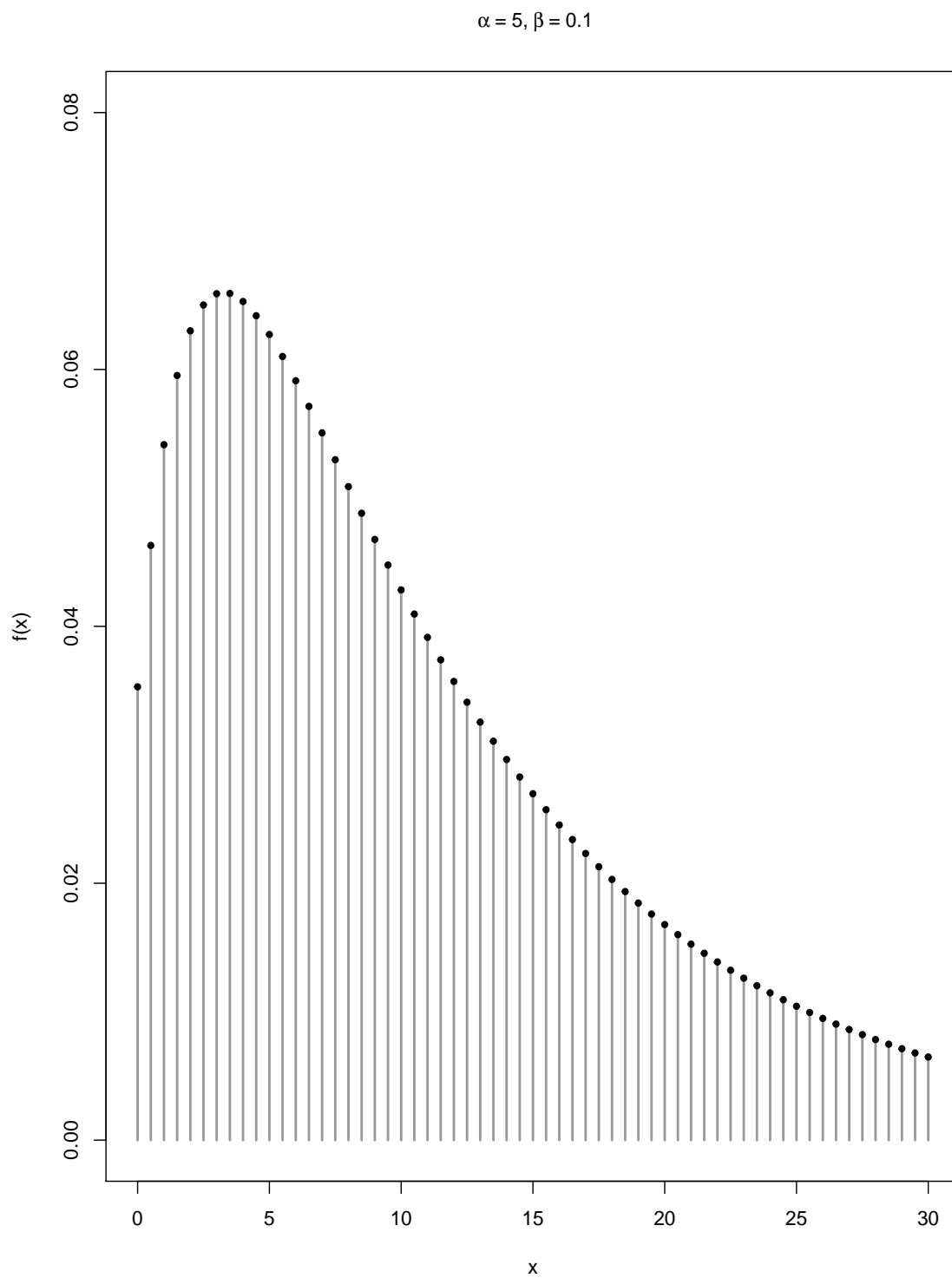


Figure 3.3: Pmf plots of the ZTPWE distribution with parameter α and β

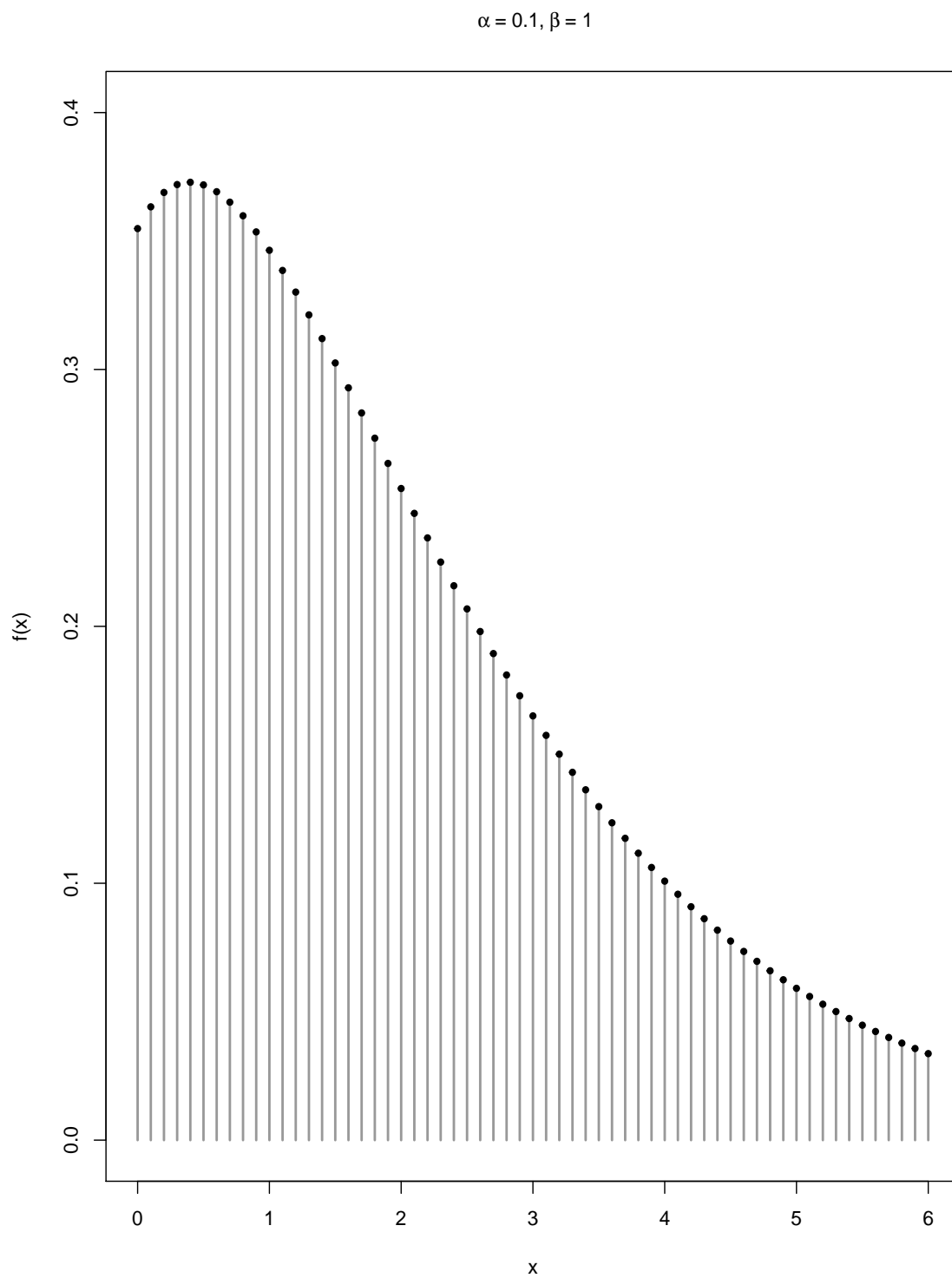


Figure 3.4: Pmf plots of the ZTPWE distribution with parameter α and β

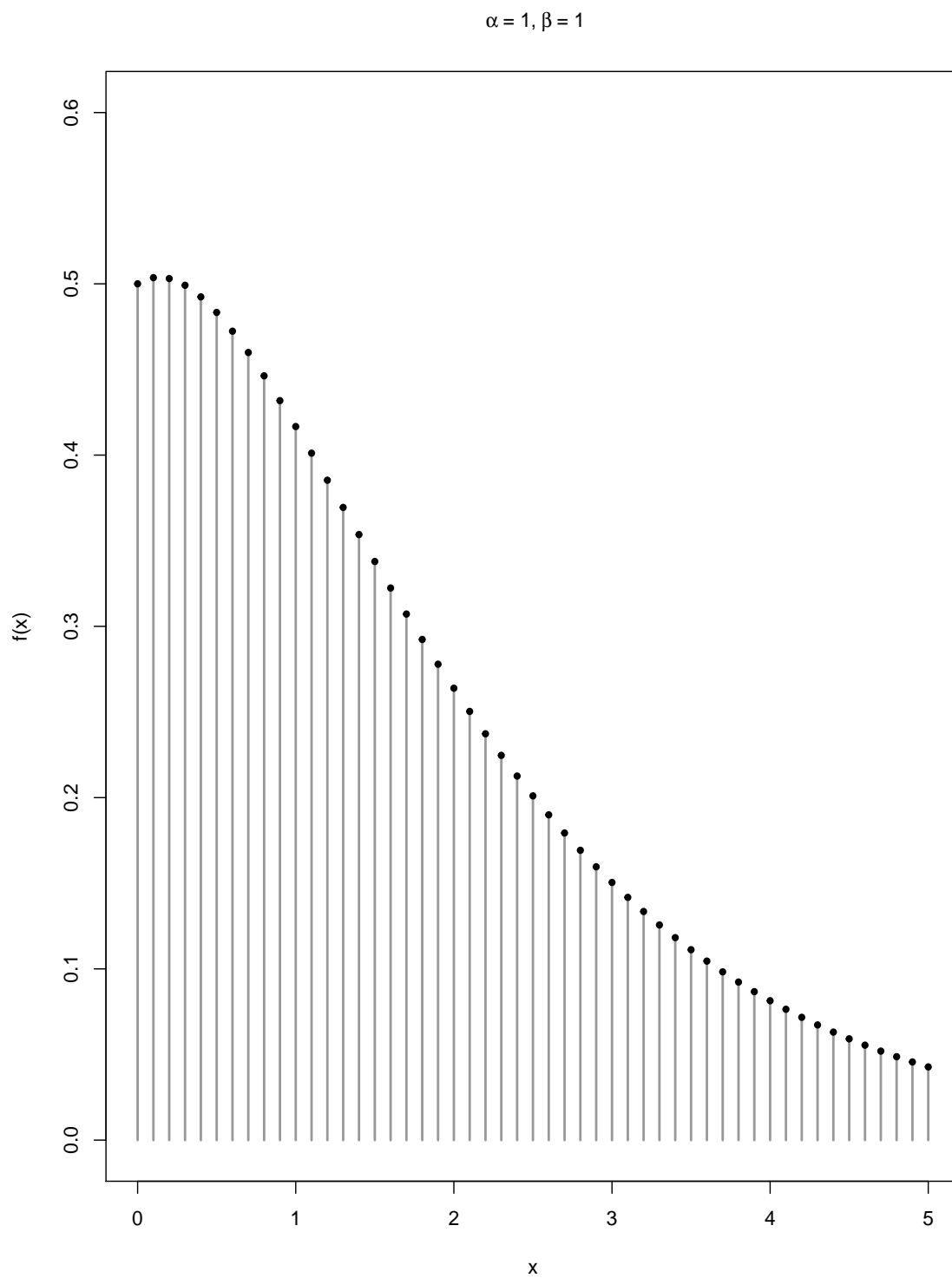


Figure 3.5: Pmf plots of the ZTPWE distribution with parameter α and β

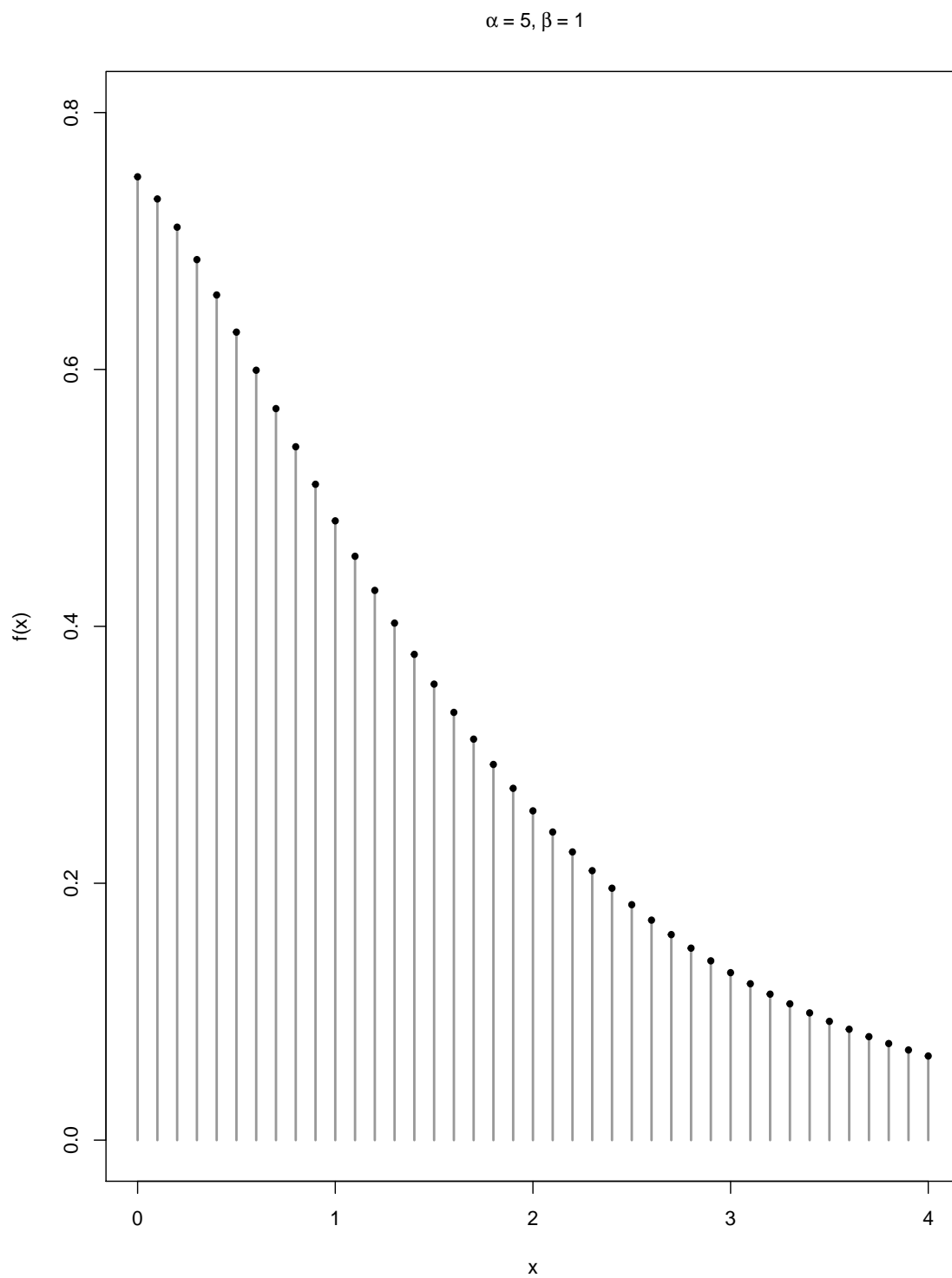


Figure 3.6: Pmf plots of the ZTPWE distribution with parameter α and β

right, seem like pulled out to the right. Each of the plots has a peak value or mode, the location of the mode is decreasing as α and β increase. Meanwhile, the magnitude of the mode increases as α and β increase. The plots clearly indicate that the ZTPWE distribution has only one mode, so it is unimodal (Johnson, Kemp and Kotz, p. 51).

Take a look at the ratio:

$$\begin{aligned} \frac{f(x+1)}{f(x)} &= \frac{\frac{(\alpha+1)\beta}{\alpha((\alpha+2)\beta+1)} \frac{(\alpha\beta+\beta+1)^{x+2} - (\beta+1)^{x+2}}{(\beta+1)^{x+1}(\alpha\beta+\beta+1)^{x+1}}}{\frac{(\alpha+1)\beta}{\alpha((\alpha+2)\beta+1)} \frac{(\alpha\beta+\beta+1)^{x+1} - (\beta+1)^{x+1}}{(\beta+1)^x(\alpha\beta+\beta+1)^x}} \\ &= \frac{1}{(\beta+1)(\alpha\beta+\beta+1)} \left[\frac{(\alpha\beta+\beta+1)^{x+2} - (\beta+1)^{x+2}}{(\alpha\beta+\beta+1)^{x+1} - (\beta+1)^{x+1}} \right] \end{aligned}$$

and by fact that

$$\begin{aligned} \frac{f(x+2)f(x)}{f^2(x+1)} &= \frac{f(x+2)}{f(x+1)} \frac{f(x)}{f(x+1)} \\ &= \frac{\frac{1}{(\beta+1)(\alpha\beta+\beta+1)} \frac{(\alpha\beta+\beta+1)^{x+3} - (\beta+1)^{x+3}}{(\beta+1)^{x+2}(\alpha\beta+\beta+1)^{x+2}}}{\frac{1}{(\beta+1)(\alpha\beta+\beta+1)} \frac{(\alpha\beta+\beta+1)^{x+2} - (\beta+1)^{x+2}}{(\beta+1)^{x+1}(\alpha\beta+\beta+1)^{x+1}}} \frac{\frac{1}{(\beta+1)(\alpha\beta+\beta+1)} \frac{(\alpha\beta+\beta+1)^{x+1} - (\beta+1)^{x+1}}{(\beta+1)^x(\alpha\beta+\beta+1)^x}}{\frac{1}{(\beta+1)(\alpha\beta+\beta+1)} \frac{(\alpha\beta+\beta+1)^{x+2} - (\beta+1)^{x+2}}{(\beta+1)^{x+1}(\alpha\beta+\beta+1)^{x+1}}} \\ &= \frac{(\alpha\beta+\beta+1)^{x+3} - (\beta+1)^{x+3}}{(\alpha\beta+\beta+1)^{x+2} - (\beta+1)^{x+2}} \frac{(\alpha\beta+\beta+1)^{x+1} - (\beta+1)^{x+1}}{(\alpha\beta+\beta+1)^{x+2} - (\beta+1)^{x+2}} \end{aligned}$$

let $A = \alpha\beta + \beta + 1, B = \beta + 1, A, B > 0, A \neq B$,

$$\begin{aligned} &= \frac{A^{2x+4} + B^{2x+4} - A^{x+3}B^{x+1} - A^{x+1}B^{x+3}}{A^{2x+4} + B^{2x+4} - 2A^{x+2}B^{x+2}} \\ &= \frac{A^{2x+4} + B^{2x+4} - A^{x+1}B^{x+1}(A^2 + B^2)}{A^{2x+4} + B^{2x+4} - A^{x+1}B^{x+1}(2AB)} \end{aligned}$$

Now compare $A^2 + B^2$ and $2AB$, it is known that $(A - B)^2 > 0$ with $A \neq B$, thus

$A^2 + B^2 - 2AB > 0 \Rightarrow A^2 + B^2 > 2AB \Rightarrow \frac{f(x+2)f(x)}{f^2(x+1)} < 1$, which implies the ZTPWE distribution is logconcave (John, Kemp and Kotz p.43).

3.1.2 Cumulative Distribution Function

By definition, the cumulative distribution function(cdf) can be derived by $F(x) = \sum f(x), x = 1, 2, 3, \dots$, the partial sums of geometric sequence terms are applied twice,

$$\begin{aligned}
F(x) &= \sum_x \frac{(\alpha + 1)\beta}{\alpha((\alpha + 2)\beta + 1)} \left[\frac{(\alpha\beta + \beta + 1)^{x+1} - (\beta + 1)^{x+1}}{(\beta + 1)^x(\alpha\beta + \beta + 1)^x} \right] \\
&= \frac{(\alpha + 1)\beta}{\alpha((\alpha + 2)\beta + 1)} \left[\sum_x \frac{(\alpha\beta + \beta + 1)^{x+1}}{(\beta + 1)^x(\alpha\beta + \beta + 1)^x} - \sum_x \frac{(\beta + 1)^{x+1}}{(\beta + 1)^x(\alpha\beta + \beta + 1)^x} \right] \\
&= \frac{(\alpha + 1)\beta}{\alpha((\alpha + 2)\beta + 1)} \left[\sum_x \frac{\alpha\beta + \beta + 1}{(\beta + 1)^x} - \sum_x \frac{\beta + 1}{(\alpha\beta + \beta + 1)^x} \right] \\
&= \frac{((\alpha + 1)\beta)(\alpha\beta + \beta + 1)}{\alpha((\alpha + 2)\beta + 1)} \sum_x \frac{1}{(\beta + 1)^x} - \frac{((\alpha + 1)\beta)(\beta + 1)}{\alpha((\alpha + 2)\beta + 1)} \sum_x \frac{1}{(\alpha\beta + \beta + 1)^x}
\end{aligned}$$

The above equation is a combination of two sums of geometric sequence. The general term of a geometric sequence can be re-written as $a_1 r^{n-1}$, where a_1 is

the first term when $x = 1$. The summation $S_n = \frac{a_1 r^n}{1 - r}$.

The first summation has ratio $r = \frac{1}{\beta + 1}$, the second summation has ratio $r = \frac{1}{\alpha\beta + \beta + 1}$.

The above equation can be re-written as:

$$\begin{aligned}
&= \frac{\frac{(\alpha+1)(\alpha\beta+\beta+1)\beta}{\alpha((\alpha+2)\beta+1)(\beta+1)}}{\frac{\beta}{\beta+1}} \left(1 - \left(\frac{1}{\beta+1}\right)^x\right) - \frac{\frac{(\alpha+1)(\beta+1)\beta}{\alpha((\alpha+2)\beta+1)(\alpha\beta+\beta+1)}}{\frac{(\alpha+1)\beta}{\alpha\beta+\beta+1}} \left(1 - \left(\frac{1}{\alpha\beta+\beta+1}\right)^x\right) \\
&= \frac{(\alpha+1)(\alpha\beta+\beta+1)}{\alpha((\alpha+2)\beta+1)} \left[1 - \left(\frac{1}{\beta+1}\right)^x\right] - \frac{\beta+1}{\alpha((\alpha+2)\beta+1)} \left[1 - \left(\frac{1}{\alpha\beta+\beta+1}\right)^x\right] \\
&= \frac{\alpha((\alpha+2)\beta+\alpha+(\beta+1)(\alpha\beta+\beta+1)^{-x} - (\alpha+1)(\alpha\beta+\beta+1)(\beta+1)^{-x}}{\alpha((\alpha+2)\beta+1)}
\end{aligned}$$

$x = 1, 2, 3, \dots$

3.2 Theoretical Probabilistic Properties of the ZTPWE Distribution

3.2.1 The Probability Generating Function of ZTPWE Distribution

Theorem 3.1 *If $X \sim ZTPWE(\alpha, \beta)$, then the probability generating function of X is*

$$G(s) = \frac{(\alpha+1)\beta}{\alpha((\alpha+2)\beta+1)} \left[\frac{s(\alpha\beta+\beta+1)}{\beta+1-s} - \frac{s(\beta+1)}{\alpha\beta+\beta+1-s} \right],$$

where $|s| < \beta+1$.

Proof. By definition, the pgf of X is given by:

$$\begin{aligned}
G(s) &= E(s^x) = \sum_{x=1}^{\infty} s^x f(x) \\
&= \sum_{x=1}^{\infty} s^x \frac{(\alpha+1)\beta}{\alpha((\alpha+2)\beta+1)} \left[\frac{(\alpha\beta+\beta+1)^{x+1} - (\beta+1)^{x+1}}{(\beta+1)^x(\alpha\beta+\beta+1)^x} \right]
\end{aligned}$$

$$= \frac{(\alpha + 1)\beta}{\alpha((\alpha + 2)\beta + 1)} \left\{ (\alpha\beta + \beta + 1) \sum_{x=1}^{\infty} \left(\frac{s}{\beta + 1}\right)^x - (\beta + 1) \sum_{x=1}^{\infty} \left(\frac{s}{\alpha\beta + \beta + 1}\right)^x \right\}$$

again calculate two sums of geometric series, $S_{\infty} = \frac{a_1}{1 - r}$, where a_1 is the

first term and r is the geometric ratio.

$$\begin{aligned} &= \frac{(\alpha + 1)\beta}{\alpha((\alpha + 2)\beta + 1)} \left[\frac{\frac{((\alpha\beta + \beta + 1)s)}{\beta + 1}}{1 - \frac{s}{\beta + 1}} - \frac{\frac{(\beta + 1)s}{\alpha\beta + \beta + 1}}{1 - \frac{s}{\alpha\beta + \beta + 1}} \right] \\ &= \frac{(\alpha + 1)\beta}{\alpha((\alpha + 2)\beta + 1)} \left[\frac{(\alpha\beta + \beta + 1)s}{\beta + 1 - s} - \frac{(\beta + 1)s}{\alpha\beta + \beta + 1 - s} \right] \end{aligned}$$

The moment generating function and the characteristic function of X can be obtained by the similar mathematic method as the pgf.

3.2.2 The Moment Generating Function of ZTPWE Distribution

Theorem 3.2 *If $X \sim ZTPWE(\alpha, \beta)$, then the moment generating function of X is*

$$\psi(t) = \frac{(\alpha + 1)\beta}{\alpha((\alpha + 2)\beta + 1)} \left[\frac{e^t(\alpha\beta + \beta + 1)}{\beta + 1 - e^t} - \frac{e^t(\beta + 1)}{\alpha\beta + \beta + 1 - e^t} \right],$$

where $t < \log(\beta + 1)$.

3.2.3 The characteristic Function of ZTPWE Distribution

Theorem 3.3 *If $X \sim ZTPWE(\alpha, \beta)$, then the characteristic function of X is*

$$\varphi(t) = \frac{(\alpha + 1)\beta}{\alpha((\alpha + 2)\beta + 1)} \left[\frac{e^{it}(\alpha\beta + \beta + 1)}{\beta + 1 - e^{it}} - \frac{e^{it}(\beta + 1)}{\alpha\beta + \beta + 1 - e^{it}} \right].$$

3.2.4 Methods of Moments

The methods of moments implies that the k th derivative of the mgf with $t = 0$ equals to $E(x^k)$. Thus the mean and variance of ZTPWE distribution can be derived by:

$$\begin{aligned} E(x) &= \left(\frac{d\psi(t)}{dt} \right)_{t=0} \\ &= \frac{d}{dt} \left(\frac{(\alpha + 1)\beta}{\alpha((\alpha + 2)\beta + 1)} \left[\frac{e^t(\alpha\beta + \beta + 1)}{\beta + 1 - e^t} - \frac{e^t(\beta + 1)}{\alpha\beta + \beta + 1 - e^t} \right] \right)_{t=0} \\ &= \frac{(\alpha + 1)\beta}{\alpha((\alpha + 2)\beta + 1)} \left(\frac{(\alpha\beta + \beta + 1)e^t(\beta + 1 - e^t) + e^{2t}(\alpha\beta + \beta + 1)}{(\beta + 1 - e^t)^2} \right. \\ &\quad \left. - \frac{(\beta + 1)e^t(\alpha\beta + \beta + 1 - e^t) + e^{2t}(\beta + 1)}{(\alpha\beta + \beta + 1 - e^t)^2} \right)_{t=0} \\ &= \frac{(\alpha + 1)\beta(\beta + 1)(\alpha\beta + \beta + 1)}{\alpha((\alpha + 2)\beta + 1)} \left(\frac{e^t}{(\beta - e^t + 1)^2} - \frac{e^t}{(\alpha\beta + \beta + 1 - e^t)^2} \right)_{t=0} \\ &= \frac{(\alpha + 1)\beta(\beta + 1)(\alpha\beta + \beta + 1)}{\alpha((\alpha + 2)\beta + 1)} \left(\frac{1}{\beta^2} - \frac{1}{(\alpha\beta + \beta)^2} \right) \\ &= \frac{(\alpha\beta + \beta + 1)(\beta + 1)(\alpha + 2)}{((\alpha + 2)\beta + 1)\beta(\alpha + 1)} \end{aligned}$$

In order to obtain the variance, consider the definition of variance, it is defined as

$$\text{Var}(X) = E(X - E(X))^2 = E(X^2 - 2XE(X) + E^2(X)) = E(X^2) - 2E(XE(X)) +$$

$$E^2(X) = E(X^2) - 2E(X)E(X) + E^2(X) = E(X^2) - 2E^2(X) + E^2(X) = E(X^2) - E^2(X).$$

Thus it is necessary to derive $E(X^2)$. Again, by methods of moments, the second derivative of mgf with respect to t at $t = 0$ equals to $E(x^2)$, thus:

$$\begin{aligned} E^2(x) &= \left(\frac{d^2\psi(t)}{dt^2} \right)_{t=0} \\ &= \frac{d^2}{dt^2} \left(\frac{(\alpha+1)\beta}{\alpha((\alpha+2)\beta+1)} \left[\frac{e^t(\alpha\beta+\beta+1)}{\beta+1-e^t} - \frac{e^t(\beta+1)}{\alpha\beta+\beta+1-e^t} \right] \right)_{t=0} \\ &= \frac{(\alpha+1)\beta}{\alpha((\alpha+2)\beta+1)} \frac{d^2}{dt^2} \left(\frac{e^t(\alpha\beta+\beta+1)}{\beta+1-e^t} - \frac{e^t(\beta+1)}{\alpha\beta+\beta+1-e^t} \right)_{t=0} \\ &= \frac{(\alpha+1)\beta(\beta+1)(\alpha\beta+\beta+1)}{\alpha((\alpha+2)\beta+1)} \frac{d}{dt} \left(\frac{e^t}{(\beta-e^t+1)^2} - \frac{e^t}{(\alpha\beta+\beta+1-e^t)^2} \right)_{t=0} \\ &= \frac{(\alpha+1)\beta(\beta+1)(\alpha\beta+\beta+1)}{\alpha((\alpha+2)\beta+1)} \left(\frac{e^t(\beta+1-e^t)^2 - e^t 2(\beta+1-e^t)(-e^t)}{(\beta+1-e^t)^4} \right. \\ &\quad \left. - \frac{e^t(\alpha\beta+\beta+1-e^t)^2 - 2e^t(\alpha\beta+\beta+1-e^t)(-e^t)}{(\alpha\beta+\beta+1-e^t)^4} \right)_{t=0} \\ &= \frac{(\alpha+1)\beta(\beta+1)(\alpha\beta+\beta+1)}{\alpha((\alpha+2)\beta+1)} \left(\frac{\beta+2}{\beta^3} - \frac{(\alpha+1)\beta+2}{(\alpha+1)^3\beta^3} \right) \\ &= \frac{(\beta+1)(\alpha\beta+\beta+1)(\alpha^2\beta+2\alpha^2+3\alpha\beta+6\alpha+2\beta+6)}{(\alpha+1)^2(\alpha\beta+2\beta+1)\beta^2} \end{aligned}$$

Now deriving the variance of X is possible,

$$\begin{aligned} \text{Var}(X) &= E(X^2) - E^2(X) \\ &= \frac{(\beta+1)(\alpha\beta+\beta+1)(\alpha^2\beta+2\alpha^2+3\alpha\beta+6\alpha+2\beta+6)}{(\alpha+1)^2(\alpha\beta+2\beta+1)\beta^2} \\ &\quad - \left[\frac{(\alpha\beta+\beta+1)(\beta+1)(\alpha+2)}{((\alpha+2)\beta+1)\beta(\alpha+1)} \right]^2 \\ &= \frac{(\beta+1)(\alpha\beta+\beta+1)(\alpha^3\beta+5\alpha^2\beta+\alpha^2+9\alpha\beta+2\alpha+6\beta+2)}{(\alpha+1)^2(\alpha\beta+2\beta+1)^2\beta^2} \end{aligned}$$

Plot the mean and variance with paramter α and β .

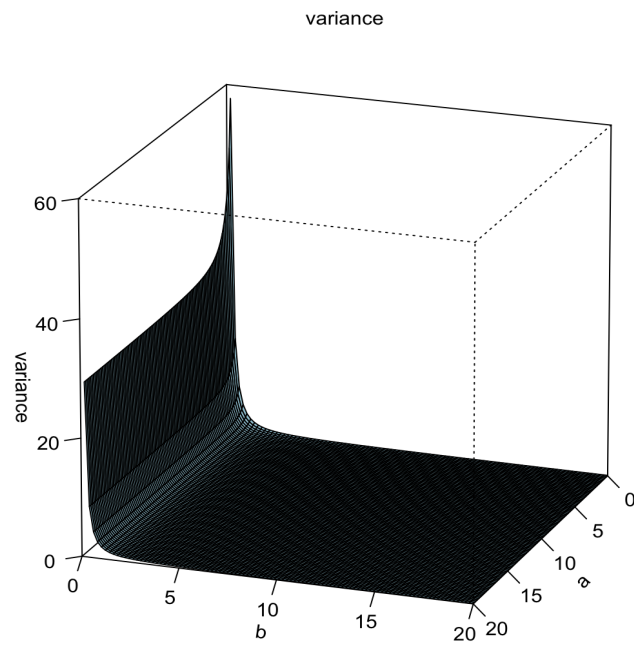
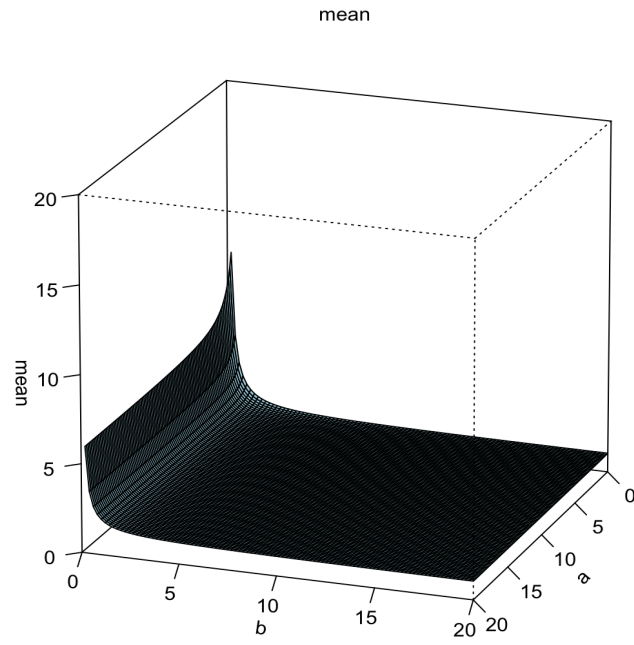


Figure 3.7: mean and variance plots of the ZTPWE distribution with parameter α and β

According to the two figures, the mean and variance values decrease as α and β increase.

3.2.5 Parameter Estimation

The maximum likelihood estimation is applied to obtain the parameter estimators of the ZTPWE distribution in this section. Assume that X_1, X_2, \dots, X_n represent a random sample of size n that follow the ZTPWE distribution with parameter α and β , the observed values are x_1, x_2, \dots, x_n respectively. Then by definition, the likelihood function of the ZTPWE distribution is:

$$\begin{aligned} L(\alpha, \beta) &= \prod_{i=1}^n f(x_i) \\ &= \prod_{i=1}^n (\alpha + 1) \beta \frac{(\beta + 1)^{-x_i} (\alpha \beta + \beta + 1)^{-x_i} ((\alpha \beta + \beta + 1)^{x_i+1} - (\beta + 1)^{x_i+1})}{\alpha ((\alpha + 2) \beta + 1)} \end{aligned}$$

Now apply the log-likelihood method,

$$\begin{aligned} \log L(\alpha, \beta) &= n \log (\alpha + 1) + n \log \beta + \sum_{i=1}^n \log ((\alpha \beta + \beta + 1)^{x_i+1} - (\beta + 1)^{x_i+1}) \\ &\quad - \left(\sum_{i=1}^n x_i \right) \log (\beta + 1) - \left(\sum_{i=1}^n x_i \right) \log (\alpha \beta + \beta + 1) - n \log \alpha \\ &\quad - n \log ((\alpha + 2) \beta + 1). \end{aligned}$$

Then take the first partial derivative of the log-likelihood function with respect to each parameter, α and β , obtain:

$$\begin{aligned}\frac{\partial \log L(\alpha, \beta)}{\partial \alpha} &= \frac{n}{\alpha + 1} + \sum_{i=1}^n \frac{\beta(x_i + 1)(\alpha\beta + \beta + 1)^{x_i}}{(\alpha\beta + \beta + 1)^{x_i+1} - (\beta + 1)^{x_i+1}} \\ &\quad - \left(\sum_{i=1}^n x_i \right) \frac{\beta}{\alpha\beta + \beta + 1} - \frac{n}{\alpha} - \frac{\beta n}{\alpha\beta + \beta + 1},\end{aligned}$$

$$\begin{aligned}\frac{\partial \log L(\alpha, \beta)}{\partial \beta} &= \frac{n}{\beta} + \sum_{i=1}^n n \frac{(\alpha + 1)(x_i + 1)(\alpha\beta + \beta + 1)^{x_i} - (x_i + 1)(\beta + 1)^{x_i}}{(\alpha\beta + \beta + 1)^{x_i+1} - (\beta + 1)^{x_i+1}} \\ &\quad - \frac{\sum_{i=1}^n x_i}{\beta + 1} - \left(\sum_{i=1}^n x_i \right) \frac{\alpha + 1}{\alpha\beta + \beta + 1} - \frac{(\alpha + 2)n}{(\alpha + 2)\beta + 1}.\end{aligned}$$

Thus the MLE can be obtained by setting the above score functions equal to 0. Although the above score functions do not have closed-form solutions, the MLE can still be obtained by the numerical methods. The BFGS method is applied from the optimx function (Nash and Varadhan) of R language.

As $n \rightarrow \infty$, the distribution of $\sqrt{n}(\hat{\alpha} - \alpha, \hat{\beta} - \beta)$ is an asymptotically bivariate normal with zero means and the variance-covariance matrix that which is calculated as the inverse of the Fisher information matrix. Due to the complexity of the Fisher information matrix, it is replaced by the observed information matrix.

Chapter 4

SIMULATION and APPLICATION

Simulation of Paramter Estimators

In this chapter the performance of the MLE, $\hat{\alpha}, \hat{\beta}$ are investigated. Generate random samples, and the steps are:

- i. Generate $u_i; i = 1, 2, \dots, n$ from $U(0, 1)$ distribution.
- ii. Set $x_i = F^{-1}(u_i)$.
- iii. Repeat n times to obtain the sample values $x_i; i = 1, 2, \dots, n$.

The simulation will be repeated 1000 times for $\alpha = 0.5, 1$. The sample size n would be 20, 50, 100, 200, and 500. The measures of accuracy of the estimators for parameter α are:

$$(i) \text{ Average Mean Square Error MSE} = \frac{\sum_{i=1}^{1,000} (\hat{\alpha}_i - \alpha)^2}{1,000},$$

$$(ii) \text{ Average Bias} = \frac{1}{1000} \sum_{i=1}^{1,000} (\hat{\alpha}_i - \alpha).$$

Table 4.1: Average MSE (average bias) of the simulated maximum likelihood estimates for α

n	α	β		
		0.5	1	2
20	0.5	0.0069 (0.0099)	0.2474 (0.4589)	2.1519 (1.3640)
50		0.0025 (0.0063)	0.2043 (0.4383)	1.6448 (1.2524)
100		0.0012 (0.0034)	0.1902 (0.4294)	1.5531 (1.2324)
200		0.0006 (0.0011)	0.1818 (0.4230)	1.4934 (1.2151)
500		0.0002 (0.0005)	0.1792 (0.4219)	1.4731 (1.2109)
20	1	0.2037 (-0.4414)	0.0505 (0.0425)	1.3581 (1.0121)
50		0.2023 (-0.4462)	0.0247 (0.0284)	0.9213 (0.9077)
100		0.2040 (-0.4499)	0.0079 (0.0117)	0.8330 (0.8867)
200		0.2056 (-0.4526)	0.0037 (0.0052)	0.7752 (0.8678)
500		0.2058 (-0.4533)	0.0015 (0.0021)	0.7467 (0.8589)

The same procedure is applied to parameter β for value of 0.5, 1, 2.

4.1 Simulation Result

Table 4.1 shows the average MSE and bias(in parentheses) of the MLE of parameter α . The MSE value should decreases as n increases. A good average bias should go as close as possible to 0 when n increases. For most combinations of α and β , the average MSE decreases when n increases and the average bias shrinks close to 0 as n increases. A similar conclusion can be made to β . Thus the estimators of α and β have pretty good accuracy.

Table 4.2: Average MSE (average bias) of the simulated maximum likelihood estimates for β

n	α	β		
		0.5	1	2
20	0.5	1.4669 (1.2034)	0.0699 (0.2355)	1.2095 (-1.0941)
50		1.4448 (1.1989)	0.0618 (0.2366)	1.1745 (-1.0816)
100		1.4427 (1.1997)	0.0604 (0.2394)	1.1692 (-1.0803)
200		1.4480 (1.2026)	0.0604 (0.2425)	1.1622 (-1.0775)
500		1.4474 (1.2028)	0.0595 (0.2425)	1.1612 (-1.0774)
20	1	1.2880 (1.1268)	0.0485 (0.1854)	1.2813 (-1.1266)
50		1.2678 (1.1228)	0.0417 (0.1880)	1.2542 (-1.1179)
100		1.2711 (1.1258)	0.0370 (0.1843)	1.2494 (-1.1167)
200		1.2744 (1.1281)	0.0359 (0.1852)	1.2428 (-1.1143)
500		1.2752 (1.1289)	0.0355 (0.1866)	1.2394 (-1.1130)

4.2 Application

Finally, fit ZTPWE distribution into two real data sets and compare the result with several competitors. The first data set is the birth order of 1800 students enrolled in elementary psychology classes at the University of California in 1924 to 1929 (Guttorp, 1991). The second data set is the number of occupants in 1469 cars (Hand, Daly, Lunn, McConway, and Ostrowsky, 1994). The competitors are ZTP, ZTPL, and ZTTPL distributions. The criteria for model selection are the lowest of the Akaike information criterion(AIC), the highest of the log-likelihood, and the highest of p -value based on the discrete Anderson-Darling(AD) goodness of fit test (Choulakian, Lockhart and Stephens, pp. 125-137). The results are shown in the Tables:

Table 4.3: The birth order of 1800 students enrolled in elementary psychology classes at the University of California in 1924 to 1929

Birth order	frequencies	ZTP	ZTPL	ZTTPPL	ZTPWE
1	797	603.8408	782.3003	807.3863	798.5009
2	455	571.8339	458.2171	453.2348	452.2384
3	265	361.0157	258.1594	245.5257	248.9240
4	125	170.9400	141.5655	135.3956	136.2151
5	68	64.7517	76.0989	74.6642	74.4478
6	37	20.4398	40.2869	41.1737	40.6786
7	26	5.5304	21.0717	22.7053	22.2258
8	8	1.3093	10.9137	12.521	12.1435
9	1	0.2755	5.6069	6.9047	6.6348
10	9	0.0522	2.8609	3.8076	3.6250
11	5	0.0090	1.4513	2.0997	1.9806
12	3	0.0014	0.7325	1.1579	1.0821
13	1	0.0002	0.3681	0.6385	0.5912
Estimated parameters		$\hat{\lambda}=1.8940$ (0.0367)	$\hat{\theta}=1.1215$ (0.0299)	$\hat{\theta}=0.8135$ (0.1637)	$\hat{\alpha}= 8.2702$ (8.8865)
(se)				$\hat{\alpha}=0.0001$ (0.1617)	$\hat{\beta} = 0.8303$ (0.0391)
Log-likelihood		-2934.301	-2762.626	-2760.31	-2760.130
AIC		5870.601	5527.252	5524.62	5524.260
AD statistic		44.8651	0.5512	0.3091	0.2031
p -value		0.01	0.4977	0.7239	0.8515

Table 4.3 illustrates that the ZTPWE distribution has the minimum AIC, maximum log-likelihood, and maximum p -value among all the competitor distributions. Notice that the ZTP distribution has p -value less than 5%, which means it cannot fit this data very well.

Table 4.4: The numbers of occupants in 1469 cars

# of occupants	freq	ZTP	ZTPL	ZTTPPL	ZTPWE
1	902	876.7183	939.3755	925.054	924.5196
2	403	419.3416	342.9273	359.950	360.6375
3	106	133.7164	121.9941	124.971	125.0469
4	38	31.9788	42.5752	40.739	40.6487
5	16	6.1183	14.6407	12.759	12.6850
6	4	0.9755	4.9759	3.887	3.8486
Estimated parameters		$\hat{\lambda}=0.9566$ (0.0315)	$\hat{\theta}=2.2600$ (0.0864)	$\hat{\theta}=2.8256$ (0.1210)	$\hat{\alpha} = 0.00002$ (0.0020)
(se)				$\hat{\alpha}=82.0071$ (190.7117)	$\hat{\beta} = 2.8453$ (0.1138)
Log-likelihood		-1487.233	-1484.41	-1481.053	-1480.958
AIC		2976.466	2970.821	2966.105	2965.915
AD statistic		1.7274	2.3561	1.0951	1.0584
p -value		0.0853	0.0380	0.1869	0.1967

Table 4.4 still indicates that the ZTPWE distribution is the most proper model to fit the data, though the p -value is not significant. For this data set, the ZTPL model has p -value less than 5%.

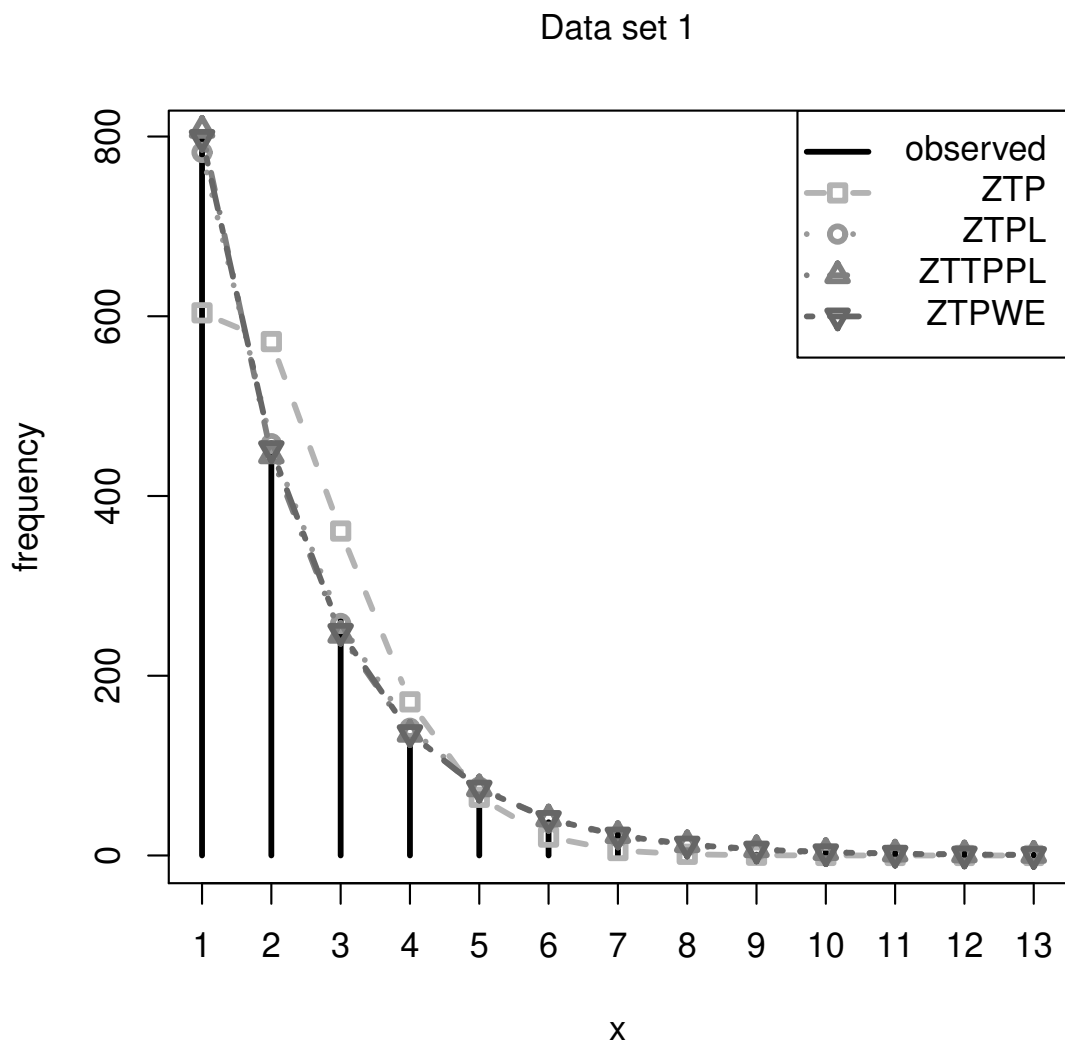


Figure 4.1: Plots of the observed data points and expected values

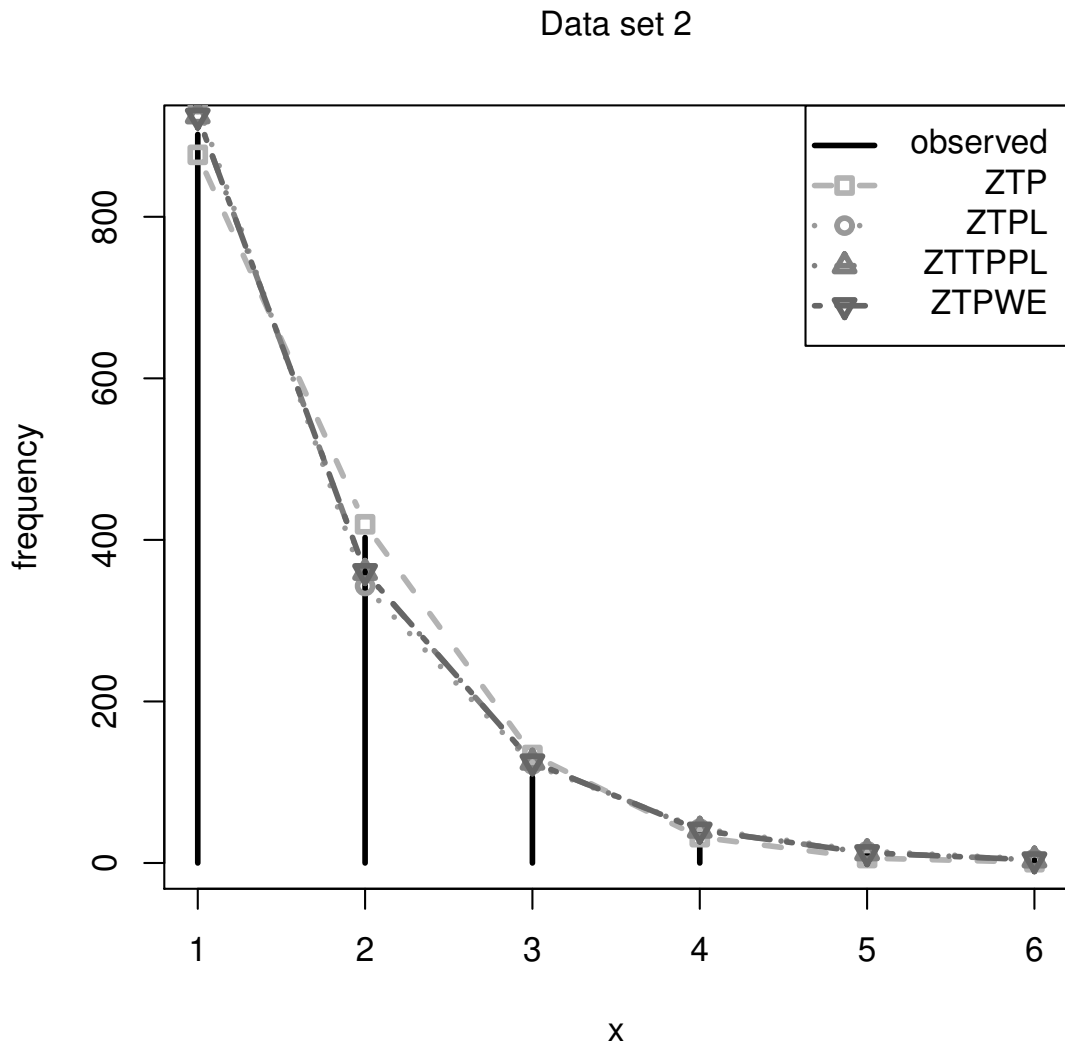


Figure 4.2: Plots of the observed data points and expected values

4.2.1 Plots of Observed Values and Models

Figure 4.1 contains observed values and expected values of four zero-truncated models fitted to the birth order data. It is noticeable that ZTP distribution cannot describe the data very well. Figure 4.2 are the models fitted to the number of occupancy in a car data, it seems most distribution fits kind of close, but ZTP still have deviation when $x = 1$, and $x = 2$. This implies that the idea of modified complex model is indeed the correct way to find a better fit of non-zero data. ZTPWE is not necessarily the best model for all non-zero datas, but compare it to the previous models in these two real data sets fitting, it appears that it can be an alternative for such kinds of data sets. Remind that there are different kinds of data sets. For data that contains a vast zero group, the other modified models should be considered too.

Chapter 5

CONCLUSIONS AND DISCUSSIONS

5.1 Theoretical Part

The objective of this research is to establish the ZTPWE distribution. The characteristics of this distribution are carefully studied, and two sets of real data points are fitted to this distribution. Several other distributions are also compared that gives the conclusion of the following:

5.1.1 Advantage for using ZTPWE distribution

First, the zero-truncated distributions are proper for data sets with no zero counts. Second, the two-parameter P-WE distribution is a more flexible model for dispersed data. The modified model, ZTPWE, performs well when all the models are compared after fitting into real data sets. Under certain circumstances, the ZTPWE distribution has its own advantage to describe positive data sets.

5.1.2 Progress for Construction of ZTPWE Model

- There are two different ways to obtain ZTPWE distribution. The first way is to substitute P-WE into zero-truncated transformation, the second way is to establish the unconditional distribution by combining a conditional random variable and its parameter with partly-exponential distribution. The two methods produce the same pmf of ZTPWE distribution is:

$$f(x) = \frac{(\alpha + 1)\beta}{\alpha((\alpha + 2)\beta + 1)}(\beta + 1)^{-x}(\alpha\beta + \beta + 1)^{-x}((\alpha\beta + \beta + 1)^{x+1} - (\beta + 1)^{x+1})$$

- The cdf of ZTPWE distribution is represented as:

$$F(x) = \frac{\alpha((\alpha + 2)\beta + \alpha + (\beta + 1)(\alpha\beta + \beta + 1)^{-x} - (\alpha + 1)(\alpha\beta + \beta + 1)(\beta + 1)^{-x}}{\alpha((\alpha + 2)\beta + 1)}$$

- After that, the pgf, mgf and, characteristic function of ZTPWE distribution can be derived:

$$\begin{aligned} G(s) &= \frac{(\alpha + 1)\beta}{\alpha((\alpha + 2)\beta + 1)} \left[\frac{s(\alpha\beta + \beta + 1)}{\beta + 1 - s} - \frac{s(\beta + 1)}{\alpha\beta + \beta + 1 - s} \right] \\ \psi(t) &= \frac{(\alpha + 1)\beta}{\alpha((\alpha + 2)\beta + 1)} \left[\frac{e^t(\alpha\beta + \beta + 1)}{\beta + 1 - e^t} - \frac{e^t(\beta + 1)}{\alpha\beta + \beta + 1 - e^t} \right] \\ \varphi(t) &= \frac{(\alpha + 1)\beta}{\alpha((\alpha + 2)\beta + 1)} \left[\frac{e^{it}(\alpha\beta + \beta + 1)}{\beta + 1 - e^{it}} - \frac{e^{it}(\beta + 1)}{\alpha\beta + \beta + 1 - e^{it}} \right] \end{aligned}$$

- Then, the mean and variance of ZTPWE distribution can be obtained:

$$E(x) = \frac{(\alpha\beta + \beta + 1)(\beta + 1)(\alpha + 2)}{((\alpha + 2)\beta + 1)\beta(\alpha + 1)}$$

$$\text{Var}(X) = \frac{(\beta + 1)(\alpha\beta + \beta + 1)(\alpha^3\beta + 5\alpha^2\beta + \alpha^2 + 9\alpha\beta + 2\alpha + 6\beta + 2)}{(\alpha + 1)^2(\alpha\beta + 2\beta + 1)^2\beta^2}$$

- Finally, the likelihood function of ZTPWE distribution is:

$$L(\alpha, \beta) = \prod_{i=1}^n (\alpha+1)\beta \frac{(\beta + 1)^{-x_i y_i} (\alpha\beta + \beta + 1)^{-x_i} ((\alpha\beta + \beta + 1)^{x_i+1} - (\beta + 1)^{x_i+1})}{\alpha((\alpha + 2)\beta + 1)}.$$

5.2 Response of Simulation

Run the simulations 1000 times for the different sizes of generated samples. The conclusion is: when the sample size (n) increases,

- The average MSE decreases.
- The average bias moves towards 0.

5.3 Interpretation of Application

Four zero-truncated distributions are fitted to two real data sets and inspected the result of three criteria information. The results are:

- The ZTPWE distribution has the minimum AIC.
- The ZTPWE distribution has the maximum log-likelihood.
- The ZTPWE distribution has the maximum p -value.

5.4 Future Study and Expectations

One crucial field that remains unrevealed is the partly distribution for the conditional distribution of random variables. It seems to have various in statistics research. Another field that needs more work is the goodness-of-fit test. There are more standards that can be used to justify the fitted model with data sets.

Bibliography

- [1] David, F. N., & N. L. Johnson, "The Truncated Poisson." *Biometrics*, 8(4), 275-285(1952).
- [2] Shanker, R., "A zero-truncated Poisson-Amarendra distribution and its application," *International Journal of Probability and Statistics*, 6(4), 82-92(2017).
- [3] Shanker, R., & K. K. Shukla, "THE POISSON-WEIGHTED AKASH DISTRIBUTION AND ITS APPLICATIONS." *Journal of Applied Quantitative Methods*, vol. 13, no. 2, 2018, p. 23.
- [4] Shanker, R., & K. K. Shukla, "A Zero-Truncated Two-Parameter Poisson-Lindley Distribution with an Application to Biological Science." *Turkiye Klinikleri Biyoistatistik*, vol. 9, no. 2, 2017, pp. 85-95.
- [5] Rohatgi, V., & A. Saleh, (2001). *An Introduction to Probability and Statistics* (2nd ed. /Vijay K. Rohatgi, A.K. Nd. Ehsanes Saleh.. ed., Wiley series in probability and statistics. Texts and references section). New York: Wiley.

- [6] Gupta, R. D., & D. Kundu, (2009). A New Class of Weighted Exponential Distributions. *Statistics (Berlinn, DDR)*, 43(6), 621-634.
- [7] H. Zamani, N. Ismail & P. Faroughi, "Poisson-Weighted Exponential univariate version and regression model with applications," *Journal of Mathematics and Statistics* 10(2), 148-154 (2014).
- [8] Gut, A., An Intermediate Course in Probability. *Springer, New York*, 2009).
- [9] Shao, J. (2003). *Mathematical Statistics* Springer Texts in Statistics. 2nd ed.
- [10] Grimmett, G., & D. J. A. Welsh, (2014). *Probability: An Introduction* (Second ed.).
- [11] R: A Language and Environment for Statistical Computing, R Core Team, R Foundation for Statistical Computing, Vienna, Austria (2018). <https://www.R-project.org/>. Accessed 2021.
- [12] R: Johnson N. L., A. W. Kemp, & S. Kotz, *Univariate Discrete Distributions* (John Wiley & Sons, New York, 2005)
- [13] Casella, G., & R. L. Berger, (1990). *Statistical Inference*. United States of America: Wadsworth.
- [14] Guttorp, P., *Statistical Inference for Branching Processes* (Wiley, New York, 1991).

- [15] Choulakian, V., R. A. Lockhart and M.A Stephens, (1994) Cramer-Von Mise Statistics for discrete distributions, *The Canadian Journal of Statistics*, 125-137.
- [16] Nash, J.C., and R. Varadhan, (2011) Unifying Optimization Algorithms to Aid Software System Users: Optimx for R, *Journal of Statistical Software*, 43(9), 1-14.
- [17] Konishi, S, & G. Kitagawa, (2008) Information Criteria and Statistical Modeling. *Springer series in statistics*. New York: Springer.
- [18] Jantschi, L. & D. D. Bolboaca, Computation of Probability Associated with Anderson-Darling Statistics. *Mathematics*. 2018, 6, 88.
- [19] Anderson, T.W., & D.A. Darling. "Asymptotic Theory of Certain 'Goodness of Fit' Criteria Based on Stochastic Processes." *The Annals of Mathematical Statistics*, Vol.23, no.2, 1952, pp. 193-212.
- [20] Jones, M. C., *Families of Distributions arising from Distributions of Order Statistics*, Test 13 (2004), pp 1-43.
- [21] Arnold, B. C., & R. J. Beaver, *Hidden Truncation Models*, Sankhya, Ser. A 62(2000), pp. 23-35.
- [22] Denuit, M., 1997. *A New Distribution of Poisson- Type for the Number of Claims*. ASTIN BULL. 27: 229-242. DOI:10.2143/AST.27.2.542049.

- [23] Deb, P. & P. K. Trivedi, (1997) *Demand for Medical Care by the Elderly: A Finite Mixture Approach*. J. Applied Econometrics, 12: 313-336.
- [24] Bjerkedal, T.,(1960) *Acquisition of Resistance in Guinea Pigs Infected with different doses of virulent tubercle bacilli*, American Journal of Epidemiology, 72(1), pp. 130-148.

Appendix. R Source Code

R Source Code for Simulation

```
test<-function(x)
{
  xnew <- unique(sort(x))

  prob_pztpwe <- pztpwe(xnew,a,b)

  StepProb_pztpwe <- stepfun(xnew, c(0,prob_pztpwe))

  library(dgof)

  dgof::cvm.test(x, StepProb_pztpwe,type="A2") -> ADPZTPWE

  p <- ADPZTPWE$p

  return(p)
}
```

```
setwd("~/Downloads/ZTP test")
```

```
t<-1000
```

```
n<-c(20,50,100,200)
```

```
a <- 1;b<- 1
```

```
k<-1
```

```
while(k <=length(n))
```

```
{
```

```
  dir_sample <- paste("Sample/",sep="\t")
```

```
  file_name_sample <- paste(dir_sample,"alpha ",a," beta ",b," n ",n[k],".txt",sep
```

```
  i<-1
```

```
  set.seed(99)
```

```
  while(i<=t) {
```



```

x <- numeric()

x <- dztpwe(n[k],a,b)

p <- test(x)

while(p<0.1)
{
  x<-numeric()

  x<-dztpwe(n[k],a,b)

  p<-test(x)
}

if(i==1) {
  write.table(cbind(i,t(x),p), file_name_sample, sep="\t", col.name=NA)
}

#else

if(i!=1) {
  write.table(cbind(i,t(x),p), file_name_sample, sep="\t", append=TRUE,col.name=NA)
}

i<-i+1
}

```

```

    k<-k+1
}

```

R code for ZTPWE MSE estimation

```

dztpwe<-function(x,a,b)
{
  p=((a + 1) *b* (b + 1) *(a* b + b + 1)* ((b + 1)^(-x - 1) - (a *b + b + 1)^(-x -
  return(p)
}

```

```

ll = function (para)
{
  a=exp(para[1])
  b=exp(para[1])
  #a=para[1]
  #b=para[2]
  f = log(dztpwe(dx,a,b))
  return(-sum(f))
}

```

```
}
```

```
library(optimx)
```

```
setwd("~/Downloads/ZTP test/")
```

```
t<-1000
```

```
n<-c(20,50,100,200)
```

```
###Parameter sets
```

```
a <- 1;b <- 1
```

```
output<-paste("alpha ",a," beta ",b,sep="")
```

```
dir.create(output)
```

```
k<-1
```

```
while(k <=length(n))
```

```
{
```

```

if(!require(optimx)){
  install.packages('optimx')
}

library(optimx)

est_alpha<-numeric()

est_beta<-numeric()

err_alpha<-numeric()

err_beta<-numeric()

bias_alpha<-numeric()

bias_beta<-numeric()

dir_sample <- paste("~/Downloads/ZTP test/",sep="\t")

file_name_sample <- paste(dir_sample,"alpha ",a," beta ",b," n ",n[k],".txt",sep=""

data<-read.table(file_name_sample,header=T)

j<-1

set.seed(99)

```

```

while(j<=t)
{

x<-data[j,4:ncol(data)-1]

dx<-t(x)


out_PNWL<-optimx(c(.1,.1),ll,hessian =T)

LL <- -out_PNWL$value

est_alpha[j]<-exp(out_PNWL$p1[1])
est_beta[j]<-exp(out_PNWL$p2[1])

err_alpha[j]<-(a-est_alpha[j])^2
err_beta[j]<-(b-est_beta[j])^2

bias_alpha[j]<-(est_alpha[j]-a)
bias_beta[j]<-(est_beta[j]-b)

```

```

####set directory for output saving

dir_output <- paste(output, "/", sep="")

file_name_output <- paste(dir_output, "mle_estimates alpha ", a, " beta ", b, " n "
file_name_loss <- paste(dir_output, "mle_error alpha ", a, " beta ", b, " n ", n[k],

if(j==1)
{

write.table(cbind(j, est_alpha, est_beta), file_name_output, sep="\t", col.names=colnames(est_alpha))
write.table(cbind(j, err_alpha, err_beta, bias_alpha, bias_beta), file_name_loss, sep="\t", col.names=colnames(err_alpha))

}

else if(j!=1)
{

write.table(cbind(j, est_alpha[j], est_beta[j]), file_name_output, sep="\t", col.names=colnames(est_alpha))
write.table(cbind(j, err_alpha[j], err_beta[j], bias_alpha[j], bias_beta[j]), file_name_loss, sep="\t", col.names=colnames(err_alpha))
}

```

```

    }

    j=j+1
}

alpha_est<-round(mean(est_alpha),4)
sd_alpha<-round(sd(est_alpha),4)

beta_est<-round(mean(est_beta),4)
sd_beta<-round(sd(est_beta),4)

MSE_alpha<-round(mean(err_alpha),4)
RMSE_alpha<-round(sqrt(MSE_alpha),4)
average_bias_alpha<-round(mean(bias_alpha),4)

MSE_beta<-round(mean(err_beta),4)
RMSE_beta<-round(sqrt(MSE_beta),4)
average_bias_beta<-round(mean(bias_beta),4)

size<-n[k]

```

```

dir_result <- paste("Result/", sep="")

file_name_result <- paste(dir_result,"mle_result alpha ",a," beta ",b," n ",n[k]

if(k==1)
{

    write.table(cbind(size,alpha_est,sd_alpha,beta_est,sd_beta,MSE_alpha,MSE_beta,

}else if(k!=1){

    write.table(cbind(size,alpha_est,sd_alpha,beta_est,sd_beta,MSE_alpha,MSE_beta,

}

k <- k+1

}

```