Bivariate Pseudo Modeling

Submitted in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy
in
Statistics

by

Mr. Banoth Veeranna, M.Sc
17MSPS01

Under the supervision of

Dr. B. G. Manjunath
Dr. B. Shobha

हैदराबाद विश्वविद्यालय
University of Hyderabad

प्रतिष्ठित संस्थान
INSTITUTION OF EMINENCE
राष्ट्रीय अपेक्षाएँ, वैश्विक मानक
National Needs, Global Standards

School of Mathematics and Statistics
University of Hyderabad
Hyderabad, India - 500 046.
February 2023

© 2023 Banoth Veeranna. All rights reserved.

To my dear beloved grandfather, I dedicate my thesis...

THE SCHOLAR'S DECLARATION

I hereby certify that:

- This thesis contains original work that I completed under the supervision of my supervisor.
- No other Institute has received the work for the purpose of awarding a degree or diploma.
- I prepared the thesis according to the instructions given by the Institute.
- I have complied with the standards and recommendations stated in the Institute's Ethical code of Conduct.
- Whenever I have used Information (data, theory, and text) from another sources, I have credited the source appropriately by citing it in the thesis body text and providing further information about it in the reference section.
- Anti-plagiarism software has reviewed the thesis.

Banoth Veeranna
Mr. Banoth Veeranna, M.Sc.
17MSPS01
February 2023

CERTIFICATE

This is to verify to the reality that the thesis titled "**Bivariate Pseudo Modeling**" by **Banoth Veeranna** bearing Reg. No. 17MSPS01 in partial fulfilment of the requirements for the award of **Doctor of Philosophy in Statistics** is a bonafide work carried out by him under my supervision and guidance.

The thesis has not been submitted previously in part or in full to this or any other University or Institution for the award of any degree or diploma.

Further, prior to submitting his thesis monograph for review, the student had the following publication(s), and he had provided documentation for each of them in the form of an acceptance latter or a reprint in the area of his research that was pertinent to the submission:

1. Barry C. Arnold, Banoth Veeranna, B. G. Manjunath and B. Shobha (2022), *"Bayesian inference for pseudo-Poisson data"*, Journal of Statistical Computation and Simulation. https://doi.org/10.1080/00949655.2022.2123918

2. Banoth Veeranna, B. G. Manjunath and B. Shobha (2022), *"Goodness of fit tests for the pseudo-Poisson distribution"*. (Under revision)

3. Banoth Veeranna (2022), *"A note on Bayesian inference for the bivariate pseudo-exponential data"*. (Submitted).

and

has made presentations in the following conferences:

1. International Virtual Conference on Prof. C.R. Rao's School of Thought on Statistical Sciences (ICON-CRRAO-STOSS 2020) organized by Department of Statistics, Pondicherry University, and Department of Mathematics and Statistics, and Mrs. A.V.N. College, Visakhapatnam, held on $21^{st} - 22^{nd}$ & $28^{th} - 29^{th}$ November, 2020. Title of the paper: *"Bayesian Analysis of the pseudo-Poisson distribution."*

2. International Conference on Knowledge Discovers on Statistical Innovations & Recent Advances in Optimization (ICON-KSRAO-2022) organized by Department of Statistics, Andhra University, Visakhapatnam on 29^{th} & 30^{th} December, 2022. Title of the paper: *"A Note on Goodness-of-Fit Tests for the Bivariate Pseudo-Poisson Model."*

3. International Conference on Statistics, Probability, Data Science and Related Areas (ICSPDS2023) in conjunction with 42^{nd} Annual Convention of Indian Society for Probability and Statistics (ISPS), held during $04^{th} - 06^{th}$ January, 2023, organized jointly by the Department of Statistics, Cochin University of Science and Technology & Indian Society for Probability and Statistics (ISPS). Title of the paper: *"Some remarks on pseudo-Poisson and its goodness-of-fit."*

Further, the student has completed the following courses to meet the coursework requirements for the doctoral degree:

Sl. No.	Course Code	Course Name	Credits	Pass/Fail
1	ST 813	Bivariate Distributions	4	Pass
2	SS 803	Analysis	5	Pass
3	SS 805	Statistics-1	5	Pass
4	CS 800	Research Methods in Computer Science	4	Pass

Thesis Supervisor:

Dr. B.G. Manjunath

Dr. B.G. MANJUNATH
Assistant Professor
School of Mathematics & Statistics
University of Hyderabad
Hyderabad-500 046. T.S.

Dean of the School:

Prof. R. Radha

संकाय-अध्यक्ष / DEAN
गणित और सांख्यिकी संकाय
School of Mathematics & Statistics
हैदराबाद विश्वविद्यालय/University of Hyderabad
हैदराबाद/Hyderabad-500 046. तेलंगाना, T.S.

Acknowledgement

First, I am grateful to Sree Sree Veerabhadra Swamy for giving me an opportunity, physical and mental strength, to pursue research work.

I sincerely thank and express my deep gratitude to my mentor Dr. B. G. Manjunath, for his guidance, constant encouragement, mental support, freedom, and valuable suggestions throughout the Ph.D. program. I appreciate his patience, kindness, and humanistic approach. I am fortunate to have pursued my Ph.D. work under the guidance of an honest man.

I have to thank my research co-supervisor, Dr. B. Shobha. Without their assistance and dedicated involvement in every step throughout the process, this work would have never been accomplished. Thank you very much for your support and understanding these past five years.

My sincere thanks to Prof. R. Radha, Dean, School of Mathematics and Statistics, as well as past school Deans for providing the infrastructure for carrying out the research work and all the faculty members for the help and encouragement at different times.

I give special thanks to my doctoral research committee members, Prof. Madhuchandra Bhattacharjee and Dr. B. Shobha, for helping in various ways.

I am grateful thank to Prof. Barry C. Arnold, University of California, USA, for their help in carrying out my collaborative research work.

I sincerely thank my teachers, Prof. V. Nageshwara Rao, Prof. A. Rajendra Prasad, Dr. Srinivas, Dr. Sri Laxmi, and Dr. Shaik. Ameen and others for building interest towards higher study.

I would like to record my juniors, B. Rajendar and Sachin Sachdev. And also thanks to my lab-mates Laba Sa, Mansoor Ali, Rubayya, Joydev Halder, Sandip Rout, Akshay

Kumar, Sreenanda S.B., and Devi Chandrika V. I would like to thank my seniors Dr. Arjit Mukarjee, and Dr. Pabitra Narayan Mandal for the fruitful discussions and suggestions. My special thanks to my senior, Mr. Laba Sa, for their help and guidance throughout my Ph.D.

I thank all non-teaching staff of the school of Mathematics and Statistics help and cooperation.

I also want to express my sincere thanks to all my brothers with whom I have some beautiful memories, and they are the people who took care of me throughout their stay.

My special thanks to my roommates P.Bharth, Abhimanyu, and Sudheer for various discussions on various topics; these are my local guardians.

I thank my friends, Bharath, Sudheer, Ravinder, Thaviti Naidu, Aditya Ankari, Satya, Yousaf Zamal, Stuti, Koyel, Atidhi Nayar, Subhangi, Dipeeka, G. Srinivas(Chinna), A. Srinivas, and others, for their help in various ways. I thank all my childhood and existing friends. When people lose some childhood friends during their Ph.D., they may get some new friends. I thank my new friends.

Getting through my thesis required more than academic support, and I have many, many people to thank for listening to me and, at times, having to tolerate me over the past five years. I cannot begin to express my gratitude and appreciation for their friendship Bharath, Ravinder, Sudheer, and Laba Sa have been unwavering in their personal ad professional support during the time I spent at the university. For many memorable evenings out and in, I must thank everyone.

Most importantly, none of this could have happened without my family. My grandfather offered his encouragement through phone calls – despite my own limited devotion to correspondence. With his own brand of humour, Guguloth Baliyya has been kind ad supportive to me over the last several years. To my parents and my sister – it would be an understatement to say that, as a family, we have experienced some ups and downs in the past five years. Every time I was ready to quit, you did not let me, and I am forever grateful. This thesis stands as a testament to your unconditional love and encouragement.

The thesis would be incomplete without my family members standing beside me on various stages of this journey. I am very grateful to my parents, Banoth Sukya and

Laxmi, and my brother and sister, Banoth Kumar and Supriya. I am fortunate that I get parents and brothers like them. I wholeheartedly thank my other family members: Banoth Ramchad & Santhi, Banoth Kishan & Koti, Banoth Laxman & Bujji & Swathi, Banoth Ramesh & Padma, Bhukya Gopi & Bhadri, Vankudoth Chinna & Guramma, Bujji, Jarpala Devsingh & Lali, Ajmeera Mangilal & Saroja, Guguloth Veeranna & Nagamani, Guguloth Naresh & Supriya(Jyothi).

Finally, I thank RGNFST for the fellowship, and I thank all the people who have helped me during this marathon journey.

Banoth Veeranna

Mr. Banoth Veeranna, M.Sc
17MSPS01
February 2023

Place: University of Hyderabad, Hyderabad
Date: February 1, 2023

Abstract

The existence of normal marginals and normal conditionals in the traditional bivariate normal distributions is widely known. It makes it natural to wonder if Poisson marginals and conditionals can experience a similar phenomenon. However, it is known from studies on conditionally specified models that Poisson marginals and both conditionals would only be seen in the scenario when the variables are independent. This thesis discusses a bivariate pseudo-Poisson model in which the conditional density of one variable and the marginal density of the other are both Poisson forms. These models are commonly used to model bivariate count data with a positive correlation. Moreover, such models have simple, flexible dependence structures and generate a sufficiently large number of parametric families. It has been a strong case made for the pseudo-Poisson model as the initial option to take into account when modeling bivariate over-dispersed data with positive correlation and having one of the marginal equi-dispersion. In the current thesis, we look at separate gamma priors for the parameters as well as pseudo-gamma priors for the Bayesian estimation of the unknown parameters of bivariate pseudo-Poisson models. Both comprehensive and sub-model investigations of potential conjugacy are verified, and conjugate priors can be found in some unique sub-instances. The effectiveness of Bayesian parameter estimates employing a range of priors, both informative and non-informative, is demonstrated through a simulation study. Two well-known bivariate count data sets are re-analyzed to illustrate the methodologies. Similarly, we also considered a bivariate pseudo-exponential model initially introduced by Arnold and Arvanitis (2019) for Bayesian analysis using pseudo gamma priors and independent gamma priors for the parameters. We also include an application to the Infant mortality and GDP data set.

Yet, before we start fitting, it is necessary to test whether the given data is compatible with the assumed pseudo-Poisson model. Hence, in the present note, we derive and propose a few goodness-of-fit tests for the bivariate pseudo-Poisson distribution. Also, we emphasize two tests, a lesser-known test based on the supremes of the absolute difference between the estimated probability generating function and its empirical counterpart. A new test has been proposed based on the difference between the estimated bivariate Fisher dispersion index and its empirical indices. However, we also consider the potential of applying the bivariate tests that depend on the generating function (like the Kocherlakota and Kocherlakota and Muñoz and Gamero tests) and the univariate goodness-of-fit tests (like the Chi-square test) to the pseudo-Poisson data. However, we analyze finite, large, and asymptotic properties for each of the tests considered. Nevertheless, we compare the power (bivariate classical Poisson and Conway-Maxwell bivariate Poisson as alternatives) of each of the tests suggested and also include examples of application to real-life data.

Keywords: Bivariate Pseudo-Poisson, MLE, Square Error Loss, Bayesian Analysis, marginal and conditional distributions, Empirical probability generating function, Goodness-of-fit, χ^2-goodness-of-fit, Neyman type A distribution, Index of dispersion, Parametric Bootstrap estimators, Consistency estimation, Thomas distribution.

List of Tables

List of Figures

List of Notations and Operations

X, Y, Z	Random Variable
$\mathbb{R}$	Real Number
N	Population Size
n	Sample Size
S_N	Random Sum
I	Indicator Function
α	Alpha
β	Beta
γ	Gamma
δ	Delta
τ	Tau
ψ	psi
χ^2	Chi-Square
λ	Lambda
θ	Theta
ϕ	phi
Σ	Covariance Matrix
Θ	Parameter Space
Γ	Gamma Function
$\underline{\theta}$	Parameter Vector
$\mathscr{N}$	Normal Distribution
$\mathscr{P}$	Poisson Distribution
$\mathscr{F}$	Pseudo-family Distribution
O	Observed Frequency
E	Expected Frequency

Abbreviations

d.f.	distribution function
p.d.f.	probability density function
p.m.f	probability mass function
c.g.f.	cumulative generating function
p.g.f.	probability generating function
e.p.g.f.	empirical probability generating function
r.v.	random variable
m.l.e.	maximum likelihood estimator
C.I.	Confidence Interval
AIC	Akaike Information Criterion
GoF	Goodness-of-Fit
asy.	Asymptotically
Ker	Kernel
HARM	Hit-And-Run Metropolis
HRS	Health retirement study
GDP	Gross domestic product
C.I.A	Central Intelligence Agency

Contents

CHAPTER 1

Introduction

The classical bivariate normal distribution is one in which both conditionals and marginals are normally distributed. That is, if random variables (X, Y) has bivariate normal distribution, $X \sim \mathcal{N}$ (normal distribution), $Y \sim \mathcal{N}$ and for each $x \in \mathbb{R}$, the conditional distribution of Y given $X = x \sim \mathcal{N}$, while for each $y \in \mathbb{R}$, the conditional distribution of X given $Y = y \sim \mathcal{N}$ also. Further the distributions of X and Y are also normal. One more example with a similar properties is the Mardia Pareto distribution in which both as marginals and conditionals have the Pareto form with dependence structure detailed in Mardia [25]. The Mardia bivariate Pareto with parameters θ_1, θ_2, and θ all of which are positive is given by

$$f(x, y, \theta_1, \theta_2, \theta) = \frac{\left[\theta(\theta+1)(\theta_1\theta_2)^{\theta+1}\right]}{(\theta_2 x + \theta_1 y - \theta_1\theta_2)^{\theta+2}} \tag{1.0.1}$$

where $x > \theta_1 > 0$ and $y > \theta_2 > 0$, $\theta > 0$, otherwise zero, and call it a bivariate Pareto distribution of type I. Note that for the above bivariate density, both marginal and conditionals again follows the Pareto distribution. The correlation between X and Y is positive. In general, the density function of k-dimensional random variables $(X_1, ..., X_k)$ is a Pareto distribution of type I with parameters $\theta_1, ..., \theta_k$ and θ is given by

$$f(x_1,...,x_k) = \frac{(\theta(\theta+1)...(\theta+k-1))}{\left(\prod_{i=1}^{k}\theta_i\right)\left\{\left(\sum_{i=1}^{k}\theta_i^{-1}x_i\right)-k+1\right\}^{\theta+k}}. \tag{1.0.2}$$

where $x_i > \theta_i > 0$, for $i = 1, 2, ..., k$ and $\theta > 0$, otherwise zero.

Now, it seems sensible to wonder if Poisson conditionals and Poisson marginals can experience a similar effect. A simple illustration of this type has two independent Poisson variables, X and Y. In reality, there are no other cases. Regarding the following theorem by Seshadri and Patil [35], which states that:

Theorem 1.1. *For a given dependent bivariate distribution of X and Y, the following statements hold:*

- *If X is Poisson, then the conditional distribution of X given $Y = y$ is not Poisson.*
- *If the conditional distribution of Y given $X = x$ is Poisson, then the marginal distribution of Y is not Poisson.*

For the detailed proof in Seshadri and Patil [35] page 216. That is, Seshadri and Patil [35] shows that it is impossible to have a non-independent bivariate distribution in which on X has a Poisson distribution and its conditional distribution of X given $Y = y$ is also Poisson distributed for each y. That is, there exists no bivariate random variables (X, Y) in which X and Y are not independent with Poisson marginals as well as Poisson conditionals. They only have Poisson marginals in the case of independence (which could be deduced using the Seshadri-Patil result).

In addition, Arnold, Castillo, and Sarabia [3] (for example) like wise consider a bivariate random variables (X, Y), the conditional distribution of Y given $X = x$ is Poisson for each x, while the conditional distribution of X given $Y = y$ is also Poisson for each y. It is called Poisson conditional distributions. In Ghosh et al.[16], one can find more information on conditional Poisson model properties and applications. For the bivariate pseudo-Poisson distribution random variables (X, Y) one of the marginals say X has Poisson distribution, while the conditional distribution of Y given $X = x$ is Poisson distribution. It is perfectly fit the bill.

1.1 Bivariate pseudo-Poisson models

The bivariate pseudo-Poisson model will be discussed in the section that follows. It was first introduced in Arnold and Manjunath [4], page 2307.

Definition 1.1. *A bivariate pseudo-Poisson distribution exists for a 2-dimensional random variable (X, Y) if there exists a positive constant λ_1 such that*

$$X \sim \mathscr{P}(\lambda_1)$$

and a function $\lambda_2 : \{0, 1, 2, ...\} \to (0, \infty)$ such that, for every non-negative integer x,

$$Y|X = x \sim \mathscr{P}(\lambda_2(x)).$$

Indeed, no constraints on the $\lambda_2(x)$ allow us to incorporate a variety of properties such as positive and negative correlation, for instance over and equi-dispersions. Hence, the joint probability mass function of X and Y is given by

$$P(X = x, Y = y) = \begin{cases} \frac{e^{-\lambda_1}\lambda_1^x}{x!}\frac{e^{-\lambda_2(x)}(\lambda_2(x))^y}{y!}; & x = 0, 1, 2, ...; y = 0, 1, 2, ... \\ 0; & Otherwise \end{cases}$$

Example 1. *Consider, a parametric family of choices for $\lambda_2(x)$ that will admit positive and negative correlation between X and Y. For example if we consider*

$$\lambda_2(x; \theta, \delta) = 1 + (2\theta - 1)(1 - e^{-\delta x}). \tag{1.1.1}$$

Where $\delta > 0$, the above function will be increasing if $\theta > 1/2$, decreasing if $\theta < 1/2$, and constant if $\theta = 1/2$. Consequently, X and Y will have positive correlation if $\theta > 1/2$, negative correlation if $\theta < 1/2$ and will be uncorrelated if $\theta = 1/2$. A more general model with the same properties can be obtained by replacing $1 - e^{-\delta x}$ by $F(x; \underline{\theta})$, a parametrized family of distribution functions with support $(0, \infty)$.

Here, we restrict the shape of the function $\lambda_2(x)$ to a polynomial with positive coefficients. In particular, the simple form we assume is that $\lambda_2(x) = \lambda_2 + \lambda_3 x$, then the bivariate distribution shown above will have the following form:

$$X \sim \mathscr{P}(\lambda_1) \tag{1.1.2}$$

and $x \in \{0, 1, 2, ...\}$

$$Y|X = x \sim \mathscr{P}(\lambda_2 + \lambda_3 x). \tag{1.1.3}$$

For this model, the parameter space is $\{(\lambda_1, \lambda_2, \lambda_3) : \lambda_1 > 0, \lambda_2 > 0, \lambda_3 \geq 0\}$. If we chose $\lambda_3 = 0$, then the corresponding random variables are independent. Then the joint probability mass function is given by

$$P(X = x, Y = y) = \begin{cases} \frac{e^{-\lambda_1}\lambda_1^x}{x!} \frac{e^{-\lambda_2 + \lambda_3 x}(\lambda_2 + \lambda_3 x)^y}{y!}; & x = 0, 1, 2, ...; y = 0, 1, 2, ... \\ 0; & Otherwise. \end{cases}$$

For this bivariate pseudo-Poisson distribution, the probability generating function (p.g.f.) is given by

$$G(t_1, t_2) = e^{\lambda_2(t_2 - 1)} e^{\lambda_1[t_1 e^{\lambda_3(t_2 - 1)} - 1]};\ t_1, t_2 \in \mathbb{R}. \tag{1.1.4}$$

Remark 1. *As noted in Arnold and Manjunath [4], for the case $\lambda_2 = 0$, the bivariate pseudo-Poisson distribution reduces to the bivariate Poisson-Poisson distribution. The corresponding Poisson-Poisson distribution was originally introduced by Leiter and Hamdani [24] in modelling traffic accidents and fatalities count data. We remark that the bivariate pseudo-Poisson model is a generalization of the Poisson-Poisson distribution.*

If $\lambda_2 = 0$, then the joint p.g.f in equation (1.1.4) reduces to

$$G_{II}(t_1, t_2) = e^{\lambda_1[t_1 e^{\lambda_3(t_2 - 1)} - 1]};\ t_1, t_2 \in \mathbb{R}. \tag{1.1.5}$$

Now, form equation (1.1.4), the marginal p.g.f of Y is

$$G(1,t_2) = G_Y(t_2) = e^{\lambda_2(t_2-1)}e^{\lambda_1[e^{\lambda_3(t_2-1)}-1]};\ t_2 \in \mathbb{R}. \tag{1.1.6}$$

Note that, in general, the p.g.f in equation (1.1.4) can not be simplified to compute all marginal distributions. Yet, we can use equation (1.1.5) to derive a few marginal distributions of Y. The derivation of marginal probability of Y is demonstrated for $Y = 0, 1, 2, 3$ in Section 1.1.3, and one can still extend the mentioned procedure to get albeit complicated values for the probability that Y assumes any positive value. Besides, the derivation of the other conditional distribution of the bivariate pseudo-Poisson, i.e., $P(X = x|Y = y)$, has been included in Section (1.1.4).

In the following sections, we discuss a few one-dimensional distributions which are derived from the bivariate pseudo-Poisson for the case $\lambda_2 = 0$. Moreover, the derived univariate distributions has classical relevance to the two-parameter Neyman Type A and Thomas distribution.

1.1.1 Neyman Type A distribution

As noted in Arnold and Manjunath [4], in the case in which $\lambda_2 = 0$, the marginal distribution is a Neyman Type A distribution with λ_3 being the index of clumping (detailed in page 403 of Johnson, Kemp, and Kotz [20]). It can also be recognized as a Poisson mixture of Poisson distributions. Now, the marginal mass function of Y is given by

$$P(Y = y) = \frac{e^{-\lambda_1}\lambda_3^y}{y!}\sum_{j=0}^{\infty}\frac{(\lambda_1 e^{-\lambda_3})^j j^y}{j!};\ y = 0, 1, 2, ... \tag{1.1.7}$$

i.e., Y has a Poisson distribution with the parameter λ_1 while λ_1 is also a Poisson random variable with parameter λ_3. We refer to Glesson and Douglas [17] and Johnson, Kemp, and Kotz [20] Section 9.6 for applications and inferential aspects of the Neyman Type A distribution.

1.1.2 Thomas distribution

Consider the joint probability generating function defined in equation (1.1.5), i.e.,

$$G_{II}(t_1, t_2) = e^{\lambda_1[t_1 e^{\lambda_3(t_2-1)}-1]};\ t_1, t_2 \in \mathbb{R}. \tag{1.1.8}$$

Take $t_1 = t_2 := t$ and the above p.g.f. reduces to

$$G^*(t) = G(t,t) = e^{\lambda_1[te^{\lambda_3(t-1)}-1]};\ t \in \mathbb{R}. \tag{1.1.9}$$

Note that the above univariate p.g.f. is the p.g.f. of the Thomas distribution with parameters λ_1 and λ_3. The probability mass function of Z that follows the Thomas distribution is

$$P(Z=z) = \frac{e^{-\lambda_1}}{z!}\sum_{j=1}^{z}\binom{z}{j}(\lambda_1 e^{-\lambda_3})^j (j\lambda_3)^{z-j},\ z = 0, 1, 2, ... \tag{1.1.10}$$

For further applications and inferential aspects of the Thomas distribution, we refer to Glesson and Douglas [17] and Johnson, Kemp, and Kotz [20] Section 9.10.

Remark 2. *The Neyman Type A and the Thomas distribution have historical relevance in modeling plant and animal populations. For example: suppose that the number of clusters of eggs an insect lays and the number of eggs per each cluster have specified probability distributions. Then for the Neyman Type A distribution and Thomas distributions, the number of clusters of eggs laid by the insect follows a Poisson distribution with parameter λ_1. For the Neyman Type A, the number of eggs per cluster is also a Poisson distribution with parameter λ_3. But for the Thomas distribution, the parent of the cluster is always to be present with the number of eggs(offspring) which has a shifted Poisson distribution with support $\{1, 2, 3, ...\}$ and parameter λ_3. Note that Neyman Type A and Thomas distributions can be generated by a mixture of distributions and also a random sum of random variables.*

Consider that the mixing distribution is a Poisson with parameter λ_1 with mixture has a Poisson with parameter λ_3, then the resultant random variable has a Neyman Type A distribution. In the sequel, if the mixing distribution is a Poisson with parameter λ_1 and the jth *distribution in the mixture has a distribution of the form* $j + Y(j)$*, where* $Y(j)$ *has a Poisson with parameter* $j\lambda_3$*, then the resultant random variable has Thomas distribution.*

However, for a random sum of random variables (also known has Stopped-Sum distributions): let us consider that the size N of the initial generation is a random variable and that each individual i of this generation independently gives a random variable Y_i*, where* $Y_1, Y_2, ...$ *has a common distribution. Then the total number of individuals is* $S_N = Y_1 + ... + Y_N$*. For the case that N is a Poisson random variable with parameter* λ_1 *and* Y_i *is a Poisson random variable with parameter* λ_3*, then the random sum* S_N *has a Neyman Type A distribution. However, if* Y_i *is a shifted Poisson with parameter* λ_3 *and support* $\{1, 2, 3, ...\}$*, then the random sum* S_N *has a Thomas distribution. A sum of N independent and identically distributed non-zero Poisson random variables, where N is also a Poisson random variable has a Thomas distribution.*

1.1.3 Marginal probability of Y

For the marginal distribution of Y, the probability that $Y = 0$ can be computed as

$$P(Y = 0) = G_Y(0) = e^{-\lambda_2} e^{\lambda_1 (e^{-\lambda_3} - 1)}. \tag{1.1.11}$$

For the probability that $Y = 1$ we have

$$\frac{d}{dt} G_Y(t_2) = G_Y(t_2) \left[\lambda_1 \lambda_3 e^{\lambda_3 (t_2 - 1)} + \lambda_2 \right]$$

$$P(Y = 1) = \frac{\frac{d}{dt} G_Y(t_2)|_{t_2=0}}{1!} = G_Y(0) \left[\lambda_1 \lambda_3 e^{-\lambda_3} + \lambda_2 \right]. \tag{1.1.12}$$

Similarly, $P(Y = 2)$ is given as

$$\frac{d^2}{dt^2} G_Y(t_2) = G_Y(t_2) \left[\left(\lambda_1 \lambda_2 e^{\lambda_2 (t_2 - 1)} + \lambda_2 \right)^2 + \lambda_1 \lambda_3^2 e^{\lambda_3 (t_2 - 1)} \right]$$

$$P(Y=2)=\frac{\frac{d}{dt}G_Y(t_2)|_{t_2=0}}{2!}=\frac{G_Y(0)}{2!}\left[\left(\lambda_1\lambda_2e^{-\lambda_2}+\lambda_2\right)^2+\lambda_1\lambda_3^2e^{-\lambda_3}\right] \tag{1.1.13}$$

and finally $P(Y=3)$ is given by

$$\begin{aligned}\frac{d^3}{dt^3}G_Y(t_2)= \quad & G_Y(t_2)\Big[\lambda_1\lambda_3((\lambda_1\lambda_3e^{\lambda_3(t_2-1)}+\lambda_2)^2+\lambda_3\\ & (2(\lambda_1\lambda_3e^{\lambda_3(t_2-1)}+\lambda_2)+\lambda_3(1+\lambda_1e^{\lambda_3(t_2-1)})))e^{\lambda_3(t_2-1)}+\\ & \lambda_2((\lambda_1\lambda_3e^{\lambda_3(t_2-1)}+\lambda_2)^2+\lambda_1\lambda_3^2e^{\lambda_3(t_2-1)})\Big]\end{aligned} \tag{1.1.14}$$

$$\begin{aligned}P(Y=3)=\frac{1}{3!}\frac{d^3}{dt^3}G_{X_2}(0)|_{t_2=0}= \quad & \frac{G_Y(0)}{6}\Big[\lambda_1\lambda_3((\lambda_1\lambda_3e^{-\lambda_3}+\lambda_2)^2+\lambda_3\\ & (2(\lambda_1\lambda_3e^{-\lambda_3}+\lambda_2)+\lambda_3(1+\lambda_1e^{-\lambda_3})))e^{-\lambda_3}+\\ & \lambda_2((\lambda_1\lambda_3e^{-\lambda_3}+\lambda_2)^2+\lambda_1\lambda_3^2e^{-\lambda_3})\Big].\end{aligned} \tag{1.1.15}$$

On a similar line, one can extend the above procedure to get albeit complicated values for the probability that Y assumes any positive value.

1.1.4 Other conditional distribution of the bivariate pseudo-Poisson

In general, in another conditional distribution of the pseudo-Poisson model, i.e., X given $Y=y$, the derivation is theoretically ambiguous. Still, for the Sub-Model II, i.e., $\lambda_2=0$, we can derive the conditional distribution. In the following, we are deriving other conditional distributions, i.e., the conditional distribution of X given $Y=y$ by induction for the Sub-model II. Consider the joint mass function of pseudo-Poisson Sub-model II

$$P(X=x,Y=y)=\begin{cases}\frac{e^{-\lambda_1}\lambda_1^x}{x!}\frac{e^{-\lambda_3x}(\lambda_3x)^y}{y!}; & x=1,2,...;y=0,1,2,...\\ e^{-\lambda_1}; & (x,y)=(0,0).\end{cases}$$

Now, consider the case in which $y = 0$ then for each $x = 0, 1, 2, ...$ the conditional mass function will be

$$\begin{aligned} P_{X|Y}(x|0) &= \frac{P(X = x, Y = 0)}{P(Y = 0)} \\ &= \frac{e^{-\lambda_1 e^{-\lambda_3}} (\lambda_1 e^{-\lambda_3})^x}{x!} \end{aligned} \tag{1.1.16}$$

Indeed the above conditional mass function is a Poisson distribution with mean equal to $\lambda_1 e^{-\lambda_3}$.

Next, consider the case with $y = 1$. For each $x = 1, 2, ...$, we have

$$\begin{aligned} P_{X|Y}(x|1) &= \frac{P(X = x, Y = 1)}{P(Y = 1)} \\ &= \frac{e^{-\lambda_1 e^{-\lambda_3}} (\lambda_1 e^{-\lambda_3})^{x-1}}{(x-1)!} \end{aligned} \tag{1.1.17}$$

which is recognizable as the distribution of 1 plus a Poisson($\lambda_1 e^{-\lambda_3}$).

For $y \geq 1$ and for each $x = 1, 2, ...$, we have a

$$\begin{aligned} P_{X|Y}(x|y) &= \frac{P(X = x, Y = y)}{P(Y = y)} \\ &= \frac{\frac{e^{-\lambda_1 e^{-\lambda_3}} (\lambda_1 e^{-\lambda_3})^{x-1} x^y}{(x-1)!}}{\mu_y} \end{aligned} \tag{1.1.18}$$

where μ_y is the yth moment of a Poisson($\lambda_1 e^{-\lambda_3}$) variable. Note that the expression μ_y can also be expressed in terms of factorial moments, and the yth factorial moment is $(\lambda_1 e^{-\lambda_3})^y$. Thus we have

$$\mu_y = \sum_{j=0}^{y} S(y, j)(\lambda_1 e^{-\lambda_3})^y \tag{1.1.19}$$

where $S(y, j)$ is a Stirling number of the second kind. Also note that if $y \geq 1$ then $S(y, 0) = 0$.

For the detailed discussion on derivation, characterization, distributional features, and inferential aspects of bivariate pseudo-Poisson, refer Arnold and Manjunath [4].

Finally, we refer recent article by Arnold and Manjunath [5] on bivariate pseudo-Poisson with concomitant variables explored distributional and inferential aspects and an example of application to real-life data.

1.2 Other bivariate count distribution with Poisson structure

In the following, we briefly introduce the other bivariate count distribution having a Poisson structure.

1.2.1 Classical Bivariate Poisson

Consider three independent Poisson random variables $X_1 \sim \mathscr{P}$, $X_2 \sim \mathscr{P}$ and $X_3 \sim \mathscr{P}$. Then, the random sum $X = X_1 + X_3$ and $Y = X_2 + X_3$ is said to follow a classical bivariate Poisson distribution with the mass function

$$P(X = x, Y = y) = \exp\big(-(\lambda_1 + \lambda_2 + \lambda_3)\big)\frac{\lambda_1^x \lambda_2^y}{x!y!}\sum_{i=1}^{min\{x,y\}} \binom{x}{i}\binom{y}{i} i \Big(\frac{\lambda_3}{\lambda_1 + \lambda_2}\Big)^i \quad (1.2.1)$$

where $\lambda_i > 0$, $i = 1, 2, 3$, $x = 0, 1, 2, ...$ and $y = 0, 1, 2, ...$ For the bivariate mass function only marginals follow Poisson distribution.

1.2.2 Both conditional Poisson

Consider the two conditional mass function

$$X|Y = y \sim \mathscr{P}(\lambda_1 \lambda_3^y) \text{ for each } Y = y, \quad (1.2.2)$$
$$Y|X = x \sim \mathscr{P}(\lambda_2 \lambda_3^x) \text{ for each } X = x. \quad (1.2.3)$$

According to the Theorem, 4.1 by Arnold et al. [3], the joint mass function of X and Y will be

$$P(X = x, Y = y) = \kappa(\lambda_1, \lambda_2, \lambda_3)\frac{\lambda_1^x \lambda_2^y \lambda_3^{xy}}{x!y!}. \tag{1.2.4}$$

where $\lambda_i > 0$, $i = 1, 2, 3$, $x = 0, 1, 2, ...$ and $y = 0, 1, 2, ...$ Note that if $\lambda_3 = 1$, then X and Y are independent. Here both conditionals are Poisson.

1.2.3 Bivariate Conway-Maxwell Poisson

Note that the above two mass functions defined are stricken to equi-dispersed data; either we assume marginals or conditionals. For flexible bivariate count model which can handle over, under or equi-dispersion is one defined in Seller et al. [33], i.e., bivariate Conway-Maxwell Poisson. The joint probability generating function is

$$\pi(t_1, t_2) = \sum_{n=0}^{\infty} \frac{\lambda^n}{(n!)^\nu Z(\lambda, \nu)} \pi(t_1, t_2|n) \tag{1.2.5}$$

where

$$\pi(t_1, t_2|n) = 1 + p_{1+}(t_1 - 1) + p_{+1}(t_2 - 1) + p_{11}(t_1 -)(t_2 - 1)^n \tag{1.2.6}$$

and

$$Z(\lambda, \nu) = \sum_{s=0}^{\infty} \frac{\lambda^s}{(s!)^\nu} \tag{1.2.7}$$

where ν is a dispersion parameter with $\nu = 1$ is equi-dispersion, $\nu > 1$ is over dispersion and $\nu < 1$ is under dispersion.

CHAPTER 2

Bayesian Inference for pseudo-Poisson data

2.1 Introduction

Conditionally specified bivariate models often provide helpful, flexible models exhibiting a variety of dependence structures. However, for bivariate count data, the unique conditionally specified Poisson distribution, which turns out to be Obrechkoff's bivariate model [32], is often inappropriate since its dependence structure is often not felt to be appropriate. In such circumstances, attention can be diverted to consider what are known as pseudo-Poisson models (advocated strongly by Filus, Filus, and Arnold [15]). We begin by reviewing the pseudo-Poisson construction, highlighting certain simplified sub-models. Bayesian inference for these sub-models is then discussed. We refer to Arnold and Manjunath [4] for classical inferential aspects and also an example of applications of the bivariate pseudo-Poisson model. Bayesian inference for a particular classical bivariate Poisson model with Poisson marginals is discussed by Karlis and Tsiamyrtzis [21]. Finally, we note that Ghosh et al. [16] present recent results on bivariate count models with both conditionals being of the Poisson form.

2.2 Bivariate pseudo-Poisson models

The mass function for a general bivariate pseudo-Poisson distribution discussed in Chapter 1 Section 1.1 (Definition 1.1) is of the following form:

$$P(X=x,Y=y)=\frac{e^{-\lambda_1}\lambda_1^x}{x!}\frac{e^{-h(x,\underline{\theta})}(h(x,\underline{\theta}))^y}{y!}I(x\in\{0,1,2,...\},y\in\{0,1,2,...\}) \quad (2.2.1)$$

where the positive function $h(x,\underline{\theta})$ is quite arbitrary and $\underline{\theta}$ is parameter vector. The sub model in which h is a linear function involving two dependence parameters will be the focus of the present this chapter. Using λs to denote the three parameters in this model, its mass function is of the form

$$P(X=x,Y=y)=\frac{e^{-\lambda_1}\lambda_1^x}{x!}\frac{e^{-\lambda_2-\lambda_3 x}(\lambda_2+\lambda_3 x)^y}{y!}I(x\in\{0,1,2,...\},y\in\{0,1,2,...\}), \quad (2.2.2)$$

where $\lambda_1>0,\lambda_2\geq 0,\lambda_3>0$. Consequently the likelihood function for a sample of size n from this distribution is given by

$$L(\lambda_1,\lambda_2,\lambda_3;(x_1,y_1),...,(x_n,y_n))=\frac{e^{-n\lambda_1}\lambda_1^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!}\frac{e^{-n\lambda_2-\lambda_3\sum_{i=1}^n x_i}\prod_{i=1}^n(\lambda_2+\lambda_3 x_i)^{y_i}}{\prod_{i=1}^n y_i!}. \quad (2.2.3)$$

Note that this likelihood factors as follows:

$$L(\lambda_1,\lambda_2,\lambda_3;(x_1,y_1),...,(x_n,y_n))=\left\{\frac{e^{-n\lambda_1}\lambda_1^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!}\right\}\left\{\frac{e^{-n\lambda_2-\lambda_3\sum_{i=1}^n x_i}\prod_{i=1}^n(\lambda_2+\lambda_3 x_i)^{y_i}}{\prod_{i=1}^n y_i!}\right\}. \quad (2.2.4)$$

The first factor only involves the parameter λ_1, while the second factor only involves the parameters λ_2 and λ_3. We will call λ_1 the marginal parameter, while λ_2 and λ_3 will be called conditional parameters. This factorization will be important in Bayesian inference for the model, as discussed below. Because of the factorization, we will know that if a priori $\tilde{\lambda}_1$ and $(\tilde{\lambda}_2,\tilde{\lambda}_3)$ are independent, then they will be independent a posteriori also.

This feature of separation of marginal and conditional parameters continues to occur for more general pseudo-$\mathscr{F}$ models defined as follows. Let $\mathscr{F} = \{F(x;\underline{\theta}) : \underline{\theta} \in \Theta \subset \mathbb{R}^m\}$ be an m-parameter family of univariate distributions. A 2-dimensional pseudo-$\mathscr{F}$ distribution can be constructed as follows

$$P(X \leq x) = F(x;\underline{\theta}^{(1)}) \tag{2.2.5}$$

and for every x

$$P(Y \leq y|X = x) = F(y;\underline{\theta}^{(2)}(x,\underline{\tau})), \tag{2.2.6}$$

where $\underline{\theta}^{(1)} \in \Theta$ and, for each $\underline{\tau}$, $\underline{\theta}^{(2)}(x,\underline{\tau}) : \mathbb{R} \to \Theta$. In this setting, $\underline{\theta}_1$ are the marginal parameters, and $\underline{\tau}$ are the conditional parameters.

Returning to the pseudo-Poisson model defined in (2.2.2), it is natural to assume a gamma prior for the marginal parameter $\tilde{\lambda}_1$, i.e., a priori, and we assume that

$$\tilde{\lambda}_1 \sim \Gamma(\alpha_1, \delta_1),\ 0 < \alpha_1 < \infty, 0 < \delta_1 < \infty. \tag{2.2.7}$$

Here and subsequently, we parametrize a gamma distribution by a shape parameter δ_1 and an intensity parameter (the reciprocal of the scale parameter α_1). Specifically the a priori density for $\tilde{\lambda}_1$ is of the form

$$f_{\tilde{\lambda}_1}(\lambda_1) = \frac{\delta_1^{\alpha_1} \lambda_1^{\alpha_1 - 1} e^{-\delta_1 \lambda_1}}{\Gamma(\alpha_1)},\ 0 < \lambda_1 < \infty.$$

The choice of an a priori joint density for $(\tilde{\lambda}_2, \tilde{\lambda}_3)$ that will be independent of $\tilde{\lambda}_1$ is not so obvious. It is not clear whether it is possible to choose such a prior that will be "conjugate" with the second factor in (2.2.4). In general, for most choices of this prior, the joint posterior of $(\tilde{\lambda}_2, \tilde{\lambda}_3)$ will need to be dealt with numerically.

In the following section, we consider independent priors and also certain two-parameter sub-models for which the choice of independent gamma prior densities turns out to be conjugate.

2.3 Independent priors

Consider an a priori joint density in which all three parameters are independent, and each has a gamma distribution. Thus we have

$$f_{\tilde{\lambda}_1,\tilde{\lambda}_2,\tilde{\lambda}_3}(\lambda_1,\lambda_2,\lambda_3) = \prod_{i=1}^{3} \frac{\delta_i^{\alpha_i}\lambda_i^{\alpha_i-1}e^{-\delta_i\lambda_i}}{\Gamma(\alpha_i)},\ 0<\lambda_i,\alpha_i,\delta_i<\infty,\ i=1,2,3. \tag{2.3.1}$$

The kernel of the posterior, which is the product of the kernel of the factored likelihood (2.2.4) and the kernel of the prior (2.3.1) is of the form

$$\begin{aligned} ker(f_{\tilde{\lambda}_1,\tilde{\lambda}_2,\tilde{\lambda}_3|\underline{X},\underline{Y}}(\lambda_1,\lambda_2,\lambda_3|\underline{x},\underline{y})) &= \left\{e^{-(\delta_1+n)\lambda_1}\lambda_1^{\alpha_1+\sum_{i=1}^n x_i-1}\right\} \\ &\times \left\{e^{-(\delta_2+n)\lambda_2-(\delta_3+\sum_{i=1}^n x_i)\lambda_3}\lambda_2^{\alpha_2-1}\lambda_3^{\alpha_3-1}\prod_{i=1}^{n}(\lambda_2+\lambda_3 x_i)^{y_i}\right\} \end{aligned} \tag{2.3.2}$$

From the first factor in (2.3.2) we recognize that a posteriori $\tilde{\lambda}_1$ has a gamma distribution, i.e.,

$$\tilde{\lambda}_1|\underline{X}=\underline{x},\underline{Y}=\underline{y} \sim \Gamma(\alpha_1+\sum_{i=1}^{n}x_i,\delta_1+n). \tag{2.3.3}$$

The second factor is the kernel of the posterior distribution of $(\tilde{\lambda}_2,\tilde{\lambda}_3)$. It will need to be dealt with numerically.

2.3.1 Sub-model I

There are two simple sub-models of the linear model (2.2.2) that merit consideration because of their simplicity while retaining dependence between X and Y. We first focus on the sub-model of (2.2.2) obtained by equating λ_2 and λ_3. The model is thus of the form

$$P(X=x,Y=y) = \frac{e^{-\lambda_1}\lambda_1^x}{x!}\frac{e^{-\lambda_3(1+x)}\lambda_3^y(1+x)^y}{y!}. \tag{2.3.4}$$

The likelihood function for a sample of size n from this distribution is given by

$$L(\lambda_1, \lambda_3; (x_1, y_1), ..., (x_n, y_n)) = \frac{e^{-n\lambda_1}\lambda_1^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!} \frac{e^{-n\lambda_3} e^{-\lambda_3 \sum_{i=1}^n x_i} \prod_{i=1}^n (1+x_i)^{y_i} \lambda_3^{\sum_{i=1}^n y_i}}{\prod_{i=1}^n y_i!}. \tag{2.3.5}$$

Here too, the likelihood factors are as follows:

$$L(\lambda_1, \lambda_3; (x_1, y_1), ..., (x_n, y_n)) = \left\{ \frac{e^{-n\lambda_1}\lambda_1^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!} \right\} \left\{ \frac{e^{-n\lambda_3} e^{-\lambda_3 \sum_{i=1}^n x_i} \prod_{i=1}^n (1+x_i)^{y_i} \lambda_3^{\sum_{i=1}^n y_i}}{\prod_{i=1}^n y_i!} \right\}. \tag{2.3.6}$$

The first factor only involves the parameter λ_1, while the second factor only involves the parameter λ_3. We will call λ_1 the marginal parameter, while λ_3 will be called the conditional parameter.

It will be observed that the joint density of (X, Y) in (2.2.2) constitutes a two-parameter exponential family. A prior conjugate density consequently exists. For such a conjugate joint prior density for the two parameters in the model, we can take one with independent gamma marginals. Thus

$$f_{\tilde{\lambda}_1, \tilde{\lambda}_3}(\lambda_1, \lambda_3) = \frac{\delta_1^{\alpha_1} \lambda_1^{\alpha_1 - 1} e^{-\delta_1 \lambda_1}}{\Gamma(\alpha_1)} \frac{\delta_3^{\alpha_3} \lambda_3^{\alpha_3 - 1} e^{-\delta_3 \lambda_3}}{\Gamma(\alpha_3)}. \tag{2.3.7}$$

The kernel of the posterior, which is the product of the kernel of the factored likelihood (2.3.6) and the kernel of the prior (2.3.7), is of the form

$$\begin{aligned} ker(f_{\tilde{\lambda}_1, \tilde{\lambda}_3 | \underline{X}, \underline{Y}}(\lambda_1, \lambda_3 | \underline{x}, \underline{y})) &= \left\{ e^{-(\delta_1 + n)\lambda_1} \lambda_1^{\alpha_1 + \sum_{i=1}^n x_i - 1} \right\} \\ &\times \left\{ e^{-(\delta_3 + n + \sum_{i=1}^n x_i)\lambda_3} \lambda_3^{\alpha_3 + \sum_{i=1}^n y_i - 1} \right\}. \end{aligned} \tag{2.3.8}$$

For examples of such prior and posterior densities, see Figure 2.1 and 2.2. We have thus confirmed that our choice of a prior with independent gamma marginals was a conjugate prior for the likelihood (2.3.6) and that a posteriori, the two parameters have independent gamma distributions. Thus

$$\tilde{\lambda}_1 | \underline{X} = \underline{x}, \underline{Y} = \underline{y} \sim \Gamma(\alpha_1 + \sum_{i=1}^n x_i, \delta_1 + n), \tag{2.3.9}$$

and, independently,

$$\tilde{\lambda}_3 | \underline{X} = \underline{x}, \underline{Y} = \underline{y} \sim \Gamma(\alpha_3 + \sum_{i=1}^{n} y_i, \delta_3 + n + \sum_{i=1}^{n} x_i). \tag{2.3.10}$$

The squared error loss estimates of the parameters (the posterior means) are thus

$$\hat{\lambda_1}^{(B)} = \frac{\alpha_1 + \sum_{i=1}^{n} x_i}{\delta_1 + n} \tag{2.3.11}$$

and

$$\hat{\lambda_3}^{(B)} = \frac{\alpha_3 + \sum_{i=1}^{n} y_i}{\delta_3 + n + \sum_{i=1}^{n} x_i}. \tag{2.3.12}$$

If we choose to use an improper prior with $\alpha_1 = \alpha_3 = \delta_1 = \delta_3 = 0$ then the resulting Bayes estimates coincide with the corresponding maximum likelihood estimates.

If we use an improper prior with $\alpha_1 = \alpha_3 = \delta_1 = \delta_3 = 0$, then the resulting Bayes estimates coincide with the corresponding maximum likelihood estimates.

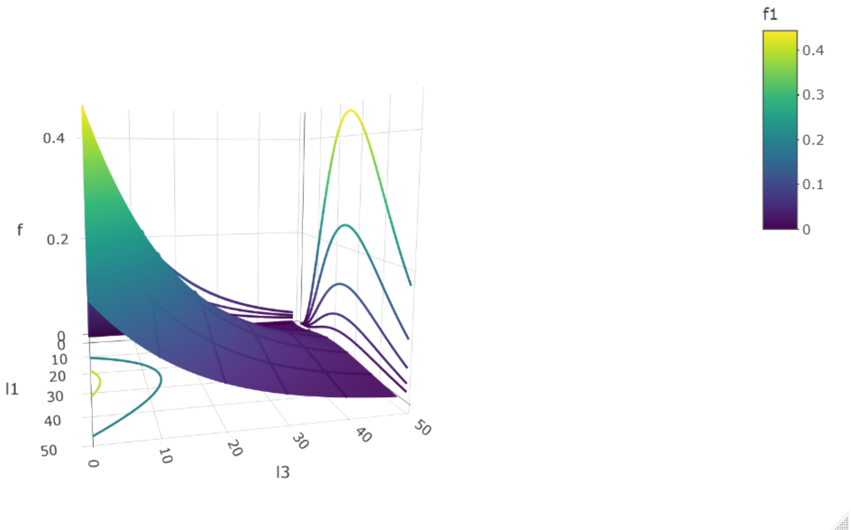


Figure 2.1: Density plot of prior (independent gamma) with parameter values $\alpha_1 = 1$, $\alpha_3 = 4$, $\delta_1 = 1$, $\delta_3 = 2$ of the Sub-model I.

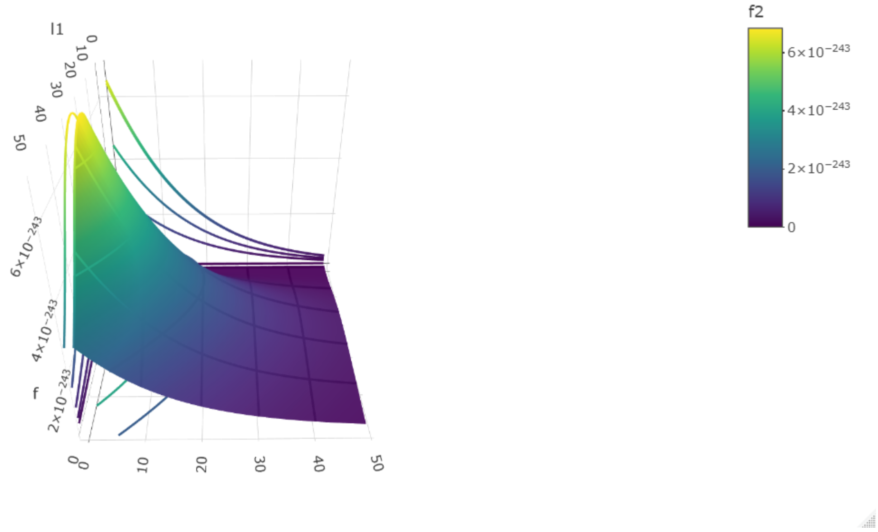


Figure 2.2: Density plot of posterior (independent gamma prior) with parameter values $\alpha_1 = 1$, $\alpha_3 = 4$, $\delta_1 = 1$, $\delta_3 = 2$, $\lambda_1 = 1$, $\lambda_3 = 4$ and $n = 20$ of the Sub-model I.

2.3.2 Sub-model II

Consider the sub-model obtained by setting $\lambda_2 = 0$. The model is thus of the form

$$P(X = x, Y = y) = \begin{cases} \frac{e^{-\lambda_1}\lambda_1^x}{x!}\frac{e^{-\lambda_3 x}\lambda_3^y x^y}{y!}; & x = 1, 2, ..., \quad y = 0, 1, 2, ... \\ e^{-\lambda_1}; & (x, y) = (0, 0) \\ 0; & otherwise \end{cases}$$

The likelihood function for a sample of size the following expression gives n from this distribution, in which n^* denotes the number of observed values of x equal to 0.

$$L(\lambda_1, \lambda_3; (x_1, y_1), ..., (x_n, y_n)) = e^{-n^*\lambda_1}\frac{e^{-(n-n^*)\lambda_1}\lambda_1^{\sum_{x_i>0} x_i}}{\prod_{x_i>0} x_i!}\frac{e^{-\lambda_3\sum_{x_i>0} x_i}[\prod_{x_i>0}(x_i)^{y_i}]\lambda_3^{\sum_{x_i>0} y_i}}{\prod_{x_i>0} y_i!}. \tag{2.3.13}$$

Here too, the likelihood factors are as follows

$$L(\lambda_1, \lambda_3; (x_1, y_1), ..., (x_n, y_n)) = \left\{\frac{e^{-n\lambda_1}\lambda_1^{\sum_{x_i>0} x_i}}{\prod_{x_i>0} x_i!}\right\}\left\{\frac{e^{-\lambda_3\sum_{x_i>0} x_i}[\prod_{x_i>0}(x_i)^{y_i}]\lambda_3^{\sum_{x_i>0} y_i}}{\prod_{x_i>0} y_i!}\right\}. \tag{2.3.14}$$

The first factor only involves the parameter λ_1, while the second factor only involves the parameter λ_3. Again, we will call λ_1 the marginal parameter, while λ_3. will be called the conditional parameter.

The kernel of the likelihood is given by

$$L(\lambda_1, \lambda_3; (x_1, y_1), ..., (x_n, y_n)) \propto e^{-n\lambda_1} \lambda_1^{\sum_{x_i>0} x_i} e^{-\lambda_3 \sum_{x_i>0} x_i} \lambda_3^{\sum_{x_i>0} y_i}. \tag{2.3.15}$$

If we use a prior with independent gamma marginals, such as

$$f_{\tilde{\lambda}_1, \tilde{\lambda}_3}(\lambda_1, \lambda_3) = \frac{\delta_1^{\alpha_1} \lambda_1^{\alpha_1 - 1} e^{-\delta_1 \lambda_1}}{\Gamma(\alpha_1)} \frac{\delta_3^{\alpha_3} \lambda_3^{\alpha_3 - 1} e^{-\delta_3 \lambda_3}}{\Gamma(\alpha_3)}, \tag{2.3.16}$$

with kernel of the form

$$f_{\tilde{\lambda}_1, \tilde{\lambda}_3}(\lambda_1, \lambda_3) \propto \lambda_1^{\alpha_1 - 1} e^{-\delta_1 \lambda_1} \lambda_3^{\alpha_3 - 1} e^{-\delta_3 \lambda_3}, \tag{2.3.17}$$

then the kernel of the posterior, which is the product of the kernel of the factored likelihood, and the kernel of the prior is of the form

$$\begin{aligned} ker(f_{\tilde{\lambda}_1, \tilde{\lambda}_3 | \underline{X}, \underline{Y}}(\lambda_1, \lambda_3 | \underline{x}, \underline{y}) &= \left\{ e^{-(\delta_1 + n)\lambda_1} \lambda_1^{\alpha_1 + \sum_{x_i>0} x_i - 1} \right\} \\ &\times \left\{ e^{-(\delta_3 + \sum_{x_i>0} x_i)\lambda_3} \lambda_3^{\alpha_3 + \sum_{x_i>0} y_i - 1} \right\}. \end{aligned}$$

For an example of such posterior densities, see Figure 2.3. We have confirmed that our choice of a prior with independent gamma marginals is a conjugate prior for the likelihood and yields a posterior density which also has independent gamma marginals.

$$\tilde{\lambda}_1 | \underline{X} = \underline{x}, \underline{Y} = \underline{y} \sim \Gamma(\alpha_1 + \sum_{x_i>0} x_i, \delta_1 + n) \tag{2.3.18}$$

and, independently,

$$\tilde{\lambda}_3 | \underline{X} = \underline{x}, \underline{Y} = \underline{y} \sim \Gamma(\alpha_3 + \sum_{x_i>0} y_i, \delta_3 + \sum_{x_i>0} x_i). \tag{2.3.19}$$

The squared error loss estimates of the parameters (the posterior means) are thus

$$\hat{\lambda_1}^{(B)} = \frac{\alpha_1 + \sum_{x_i>0} x_i}{\delta_1 + n} \tag{2.3.20}$$

and

$$\hat{\lambda_3}^{(B)} = \frac{\alpha_3 + \sum_{x_i>0} y_i}{\delta_3 + \sum_{x_i>0} x_i} \tag{2.3.21}$$

Note that, since $y_i = 0$ whenever $x_i = 0$, we can simplify the above expressions by replacing $\sum_{x_i>0} x_i$ by $\sum_{i=1}^{n} x_i$ and replacing $\sum_{x_i>0} y_i$ by $\sum_{i=1}^{n} y_i$

If we choose to use an improper prior with $\alpha_1 = \alpha_3 = \delta_1 = \delta_3 = 0$ then the resulting Bayes estimates coincide with the corresponding maximum likelihood estimates.

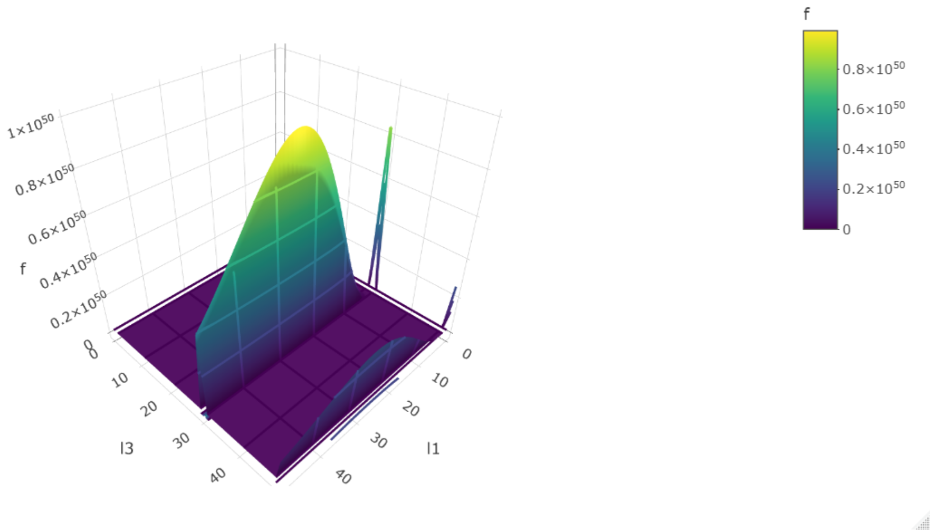


Figure 2.3: Density plot of posterior (independent gamma prior) with parameter values $\alpha_1 = 1$, $\alpha_3 = 4$, $\delta_1 = 1$, $\delta_3 = 2$, $\lambda_1 = 1$, $\lambda_3 = 4$ and $n = 20$ of the Sub-model II.

Remark 3. *It should be noted that the model with $\lambda_2 = 0$ will only be appropriate for data sets for which $Y = 0$ whenever $X = 0$.*

2.4 Pseudo-gamma priors

In this section, we introduce a bivariate pseudo-gamma prior, which allows us to incorporate external information on the dependence between the parameters. Note that the hyperparameter ψ_3(say) in the pseudo-gamma prior plays a crucial role in defining a priori dependence between λ_1 and λ_3. Also, note that Sub-models I & II are the sub-families of all bivariate pseudo-Poisson distributions having severe dependence between the variables. Hence, the importance of considering the pseudo-gamma prior to developing posterior inference. To reduce the computational complexity and have closed-form expression of marginals, we shall consider some special cases of pseudo-gamma distribution as priors. For the classical bivariate Poisson distribution, a discussion of the importance of considering dependence in the prior and its application can be found in Karlis, and Tsiamyrtzis [21]. However, the prior proposed here, i.e., bivariate pseudo-gamma prior, is very simple because of its marginal and conditional composition. Hence, analytical computations and simulation algorithms are easy to implement and render comprehensible its posterior properties.

Dubey [11] and Sen et.al.[34] provides a discussion of alternative bivariate gamma distributions. Also, reference can be made to Balakrishnan and Lai [6] for coverage of various methods of constructing bivariate continuous distributions. In addition, developments on similar lines to the bivariate pseudo-exponential distribution modeling and its applications can be found in Mohsin et. al. [27].

In the following sections, we will discuss bivariate pseudo-gamma priors and their applications in more detail for each sub-models I and II.

2.4.1 Sub-model I

Consider the sub model, specified in (2.3.4), the likelihood function for a sample of size n from this distribution is given by

$$L(\lambda_1, \lambda_3; (x_1, y_1), ..., (x_n, y_n)) = \frac{e^{-n\lambda_1}\lambda_1^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!} \frac{e^{-n\lambda_3} e^{-\lambda_3 \sum_{i=1}^{n} x_i} \prod_{i=1}^{n}(1+x_i)^{y_i} \lambda_3^{\sum_{i=1}^{n} y_i}}{\prod_{i=1}^{n} y_i!}. \tag{2.4.1}$$

This time we will consider a joint prior that is of the bivariate-pseudo-gamma form. For it we assume that λ_3 has a $\Gamma(\tau_1, \psi_1)$ density, i.e.,

$$f(\lambda_3) \propto \lambda_3^{\tau_1 - 1} e^{-\psi_1 \lambda_3} I(\lambda_3 > 0), 0 < \tau_1, \psi_1 < \infty,$$

and then for each value of λ_3, the conditional density of λ_1 given λ_3 is assumed to be of the gamma form with an intensity parameter that is a linear function of λ_3. Thus

$$f(\lambda_1|\lambda_3) \propto (\psi_2 + \psi_3\lambda_3)^{\tau_2} \lambda_1^{\tau_2 - 1} e^{-(\psi_2 + \psi_3\lambda_3)\lambda_1} I(\lambda_1 > 0), 0 < \tau_2, \psi_3 < \infty, 0 \leq \psi_2 < \infty.$$

The joint prior is thus of the form

$$f(\lambda_1, \lambda_3) \propto (\psi_2 + \psi_3\lambda_3)^{\tau_2} \lambda_1^{\tau_2 - 1} e^{-(\psi_2 + \psi_3\lambda_3)\lambda_1} \lambda_3^{\tau_1 - 1} e^{-\psi_1\lambda_3} I(\lambda_1 > 0) I(\lambda_3 > 0). \quad (2.4.2)$$

Consider the simpler prior in which we assume that $\psi_2 = 0$. This prior density will be of the form

$$f_p(\lambda_1, \lambda_3) \propto \lambda_1^{\tau_2 - 1} e^{-\psi_3\lambda_3\lambda_1} \lambda_3^{\tau_1 + \tau_2 - 1} e^{-\psi_1\lambda_3} I(\lambda_1 > 0) I(\lambda_3 > 0). \quad (2.4.3)$$

The corresponding posterior density corresponding to a sample of size n from sub-model I will be

$$\begin{aligned} f(\lambda_1, \lambda_3 | \underline{X} = \underline{x}, \underline{Y} = \underline{y})) &\propto [e^{-n\lambda_1} \lambda_1^{\sum_{i=1}^n x_i}][e^{-n\lambda_3} e^{-\lambda_3 \sum_{i=1}^n x_i} \lambda_3^{\sum_{i=1}^n y_i}] \\ &\quad \times [\lambda_1^{\tau_2 - 1} e^{-\psi_3\lambda_3\lambda_1} \lambda_3^{\tau_1 + \tau_2 - 1} e^{-\psi_1\lambda_3}] \\ &\propto \lambda_1^{\tau_2 + \sum_{i=1}^n x_i - 1} e^{-n\lambda_1} \lambda_3^{\tau_1 + \tau_2 + \sum_{i=1}^n y_i - 1} e^{-(n+\psi_1)\lambda_3} e^{-\psi_3\lambda_1\lambda_3} \\ &\propto \left[\lambda_3^{\tau_1 + \sum_{i=1}^n y_i - 1} e^{-(n+\psi_1)\lambda_3}\right] \left[\lambda_1^{\tau_2 + \sum_{i=1}^n x_i - 1} e^{-(n+\psi_3\lambda_3)\lambda_1}\right] \quad (2.4.4) \end{aligned}$$

For examples of such prior and posterior densities, see Figures 2.4, 2.5, 2.6 (for priors) and Figures 2.7, 2.8, 2.9(for posteriors). The marginal posterior distributions of λ_1 and λ_3 are

$$f^{P}_{\lambda_1}(\lambda_1) \;\propto\; \left[\lambda_1^{\tau_2+\sum_{i=1}^{n}x_i-1}e^{-n\lambda_1}\right]\frac{\Gamma(\tau_1+\sum_{i=1}^{n}y_i)}{(n+\psi_1+\psi_3\lambda_1)^{(\tau_1+\sum_{i=1}^{n}y_i)}}, \tag{2.4.5}$$

and

$$f^{P}_{\lambda_3}(\lambda_3) \;\propto\; \left[\lambda_3^{\tau_1+\sum_{i=1}^{n}y_i-1}e^{-(n+\psi_1)\lambda_3}\right]\frac{\Gamma(\tau_2+\sum_{i=1}^{n}x_i)}{(n+\psi_3\lambda_3)^{(\tau_2+\sum_{i=1}^{n}x_i)}}. \tag{2.4.6}$$

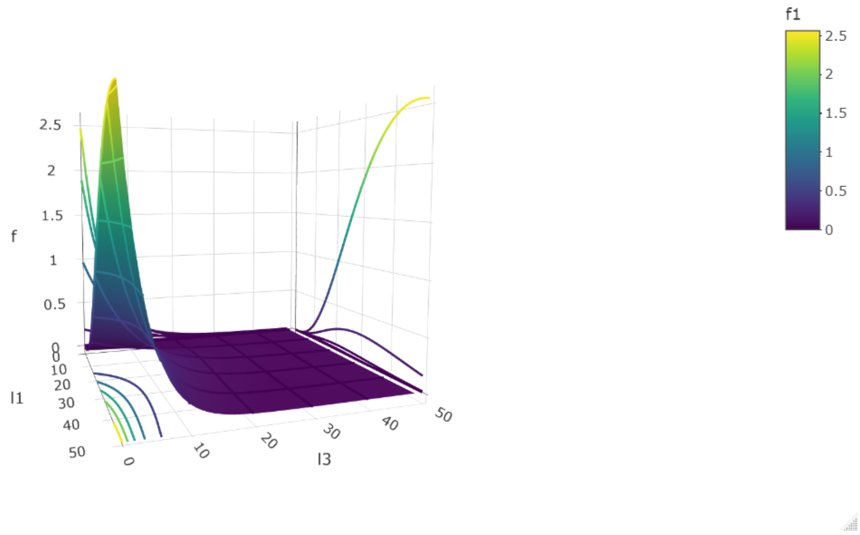


Figure 2.4: Density plot of prior (pseudo-gamma prior) with parameter values $\tau_1 = 1$, $\tau_2 = 4$, $\psi_1 = 1$, $\psi_2 = 1$(small value), $\psi_3 = 3$ of the Sub-model I.

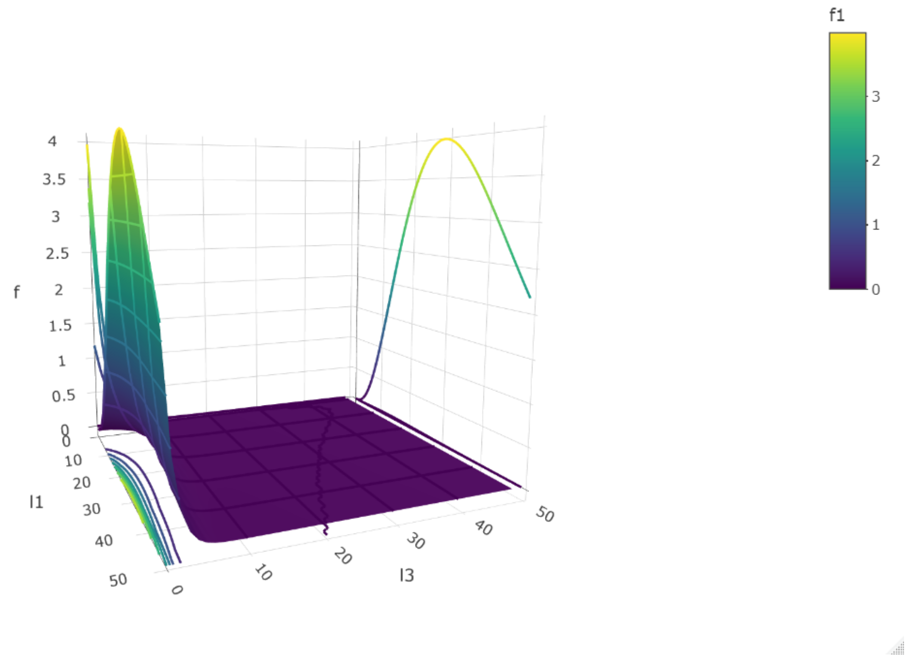


Figure 2.5: Density plot of prior (pseudo-gamma prior) with parameter values $\tau_1 = 1$, $\tau_2 = 4$, $\psi_1 = 1$, $\psi_2 = 5$ (large value), $\psi_3 = 3$ of the Sub-model I.

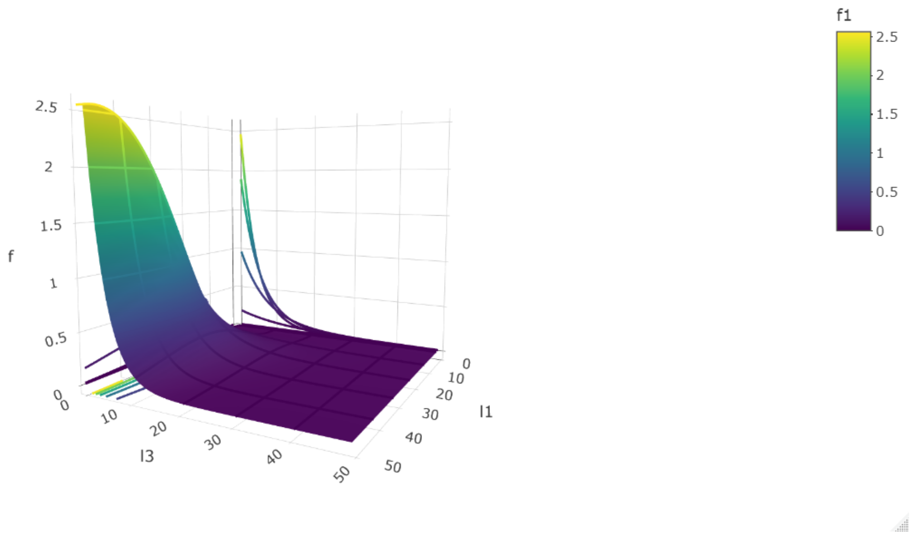


Figure 2.6: Density plot of prior (pseudo-gamma prior simple) with parameter values $\tau_1 = 1$, $\tau_2 = 4$, $\psi_1 = 1$, $\psi_2 = 0$ (simple), $\psi_3 = 3$ of the Sub-model I.

Note that the mean and variance of the marginals need to be dealt with numerically.

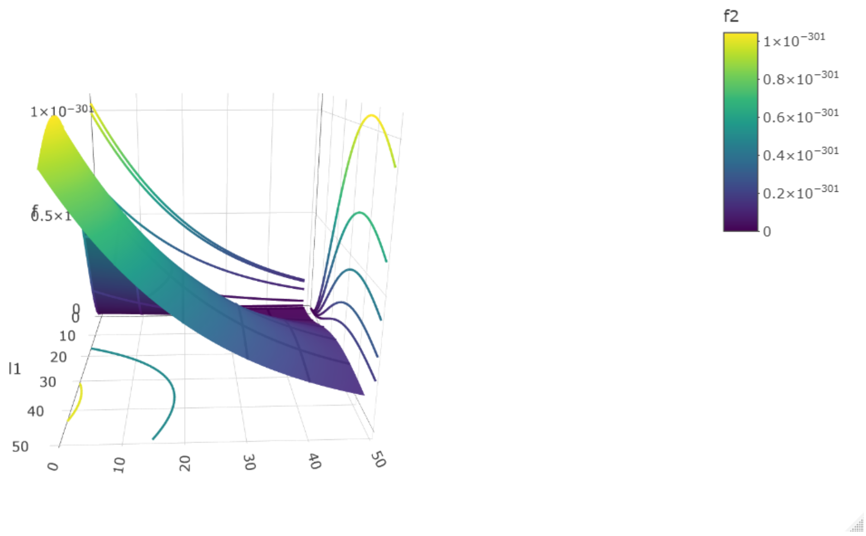


Figure 2.7: Density plot of posterior (pseudo-gamma prior) with parameter values $\tau_1 = 1$, $\tau_2 = 4$, $\psi_1 = 1$, $\psi_2 = 1$ (small value), $\psi_3 = 3$, $\lambda_1 = 1$, $\lambda_3 = 4$ and $n = 20$ of Sub-model I.

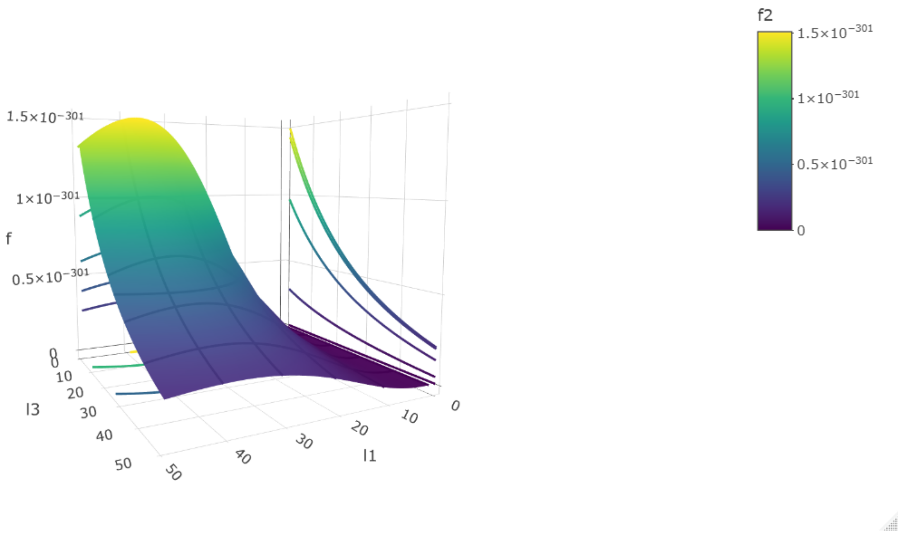


Figure 2.8: Density plot of posterior (pseudo-gamma prior) with parameter values $\tau_1 = 1$, $\tau_2 = 4$, $\psi_1 = 1$, $\psi_2 = 5$ (large value), $\psi_3 = 3$, $\lambda_1 = 1$, $\lambda_3 = 4$ and $n = 20$ of Sub-model I.

2.4.2 Sub-model II

Consider the sub-model obtained by setting $\lambda_2 = 0$. The model is thus of the form

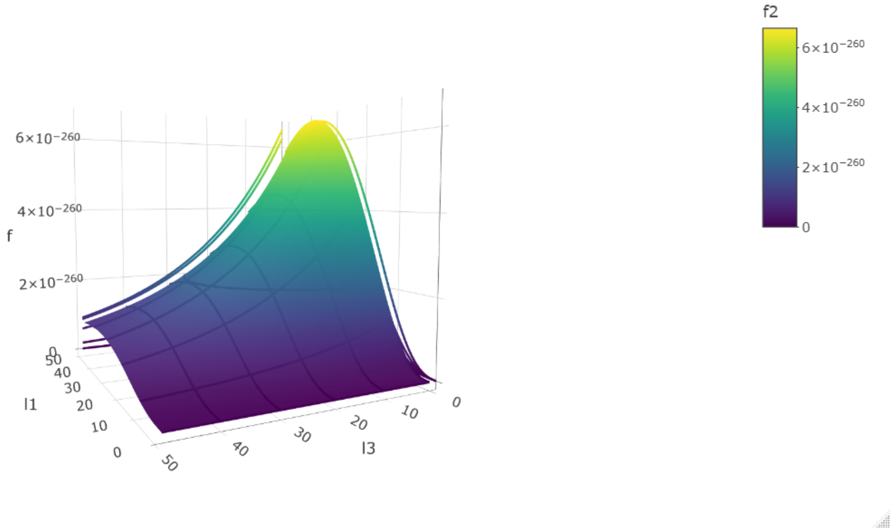


Figure 2.9: Density plot of posterior (pseudo-gamma prior simpler) with parameter values $\tau_1 = 1$, $\tau_2 = 4$, $\psi_1 = 1$, $\psi_2 = 0$ (simple), $\psi_3 = 3$, $\lambda_1 = 1$, $\lambda_3 = 4$ and $n = 20$ of the Sub-model I.

$$P(X = x, Y = y) = \begin{cases} \frac{e^{-\lambda_1}\lambda_1^x}{x!}\frac{e^{-\lambda_3 x}\lambda_3^y x^y}{y!}; & x = 1, 2, ..., \;\; y = 0, 1, 2, ... \\ e^{-\lambda_1}; & (x, y) = (0, 0) \\ 0; & otherwise \end{cases}$$

The likelihood function for a sample of size the following expression gives n from this distribution, in which n^* denotes the number of observed values of x equal to 0.

$$L(\lambda_1, \lambda_3; (x_1, y_1), ..., (x_n, y_n)) = e^{-n^*\lambda_1} \frac{e^{-(n-n^*)\lambda_1}\lambda_1^{\sum_{x_i>0} x_i}}{\prod_{x_i>0} x_i!} \frac{e^{-\lambda_3 \sum_{x_i>0} x_i}[\prod_{x_i>0}(x_i)^{y_i}]\lambda_3^{\sum_{x_i>0} y_i}}{\prod_{x_i>0} y_i!}. \tag{2.4.7}$$

Here too, the likelihood factors are as follows

$$L(\lambda_1, \lambda_3; (x_1, y_1), ..., (x_n, y_n)) = \left\{\frac{e^{-n\lambda_1}\lambda_1^{\sum_{x_i>0} x_i}}{\prod_{x_i>0} x_i!}\right\} \left\{\frac{e^{-\lambda_3 \sum_{x_i>0} x_i}[\prod_{x_i>0}(x_i)^{y_i}]\lambda_3^{\sum_{x_i>0} y_i}}{\prod_{x_i>0} y_i!}\right\}. \tag{2.4.8}$$

The first factor only involves the parameter λ_1, while the second factor only involves the parameter λ_3. Again, we will call λ_1 the marginal parameter, while λ_3 will be called the conditional parameter.

The kernel of the likelihood is given by

$$L(\lambda_1, \lambda_3) \propto e^{-n\lambda_1} \lambda_1^{\sum_{x_i>0} x_i} e^{-\lambda_3 \sum_{x_i>0} x_i} \lambda_3^{\sum_{x_i>0} y_i}. \tag{2.4.9}$$

Consider the simpler pseudo-gamma prior in which we assume that $\psi_2 = 0$. This prior density will be

$$f_p(\lambda_1, \lambda_3) \propto \lambda_1^{\tau_2 - 1} e^{-\psi_3 \lambda_3 \lambda_1} \lambda_3^{\tau_1 + \tau_2 - 1} e^{-\psi_1 \lambda_3} I(\lambda_1 > 0) I(\lambda_3 > 0). \tag{2.4.10}$$

The posterior density will be

$$\begin{aligned} f(\lambda_1, \lambda_3 | \underline{X} = \underline{x}, \underline{Y} = \underline{y}) \quad \propto \quad & e^{-n\lambda_1} \lambda_1^{\sum_{x_i>0} x_i} e^{-\lambda_3 \sum_{x_i>0} x_i} \lambda_3^{\sum_{x_i>0} y_i} \\ & \times \lambda_1^{\tau_2 - 1} e^{-\psi_3 \lambda_3 \lambda_1} \lambda_3^{\tau_1 + \tau_2 - 1} e^{-\psi_1 \lambda_3}. \end{aligned} \tag{2.4.11}$$

For examples of such posterior densities, see Figure 2.10, 2.11 and 2.12.

The marginal posterior distributions of λ_1 and λ_3 are

$$f_{\lambda_1}^{P^*}(\lambda_1) \quad \propto \quad \left[\lambda_1^{\tau_2 + \sum_{x_i>0} x_i - 1} e^{-n\lambda_1} \right] \frac{\Gamma(\tau_1 + \sum_{y_i>0} y_i)}{(\psi_1 + \psi_3 \lambda_1)^{(\tau_1 + \sum_{y_i>0} y_i)}}, \tag{2.4.12}$$

and

$$f_{\lambda_3}^{P^*}(\lambda_3) \quad \propto \quad \left[\lambda_3^{\tau_1 + \sum_{y_i>0} y_i - 1} e^{-(n+\psi_1)\lambda_3} \right] \frac{\Gamma(\tau_2 + \sum_{x_i>0} x_i)}{(n + \psi_3 \lambda_3)^{(\tau_2 + \sum_{x_i>0} x_i)}}. \tag{2.4.13}$$

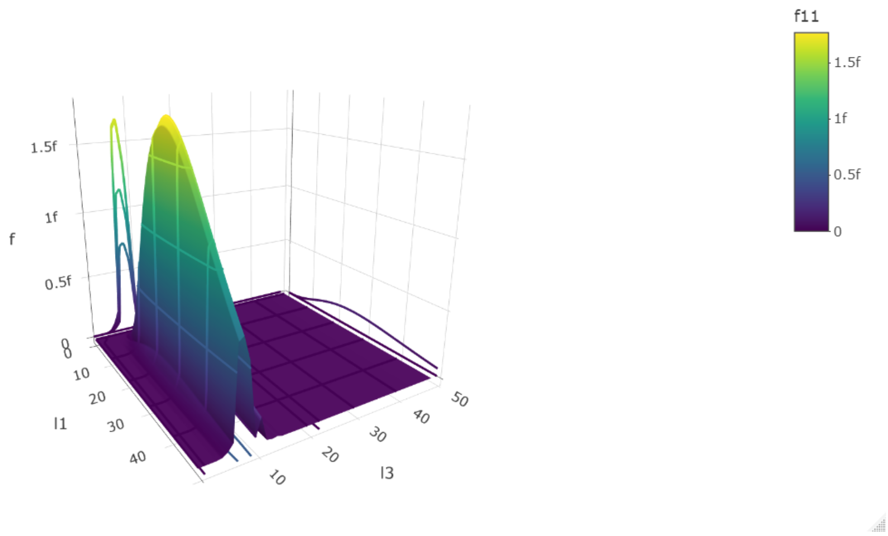


Figure 2.10: Density plot of posterior (pseudo-gamma prior) with parameter values $\tau_1 = 1$, $\tau_2 = 4$, $\psi_1 = 1$, $\psi_2 = 1$ (small value), $\psi_3 = 3$, $\lambda_1 = 1$, $\lambda_3 = 4$ and $n = 20$ of the Sub-model II.

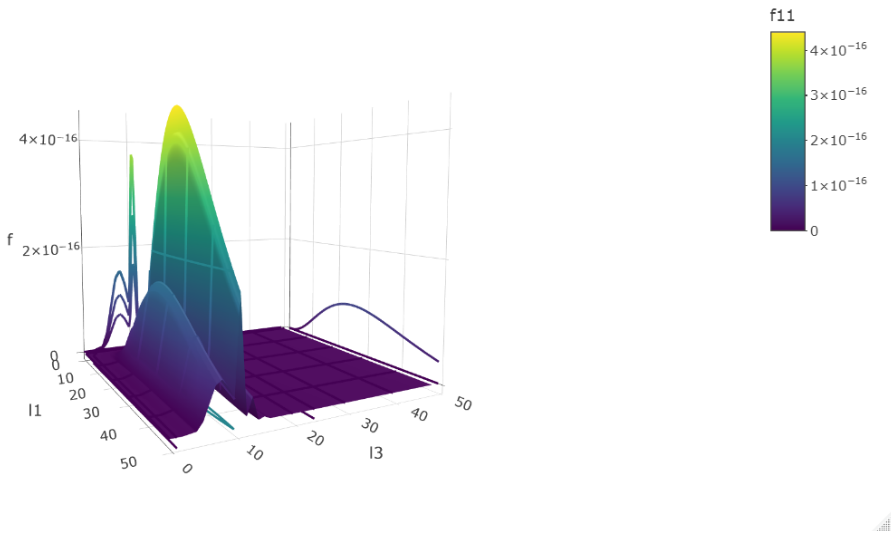


Figure 2.11: Density plot of posterior (pseudo-gamma prior) with parameter values $\tau_1 = 1$, $\tau_2 = 4$, $\psi_1 = 1$, $\psi_2 = 5$ (large value), $\psi_3 = 3$, $\lambda_1 = 1$, $\lambda_3 = 4$ and $n = 20$ of the Sub-model II.

Note that the mean and variance of the marginals need to be dealt with numerically.

Some comparisons of Bayesian posterior analyses using a variety of prior densities, including non-informative, independent gammas, and pseudo-gamma, are provided in the following section.

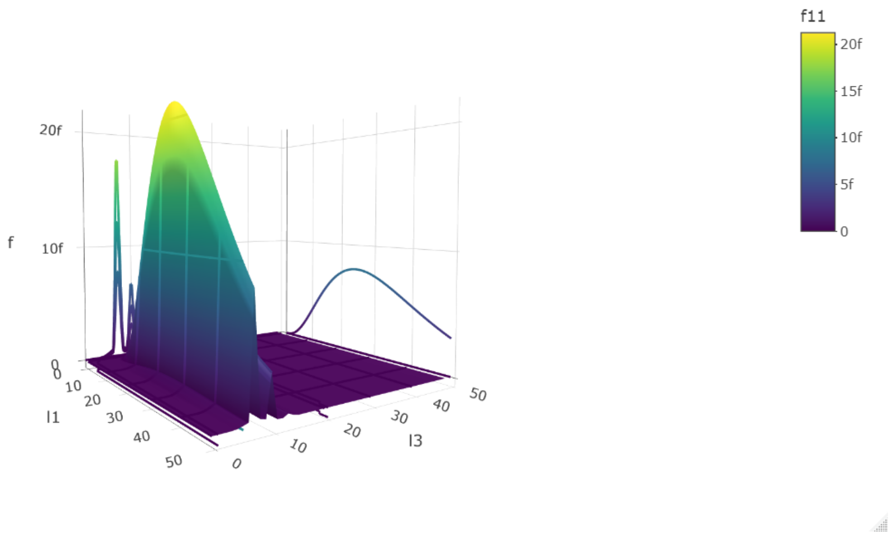


Figure 2.12: Density plot of posterior (pseudo-gamma prior simpler) with parameter values $\tau_1 = 1$, $\tau_2 = 4$, $\psi_1 = 1$, $\psi_2 = 0$ (simple), $\psi_3 = 3$, $\lambda_1 = 1$, $\lambda_3 = 4$ and $n = 20$ of the Sub-model II.

2.5 Examples

2.5.1 Simulation study

Due to the marginal and conditional compositions of the pseudo-Poisson distributions, simulation can be done in two steps: first, simulate x from $Poisson(\lambda_1)$ and next, simulate y from $Poisson(\lambda_2 + \lambda_3 x)$. However, inference for the posterior distributions is more complex for a complete model, even under independent gamma priors. We need to rely on numerical algorithms to compute marginal distributions and their moments. In the current work, we will use a Hit-And-Run Metropolis (HARM) algorithm to simulate observations from the posterior distributions under all priors (improper, independent, and pseudo) and for its full and sub-models. We refer to Chen [9] for the HARM algorithm implementation and its comparison with Gibbs and metropolis sampling. However, for each set of samples from the posterior densities, the convergence of the HARM algorithm is verified and assured. Although, due to the limitations in the content of the current work, all convergence plots are not included. However, one can refer to Hall [18] for all contemporary Bayesian simulation algorithms, including

HARM. Also, convergence plots of each of the algorithms can be verified in Laplaces-Demon R package [36].

We have simulated $10,000$ data sets using the HARM algorithm with thinning at the 10th sample from each posterior distribution of λ's with varying sample sizes from $n = 10, 20, 30, 50, 100, 500$. Mean and posteriori 95% confidence intervals are mentioned in Table 2.1 and 2.2 are computed from 1000 iterations of the aforementioned procedure.

Full-model : For the parameter values $\lambda_1 = 1$, $\lambda_2 = 3$ and $\lambda_3 = 4$, we consider the following priors

FP_1: Uniform prior (improper prior)

FP_2: Independent gamma prior, i.e., $\lambda_1 \sim \Gamma(\alpha_1, \delta_1)$, $\lambda_2 \sim \Gamma(\alpha_2, \alpha_2)$ and $\lambda_3 \sim \Gamma(\alpha_3, \delta_3)$ for the prior parameter values $\alpha_1 = 1, \delta_1 = 1, \alpha_2 = 3, \delta_2 = 1, \alpha_3 = 4$ and $\delta_3 = 2$.

The posterior density plots of λ_1, λ_2 and λ_3 with improper (Uniform prior) and independent gamma priors c.f. Figures 2.13, 2.14 and 2.15 [1]. The plot has been illustrated for one set of 100000 observations with thinning at 10th sample to understand the large sample behaviour of posterior distribution.

Sub-model I (i.e. when $\lambda_2 = \lambda_3$): for the parameter values $\lambda_1 = 1$ and $\lambda_3 = 4$, we consider the following priors:

SIP_1: Uniform prior (improper prior)

SIP_2: Independent gamma prior, i.e.,$\lambda_1 \sim \Gamma(\alpha_1, \delta_1)$ and $\lambda_3 \sim \Gamma(\alpha_3, \delta_3)$, for the prior parameter values $\alpha_1 = 1, \delta_1 = 1, \alpha_3 = 4$ and $\delta_3 = 2$.

SIP_3: Bivariate pseudo-gamma prior, i.e., $\lambda_3 \sim \Gamma(\tau_1, \psi_1)$ and $\lambda_1|\lambda_3 \sim \Gamma(\tau_2, (\psi_2 + \psi_3\lambda_3))$, for the prior parameter values $\tau_1 = 1, \psi_1 = 1, \tau_2 = 4$ $\psi_2 = 1, \psi_3 = 3$

SIP_4: Bivariate pseudo-gamma prior, i.e., $\lambda_3 \sim \Gamma(\tau_1, \psi_1)$ and $\lambda_1|\lambda_3 \sim \Gamma(\tau_2, (\psi_2 + \psi_3\lambda_3))$, for the prior parameter values $\tau_1 = 1, \psi_1 = 1, \tau_2 = 4$ $\psi_2 = 5, \psi_3 = 3$

[1] li_imp: **posteriori observations of λ_i by improper prior ;** li_ind: **posterior observations from λ_i by independent gamma prior, $i = 1, 2, 3$.**

SIP_5: Bivariate pseudo-gamma prior, i.e., $\lambda_3 \sim \Gamma(\tau_1, \psi_1)$ and $\lambda_1|\lambda_3 \sim \Gamma(\tau_2, (\psi_2 + \psi_3\lambda_3))$, for the prior parameter values $\tau_1 = 1$, $\psi_1 = 1$, $\tau_2 = 4$ $\psi_2 = 0$, $\psi_3 = 3$

We refer to Table 2.1 for the comparison of different prior effects on posteriori means distribution. Also, c.f. Figure 2.16 and 2.17 for posteriori density plots and the plots has been illustrated for one set of 100000 observations with thinning at 10th sample to understand the large sample behaviour of posteriori distribution.

Sub-Model II (i.e. when $\lambda_2 = 0$): for the parameter values $\lambda_1 = 1$ and $\lambda_3 = 4$, we consider the following priors:

$SIIP_1$: Uniform prior (improper prior)

$SIIP_2$: Independent gamma prior, i.e.,$\lambda_1 \sim \Gamma(\alpha_1, \delta_1)$ and $\lambda_3 \sim \Gamma(\alpha_3, \delta_3)$, for the parameter values $\alpha_1 = 1$, $\delta_1 = 1$, $\alpha_3 = 4$ and $\delta_3 = 2$.

$SIIP_3$: Bivariate pseudo-gamma prior, i.e., $\lambda_3 \sim \Gamma(\tau_1, \psi_1)$ and $\lambda_1|\lambda_3 \sim \Gamma(\tau_2, (\psi_2 + \psi_3\lambda_3))$, for the parameter values $\tau_1 = 1$, $\psi_1 = 1$, $\tau_2 = 4$ $\psi_2 = 1$, $\psi_3 = 3$

$SIIP_4$: Bivariate pseudo-gamma prior, i.e., $\lambda_3 \sim \Gamma(\tau_1, \psi_1)$ and $\lambda_1|\lambda_3 \sim \Gamma(\tau_2, (\psi_2 + \psi_3\lambda_3))$, for the parameter values $\tau_1 = 1$, $\psi_1 = 1$, $\tau_2 = 4$ $\psi_2 = 5$, $\psi_3 = 3$

$SIIP_5$: Bivariate pseudo-gamma prior, i.e., $\lambda_3 \sim \Gamma(\tau_1, \psi_1)$ and $\lambda_1|\lambda_3 \sim \Gamma(\tau_2, (\psi_2 + \psi_3\lambda_3))$, for the parameter values $\tau_1 = 1$, $\psi_1 = 1$, $\tau_2 = 4$ $\psi_2 = 0$, $\psi_3 = 3$

We refer to Table 2.2 for the comparison of different prior effects on posteriori. Also, refer Figure 2.18 and 2.19 for the posterior density plots and the plots has been illustrated for one set of 100000 observations with thinning at 10th sample to understand the large sample behaviour of posteriori distribution.

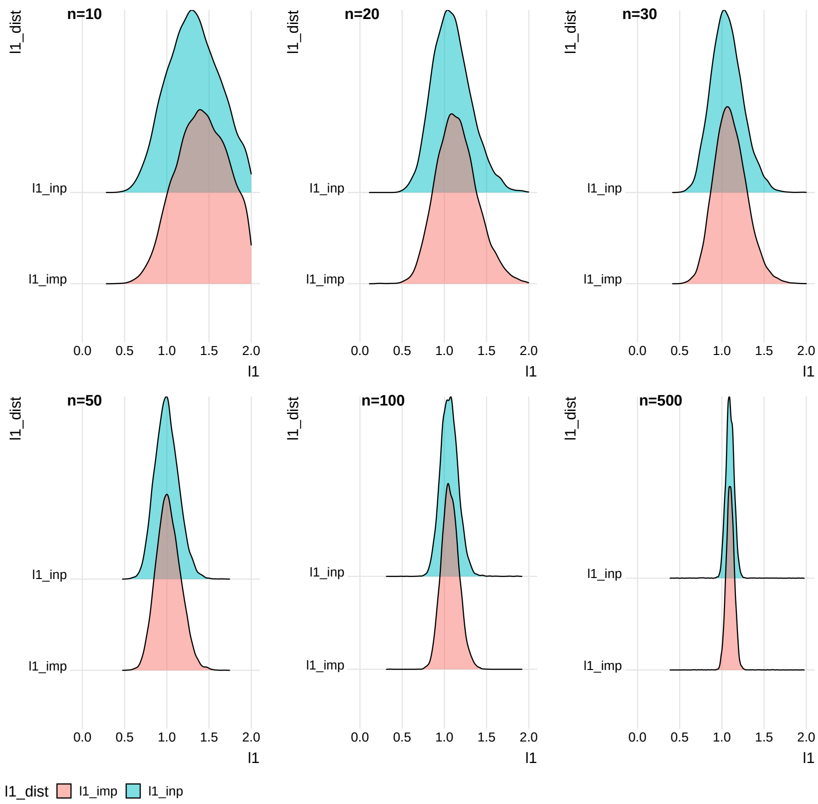


Figure 2.13: Posterior density plot of λ_1 (independent gamma and improper priors)

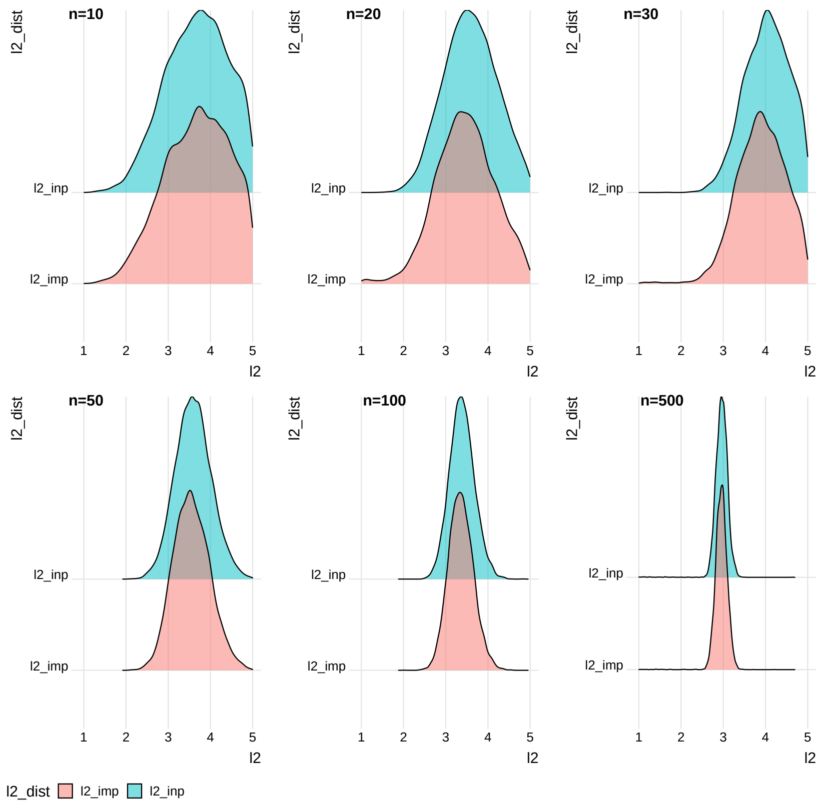


Figure 2.14: Posterior density plot of λ_2 (independent gamma and improper priors)

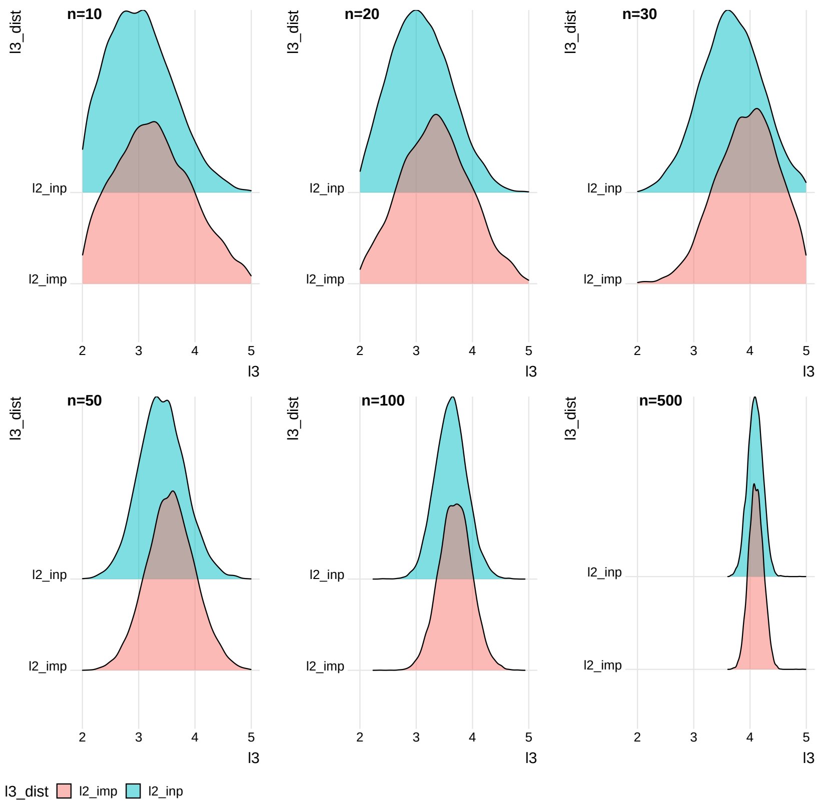


Figure 2.15: Posterior density plot of λ_3 (independent gamma and improper priors)

Table 2.1: Simulation (Sub-model I)

Sample size (n), parameters (P), posterior mean by improper prior (SIP_1), posterior mean by independent gamma prior (SIP_2), posterior mean by pseudo-gamma prior for $\psi_2 = 1$ (SIP_3), posterior mean by pseudo-gamma prior for $\psi_2 = 5$ (SIP_4), posterior mean by pseudo-gamma prior for $\psi_2 = 0$ (SIP_5), 95% confidence interval of SIP_1 ($CISIP_1$), 95% confidence interval of SIP_2 ($CISIP_2$), 95% confidence interval of SIP_3 ($CISIP_3$), 95% confidence interval of SIP_4 ($CISIP_4$), 95% confidence interval of SIP_5 ($CISIP_5$)

n	P	SIP_1	SIP_2	SIP_3	SIP_4	SIP_5	$CISIP_1$	$CISIP_2$	$CISIP_3$	$CISIP_4$	$CISIP_5$
10	λ_1	1.535	1.414	0.954	0.961	0.955	$(1.478, 1.791)$	$(1.349, 1.638)$	$(0.914, 1.110)$	$(0.912, 1.212)$	$(0.905, 1.161)$
	λ_3	3.959	3.588	3.203	3.194	3.213	$(3.896, 4.121)$	$(3.511, 3.877)$	$(3.145, 3.336)$	$(3.122, 3.335)$	$(3.132, 3.375)$
20	λ_1	1.176	1.128	0.875	0.866	0.865	$(1.135, 1.321)$	$(1.086, 1.310)$	$(0.829, 1.077)$	$(0.826, 1.020)$	$(0.827, 1.063)$
	λ_3	4.261	4.002	3.731	3.723	3.733	$(4.189, 4.492)$	$(3.928, 4.312)$	$(3.657, 3.858)$	$(3.663, 3.818)$	$(3.665, 3.957)$
30	λ_1	1.133	1.108	0.928	0.905	0.907	$(1.092, 1.347)$	$(1.059, 1.292)$	$(0.870, 1.320)$	$(0.868, 1.080)$	$(0.869, 1.135)$
	λ_3	4.054	3.896	3.688	3.682	3.684	$(4.012, 4.148)$	$(3.815, 4.152)$	$(3.626, 3.846)$	$(3.630, 3.863)$	$(3.626, 3.842)$
50	λ_1	1.048	1.042	0.913	0.909	0.907	$(1.010, 1.248)$	$(0.993, 1.297)$	$(0.874, 1.128)$	$(0.874, 1.073)$	$(0.874, 1.052)$
	λ_3	4.045	3.936	3.798	3.796	3.795	$(3.995, 4.236)$	$(3.879, 4.143)$	$(3.746, 3.989)$	$(3.744, 3.954)$	$(3.745, 3.931)$
100	λ_1	1.100	1.112	1.006	1.009	1.008	$(1.066, 1.303)$	$(1.054, 1.402)$	$(0.968, 1.213)$	$(0.968, 1.258)$	$(0.968, 1.279)$
	λ_3	4.540	4.482	4.393	4.395	4.391	$(4.499, 4.700)$	$(4.434, 4.677)$	$(4.347, 4.572)$	$(4.348, 4.625)$	$(4.345, 4.581)$
500	λ_1	1.120	1.127	1.113	1.110	1.111	$(1.094, 1.266)$	$(1.092, 1.326)$	$(1.075, 1.333)$	$(1.074, 1.298)$	$(1.075, 1.327)$
	λ_3	3.994	3.989	3.976	3.980	3.975	$(3.963, 4.151)$	$(3.953, 4.140)$	$(3.935, 4.181)$	$(3.936, 4.180)$	$(3.936, 4.172)$

Table 2.2: Simulation (Sub-model II)

Sample size (n), parameters (P), posterior mean by improper prior (SIP_1), posterior mean by independent gamma prior (SIP_2), posterior mean by pseudo-gamma prior for $\psi_2 = 1$ (SIP_3), posterior mean by pseudo-gamma prior for $\psi_2 = 5$ (SIP_4), posterior mean by pseudo-gamma prior for $\psi_2 = 0$ (SIP_5), 95% confidence interval of SIP_1 (CISIP_1), 95% confidence interval of SIP_2 (CISIP_2), 95% confidence interval of SIP_3 (CISIP_3), 95% confidence interval of SIP_4 (CISIP_4), 95% confidence interval of SIP_5 (CISIP_5)

n	P	SIP_1	SIP_2	SIP_3	SIP_4	SIP_5	CISIP_1	CISIP_2	CISIP_3	CISIP_4	CISIP_5
10	λ_1	1.531	1.402	1.035	1.040	1.027	$(1.479, 1.731)$	$(1.343, 1.589)$	$(1.002, 1.139)$	$(1.002, 1.213)$	$(0.995, 1.070)$
	λ_3	3.105	2.911	2.626	2.628	2.637	$(3.041, 3.286)$	$(2.847, 3.126)$	$(2.576, 2.703)$	$(2.577, 2.778)$	$(2.575, 2.816)$
20	λ_1	1.202	1.123	0.850	0.855	0.867	$(1.139, 1.536)$	$(1.082, 1.300)$	$(0.817, 0.982)$	$(0.818, 1.047)$	$(0.819, 1.119)$
	λ_3	4.244	4.029	3.834	3.836	3.832	$(4.201, 4.359)$	$(3.976, 4.189)$	$(3.772, 3.957)$	$(3.766, 4.112)$	$(3.773, 4.046)$
30	λ_1	1.127	1.096	0.921	0.937	0.928	$(1.090, 1.300)$	$(1.055, 1.255)$	$(0.886, 1.122)$	$(0.885, 1.226)$	$(0.885, 1.169)$
	λ_3	3.742	3.621	3.464	3.463	3.482	$(3.697, 3.889)$	$(3.570, 3.836)$	$(3.423, 3.547)$	$(3.411, 3.631)$	$(3.421, 3.616)$
50	λ_1	1.050	1.033	0.906	0.900	0.907	$(1.013, 1.219)$	$(0.992, 1.236)$	$(0.863, 1.158)$	$(0.862, 1.105)$	$(0.864, 1.136)$
	λ_3	4.250	4.149	4.032	4.033	4.030	$(4.199, 4.471)$	$(4.098, 4.370)$	$(3.978, 4.203)$	$(3.983, 4.220)$	$(3.986, 4.190)$
100	λ_1	1.115	1.098	1.003	1.002	1.003	$(1.064, 1.369)$	$(1.053, 1.338)$	$(0.967, 1.204)$	$(0.967, 1.198)$	$(0.967, 1.185)$
	λ_3	4.536	4.472	4.403	4.403	4.404	$(4.484, 4.753)$	$(4.428, 4.661)$	$(4.358, 4.600)$	$(4.358, 4.581)$	$(4.358, 4.598)$
500	λ_1	1.139	1.131	1.109	1.105	1.107	$(1.094, 1.380)$	$(1.092, 1.353)$	$(1.075, 1.314)$	$(1.074, 1.274)$	$(1.074, 1.299)$
	λ_3	4.052	4.039	4.029	4.022	4.028	$(4.014, 4.225)$	$(4.005, 4.192)$	$(3.990, 4.222)$	$(3.989, 4.170)$	$(3.989, 4.183)$

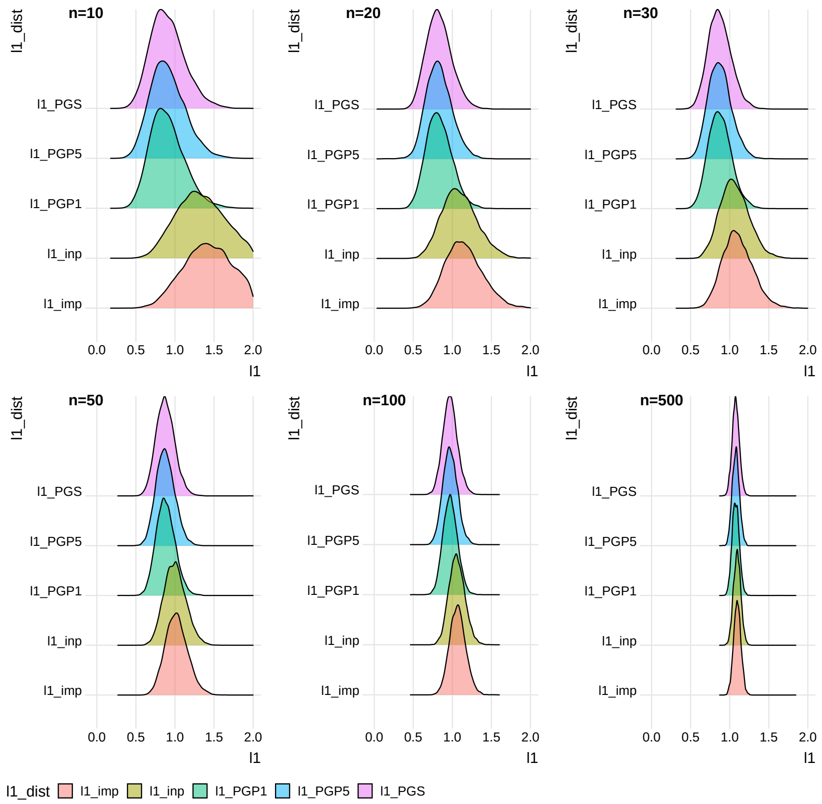


Figure 2.16: Posterior density plot of λ_1 of Sub-model I.

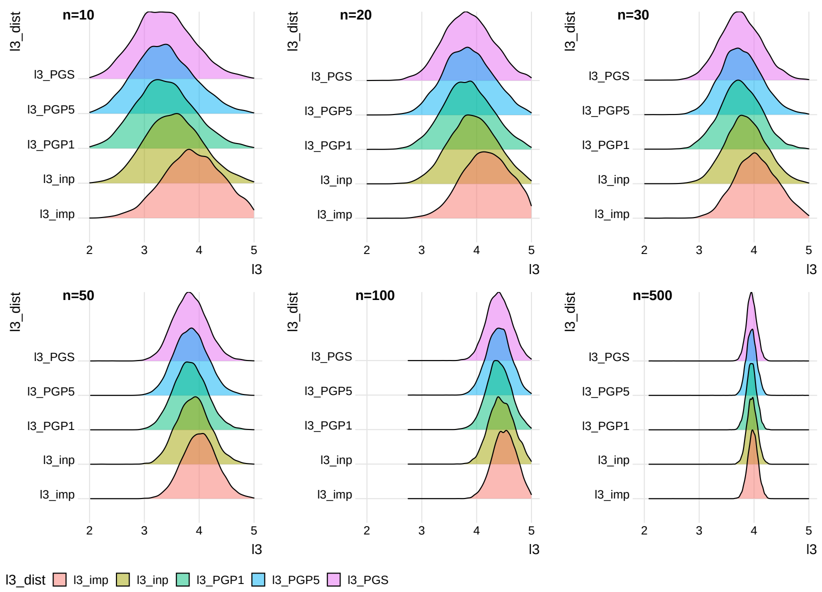


Figure 2.17: Posterior density plot of λ_3 of Sub-model I.

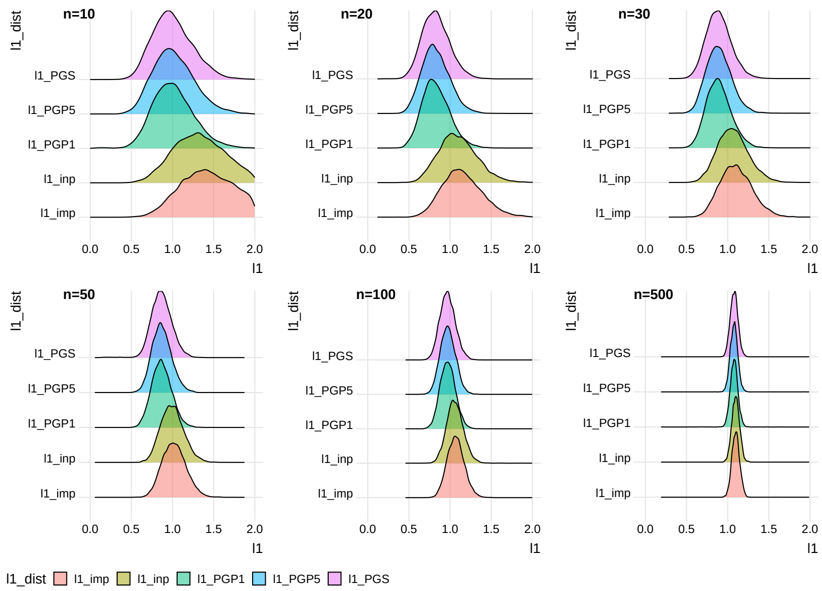


Figure 2.18: Posterior density plot of λ_1 of Sub-model II

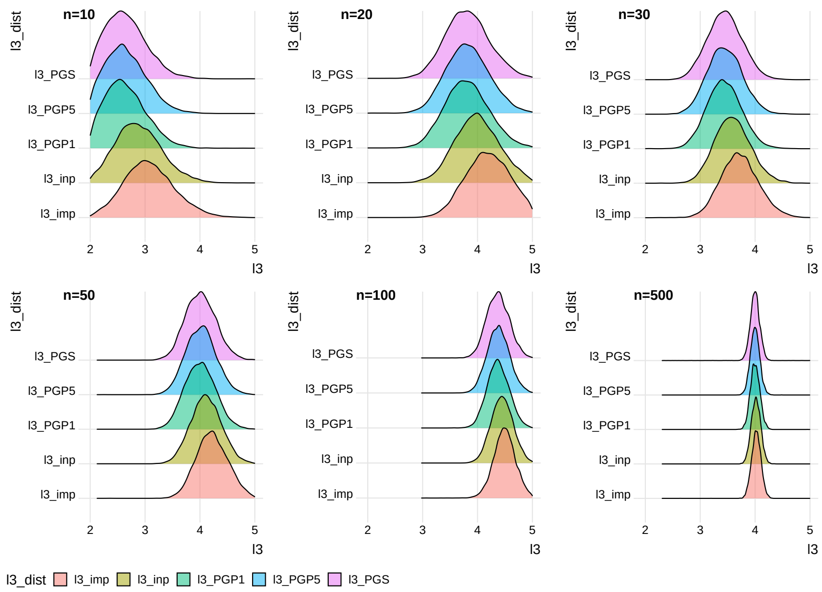


Figure 2.19: Posterior density plot of λ_3 of Sub-model II

Note that from the above simulation, Table 2.1 and 2.2 and posterior density plots with varying sample sizes summarize that with a sufficient sample size pseudo-gamma prior are closer to the true value than improper and independent gamma priors. Besides, in most cases, the estimated posterior confidence interval contains the true value. For example, in a sample size of $n = 10$, both improper and independent gamma prior posterior confident interval doesn't contain the true value. It is expected that with an increase in sample size, all five posterior distributions converge to true values. We also remark that posterior distribution behaviour is similar under pseudo-gamma prior under different values of ψ and varying sample sizes. We strongly advocate that from small sample sizes to sufficiently large samples, the estimated length of the confidence intervals is shorter under pseudo-gamma priors. Hence, we concluded that the standard errors of posterior means under pseudo-gamma priors are smaller than the improper and independent gamma priors. Finally, the simulation study insists we consider pseudo-gamma prior as the first choice under dependence priors.

2.5.2 Real data sets

In the following section, we consider two data sets which are mentioned in Karlis and Tsiamyrtzis [21], Islam and Chowdhury [19], Leiter and Hamdani [24] and also in Arnold and Manjunath [4]. We also consider an additional quantity of interest under each particular example: the ratio of the two marginal means, denoted by ϕ. For the importance of analyzing the posterior distribution of ϕ, see Karlis and Tsiamyrtzis [21] page 33.

Also, remark that pseudo-gamma prior with particular parameter value, i.e., $\psi_2 = 0$, the density function of ϕ is equal to λ_3 under Sub-models I & II. The exact distribution can be derived, c.f. equation (2.4.6) and (2.4.13). However, for further analysis, like moment computations, one has to rely on numerical methods. We conclude that under dependence priors, we can derive the exact density of the marginals for particular hyperparameters (i.e., $\psi_2 = 0$). Nevertheless, the prior mixture mentioned in Karlis and Tsiamyrtzis [21] does not have closed-form expressions for marginals, so also, for ϕ, one needs to depend on numerical computations.

A particular data set I

We consider a data sets which is mentioned in Islam and Chowdhury [19] and also in Arnold and Manjunath [4]; the source of the data is from the tenth wave of the Health and Retirement Study (HRS). The data represents the number of conditions that ever had (X) as mentioned by the doctors and utilization of healthcare services (say, hos-

Table 2.3: Health and retirement study data (Full Model)

n	P	$\mathbf{SIP}_1$	$\mathbf{SIP}_2$	$\mathbf{CISIP}_1$	$\mathbf{CISIP}_2$
	λ_1	2.642	2.642	$(2.608, 2.673)$	$(2.608, 2.678)$
5567	λ_2	0.639	0.628	$(0.598, 0.683)$	$(0.585, 0.667)$
	λ_3	0.049	0.054	$(0.033, 0.642)$	$(0.039, 0.070)$

pital, nursing home, doctor and home care) (Y). The Pearson correlation coefficient between X and Y is 0.063. The test for independence, classical inference (m.l.e and moment estimates), and AIC values for full and its sub-models c.f. Arnold and Manjunath [4] page 16 and 18 (Table 10).

In the following, we will consider the following two models. The criteria for selecting below two models are discussed in Arnold, and Manjunath [4] on page 18 and Table 10.

Full-model : For the full model, the ratio of two marginal means is

$$\phi = \frac{\lambda_2 + \lambda_3\lambda_1}{\lambda_1}$$

We refer to Table 2.3 and Figure 2.20 for posteriori analysis and the sample density of ϕ for the Full-model.

Sub-model II: Similarly, for the Sub-model II, the ratio of two marginal means will be equal to

$$\phi = \lambda_3$$

We refer to Table 2.4 and Figure 2.21 for posteriori analysis and the sample density of ϕ for the Full and its Sub-model II.

Table 2.4: Health and retirement study data (Sub-model II)

n	P	SIP_1	SIP_2	SIP_3	SIP_4	SIP_5	CISIP_1	CISIP_2	CISIP_3	CISIP_4	CISIP_5
5567	λ_1	2.674	2.642	2.648	2.767	2.642	$(2.610, 2.681)$	$(2.605, 2.677)$	$(2.610, 2.680)$	$(2.605, 2.681)$	$(2.606, 2.681)$
	λ_3	0.291	0.290	0.309	0.302	0.291	$(0.284, 0.298)$	$(0.282, 0.298)$	$(0.283, 0.298)$	$(0.284, 0.298)$	$(0.284, 0.298)$

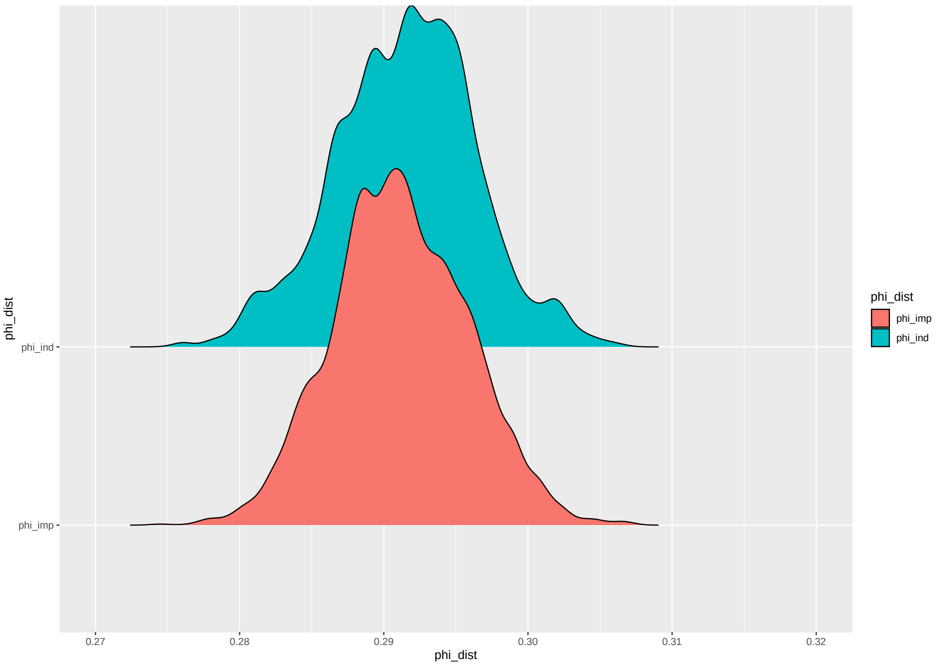


Figure 2.20: Posterior density plot of ϕ of Full-model (improper and independent-gamma prior)

A particular data set II

Now, we consider a data set in Leiter and Hamdani [24]; the data source is a 50-mile stretch of Interstate 95 in Prince William, Stafford, and Spotsylvania counties in Eastern Virginia. The data represents the number of accidents categorized as fatal accidents, injury accidents, or property damage accidents, along with the corresponding number of fatalities and injuries for the period 1 January 1969 to 31 October 1970. For classical inference (m.l.e and moment estimates) and AIC values for full and its sub-models c.f. Arnold and Manjunath [4] page 17 and 19 (Table 11). The ratio of two marginal means, i.e., ϕ, refers to the number of fatalities per accident. The criteria for selecting below two models are discussed in Arnold, and Manjunath [4] on page 19 and Table 11. Moreover, the Mirror Sub-model II suggested below is the same model considered

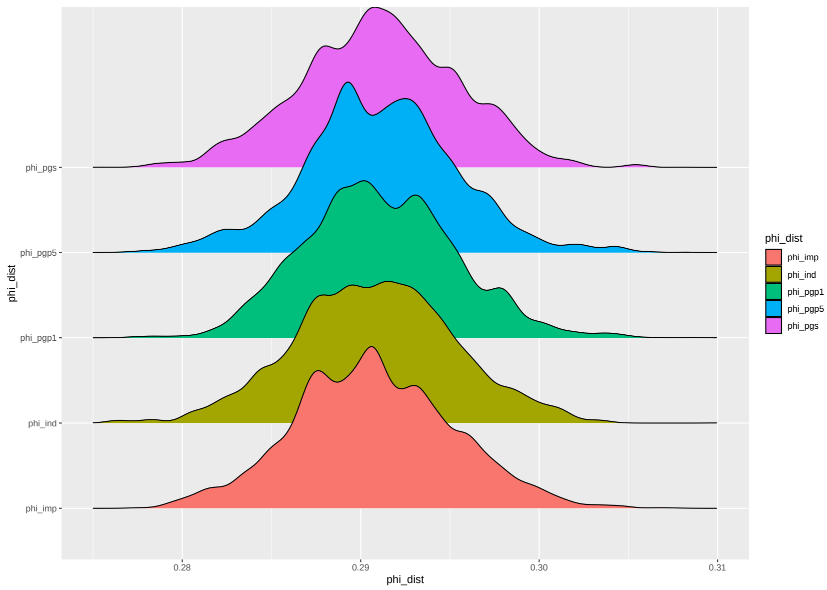


Figure 2.21: Posterior density plot of ϕ of Sub-model II

Table 2.5: Accidents and fatalities study data (Full Model)

n	P	SIP_1	SIP_2	CISIP_1	CISIP_2
	λ_1	0.059	0.060	$(0.044, 0.076)$	$(0.044, 0.077)$
639	λ_2	0.814	0.813	$(0.753, 0.879)$	$(0.753, 0.973)$
	λ_3	0.881	0.952	$(0.534, 1.269)$	$(0.641, 1.268)$

in Leiter and Hamdani [24]. It has been emphasized in Leiter and Hamdani [24] and Arnold and Manjunath [4] that mirror Sub-model II fits the data better than any other sub-models.

Full-model : We refer to Table 2.5 for posteriori analysis and the sample density of ϕ for the Full-model. Note that, since Mirror Sub-model II fits the data and for the Full model computation ϕ or reference of two marginal means is of no significance to the application.

Mirror Sub-model II: We refer Table 2.6 and Figure 2.22 for posteriori analysis and the sample density of ϕ for the mirror Sub-model II.

As mentioned earlier, under pseudo-gamma prior (dependence), the estimated confidence interval includes posterior means. Also, under specific hyper-parameters, i.e.,

Table 2.6: Accidents and fatalities study data (Mirrored Sub-model II)

n	P	SIP$_1$	SIP$_2$	SIP$_3$	SIP$_4$	SIP$_5$	CISIP$_1$	CISIP$_2$	CISIP$_3$	CISIP$_4$	CISIP$_5$
639	λ_1	0.860	0.908	0.867	0.867	0.871	$(0.802, 0.917)$	$(0.803, 0.921)$	$(0.807, 0.928)$	$(0.800, 0.928)$	$(0.807, 0.927)$
	λ_3	0.069	0.076	0.069	0.072	0.132	$(0.052, 0.088)$	$(0.056, 0.095)$	$(0.052, 0.089)$	$(0.051, 0.089)$	$(0.057, 0.095)$

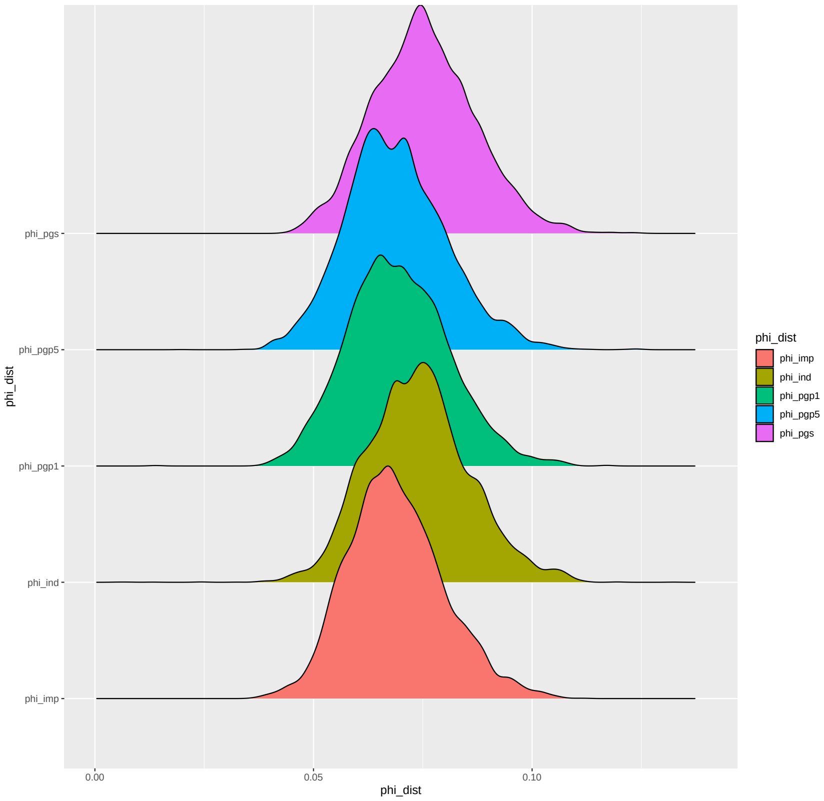


Figure 2.22: Posterior density plot of ϕ of Mirror Sub-model II

$\psi_2 = 0$, we have an exact density of the ϕ, so simulation and further analysis are reliable. Due to the simplicity of its structure, the pseudo-gamma prior allows for simple simulation and marginal distribution computations. Unlike the mixture prior (dependence) suggested in Karlis and Tsiamyrtzis [21], the computation of marginal distributions or simulating from the same, one has to rely on numerical computations. The pseudo-gamma discussed in this paper may be feasible alternatives to the other dependence priors for analyzing count data sets.

In the following, we extend bivariate pseudo-Poisson to a high dimension. Due to the simplicity of trivariate pseudo-Poisson, we have considered improper and independent gamma priors for an illustration of higher dimensions of the above-mentioned approaches.

2.6 Trivariate pseudo-Poisson model

Analogously, we could consider a trivariate model with:

$$X \sim P(\lambda_1)$$

$$Y|X = x \sim P(\lambda_2 + \lambda_3 x)$$

$$Z|X = x, Y = y \sim P(\lambda_4 + \lambda_5 x + \lambda_6 y)$$

For the parameter values $\lambda_1 = 1$, $\lambda_2 = 3$, $\lambda_3 = 4$, $\lambda_4 = 4$, $\lambda_5 = 2$ and $\lambda_6 = 5$, we consider the following priors:

FP_1: Uniform prior (improper prior)

FP_2: Independent gamma prior, i.e., $\lambda_1 \sim \Gamma(\alpha_1, \delta_1)$, $\lambda_2 \sim \Gamma(\alpha_2, \alpha_2)$, $\lambda_3 \sim \Gamma(\alpha_3, \delta_3)$, $\lambda_4 \sim \Gamma(\alpha_4, \delta_4)$, $\lambda_5 \sim \Gamma(\alpha_5, \delta_5)$ and $\lambda_6 \sim \Gamma(\alpha_6, \delta_6)$ for the prior values $\alpha_1 = 1$, $\delta_1 = 1$, $\alpha_2 = 3$, $\delta_2 = 1$, $\alpha_3 = 4$,$\delta_3 = 2$, $\alpha_4 = 2$, $\delta_4 = 3$, $\alpha_5 = 1$, $\delta_5 = 3$, $\alpha_6 = 1$ and $\delta_6 = 1$.

The posterior density plots of λ_1, λ_2 and λ_3 with improper (Uniform prior) and independent gamma priors c.f. Figures 2.23, 2.24, 2.25, 2.26, 2.27 and 2.28 . The plot has been generated with one set of 100000 observations with thinning at 10th sample from posteriori with varying sample size. As mentioned in the bivariate case, posteriori estimates are close to the true value for a sufficiently large sample size.

The sub model in which $\lambda_2 = \lambda_3$ and $\lambda_4 = \lambda_5 = \lambda_6$ is particularly easy to analyze. Further analysis and applications of trivariate pseudo-Poisson models will be discussed in a separate report.

The analysis's overall summary and detailed overview are included in the Conclusion Section.

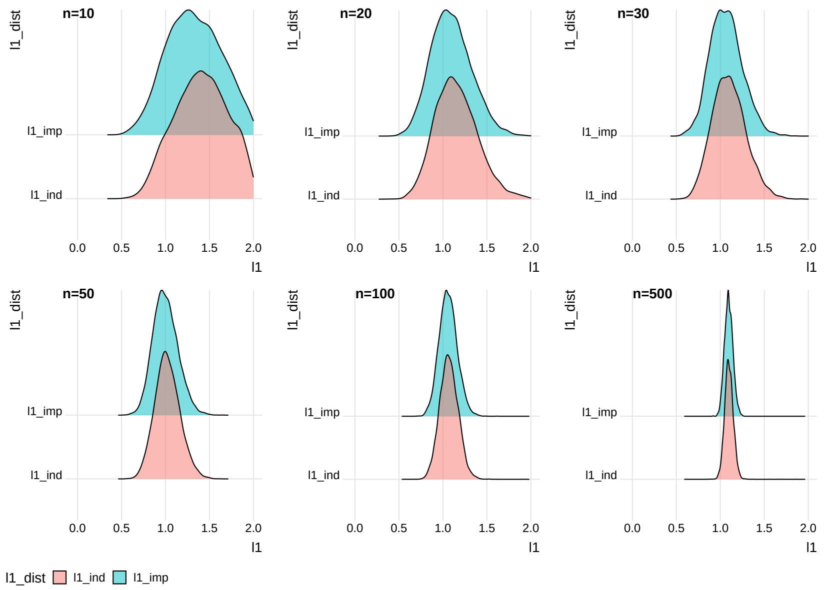


Figure 2.23: Posterior density plot of λ_1 (independent gamma and improper priors)

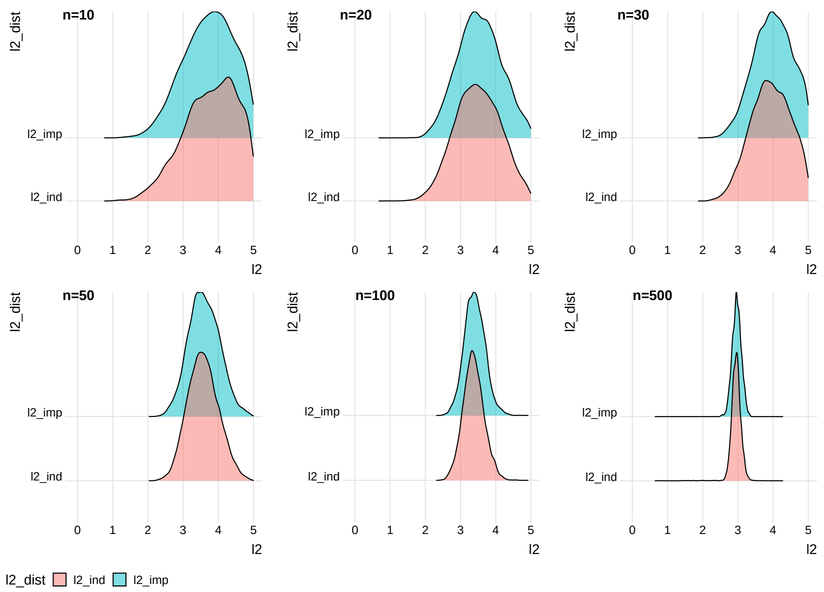


Figure 2.24: Posterior density plot of λ_2 (independent gamma and improper priors)

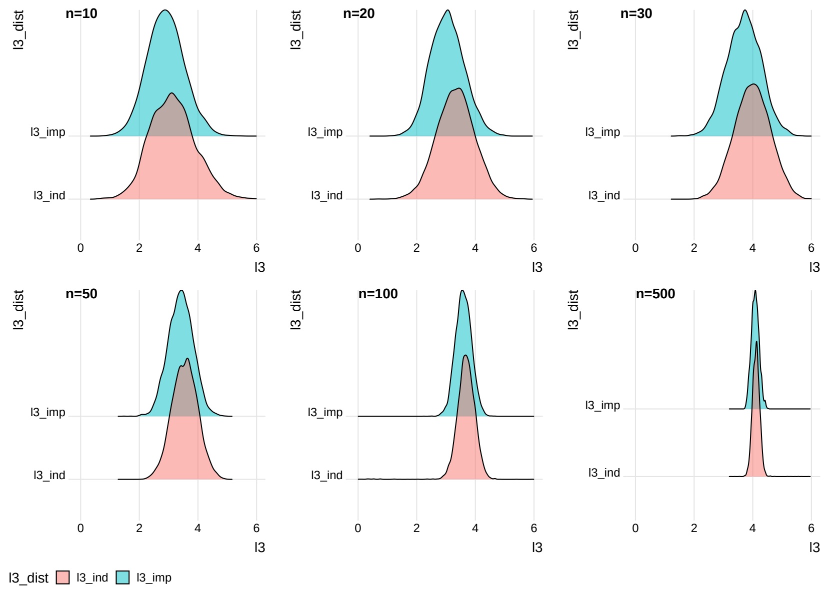


Figure 2.25: Posterior density plot of λ_3 (independent gamma and improper priors)

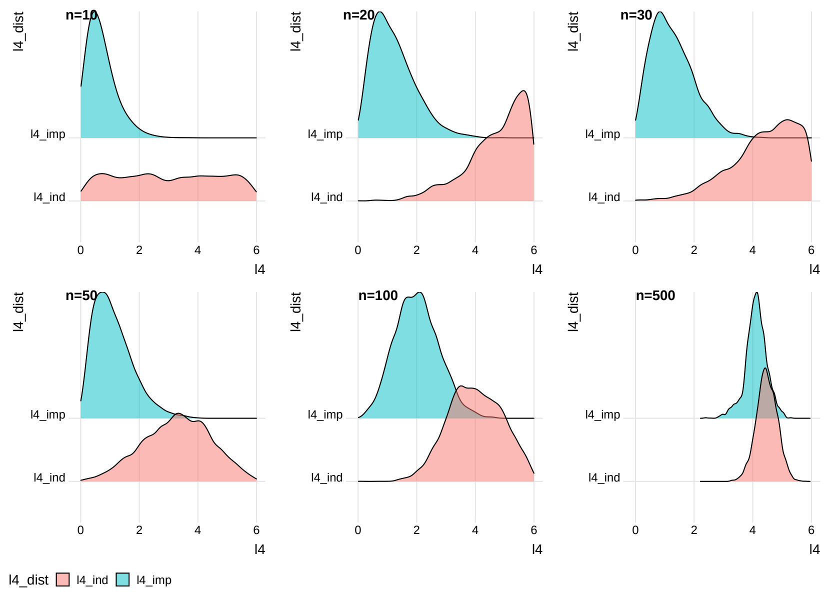


Figure 2.26: Posterior density plot of λ_4 (independent gamma and improper priors)

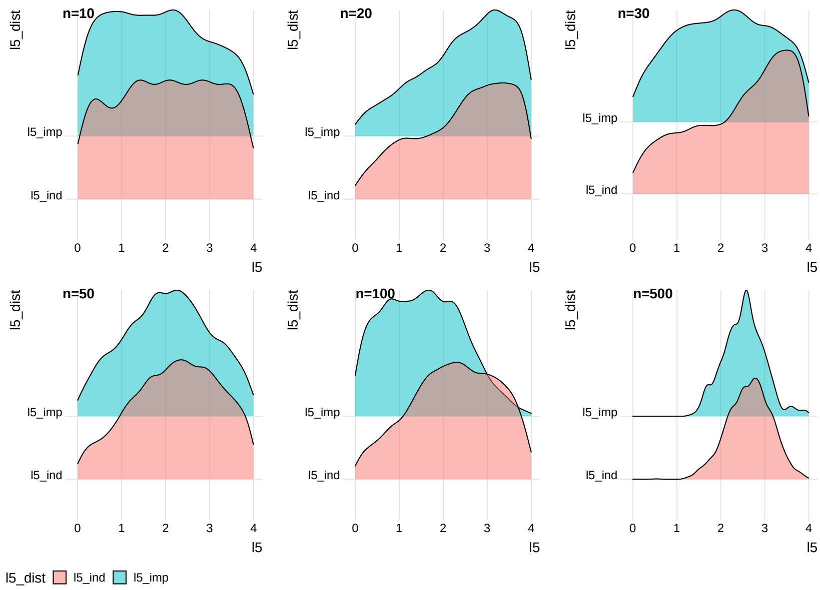


Figure 2.27: Posterior density plot of λ_5 (independent gamma and improper priors)

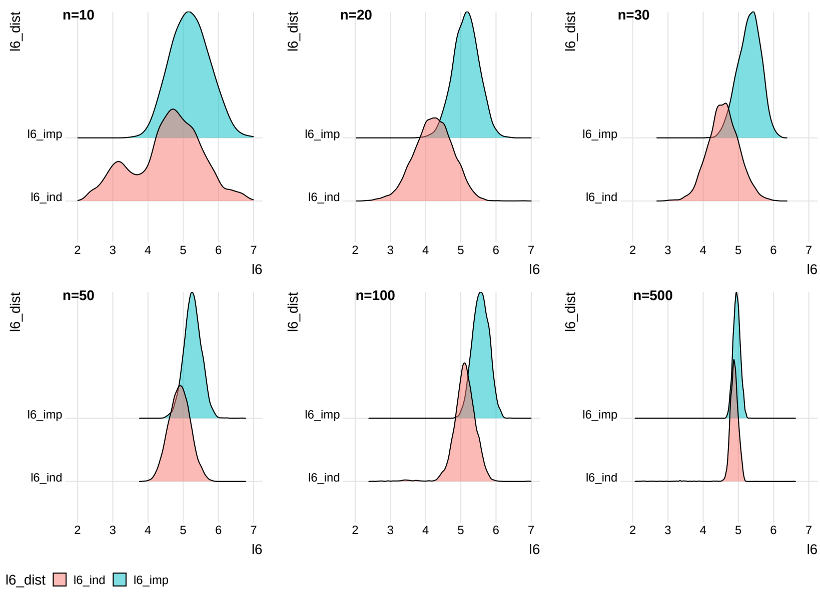


Figure 2.28: Posterior density plot of λ_6 (independent gamma and improper priors)

CHAPTER 3

On pseudo-Poisson goodness-of-fit tests

3.1 Introduction

Indeed, goodness-of-fit (GoF) is a statistical procedure to test whether the given data is compatible with the assumed distribution. Any GoF test requires the following three steps: (1) identifying the unique characteristic of the assumed model (examples: distribution function, generating function, or density function); (2) computing the empirical version of the assumed characteristic; (3) with the pre-assumed measure(examples: L_1 - or L_2-space), measure the distance between the assumed item in Step (1) and its empirical one, in Step (2). A rejection region can be computed with a given level and the cut-off value for the distance measure determined. However, if the rejection region can not be derived explicitly, then one can use the Bootstrapping technique to generate a critical region. The general steps required to simulate a rejection region using the Bootstrapping are discussed detail in Section 3.4. We refer to Meintanis [26] and Nikitin [31] for a detailed discussion of the GoF tests, which involve the aforementioned steps. Besides, there exist or can be constructed tests that are not based on a unique characteristic of the assumed distribution. For example, considering the univariate Poisson distribution, there exists a GoF test which depends on the Fisher index of depression. We also know that the Poisson distribution belongs to the class of equi-dispersed models, but this property does not characterize the Poisson distribution. Hence, such tests, which are not based on a unique characteristic of the assumed distributions, are not consistent tests.

The literature on GoF tests for bivariate count data is sparse. For the classical bivariate and multivariate Poisson distributions, a GoF test using the probability generating function is discussed by Muñoz and Gamero [29] and Muñoz and Gamero [30]. More recently, Muñoz [28] contains a review of the available bivariate GoF tests and also a new test using the differentiation of the probability generating function(p.g.f.).

In the following sections, we are starting with a test defined in Kocherlakota and Kocherlakota [22] and a few bivariate GoF tests reviewed in Muñoz [28]. In addition to the classical GoF tests using probability generating function (p.g.f.), we considered a less known test statistic which is the supremum of the absolute difference between estimated p.g.f. and empirical ones. In addition, we are introducing a non-consistent test which is based on the moments, in particular, defining the test taking the difference between the estimated bivariate Fisher index and its empirical counterpart. We examine each test's finite, large, and asymptotic properties and recommend a few tests based on their power and robustness analysis.

Before we start a discussion on GoF tests, we would like to make a few remarks on the bivariate pseudo-Poisson model and its relevance in the literature. Finally, we refer to Arnold and Manjunath [4] and Arnold et. al. [1] for classical inferential aspects, characterization, Bayesian analysis, and also an example of applications of the bivariate pseudo-Poisson model.

3.2 Bivariate pseudo-Poisson models

As discussed in Chapter 1 Section 1.1 (Definition 1.1), the bivariate pseudo-Poisson definition is given by

Definition 3.1. *A* 2*-dimensional random variable* (X, Y) *is said to have a bivariate Pseudo-*

Poisson distribution if there exists a positive constant λ_1 *such that*

$$X \sim \mathscr{P}(\lambda_1)$$

and a function $\lambda_2 : \{0, 1, 2, ...\} \to (0, \infty)$ *such that, for every non-negative integer* x*,*

$$Y|X = x \sim \mathscr{P}(\lambda_2(x)).$$

Here we restrict the form of the function $\lambda_2(x)$ to be a polynomial with unknown coefficients. In particularly the simple form we assume is that $\lambda_2(x) = \lambda_2 + \lambda_3 x$, $\forall x \in \{0, 1, 2, ...\}$, then the above bivariate distribution will be of the form

$$X \sim \mathscr{P}(\lambda_1) \tag{3.2.1}$$

and $x \in \{0, 1, 2, ...\}$,

$$Y|X = x \sim \mathscr{P}(\lambda_2 + \lambda_3 x). \tag{3.2.2}$$

The parameter space for this model is $\{(\lambda_1, \lambda_2, \lambda_3) : \lambda_1 > 0, \lambda_2 > 0, \lambda_3 \geq 0\}$. The case in which the variables are independent corresponds to the choice $\lambda_3 = 0$. The probability generating function (p.g.f) for this bivariate Pseudo-Poisson distribution is given by

$$G(t_1, t_2) = e^{\lambda_2(t_2 - 1)} e^{\lambda_1 [t_1 e^{\lambda_3(t_2 - 1)} - 1]};\ t_1, t_2 \in \mathbb{R}. \tag{3.2.3}$$

Now, the marginal p.g.f of Y is

$$G(1, t_2) = G_Y(t_2) = e^{\lambda_2(t_2 - 1)} e^{\lambda_1 [e^{\lambda_3(t_2 - 1)} - 1]};\ t_2 \in \mathbb{R}. \tag{3.2.4}$$

Note that, in general, the p.g.f. in equation (3.2.3) can not be simplified to compute all marginal probabilities. Yet, we can use equation (3.2.4) to derive a few marginal probabilities of Y. The derivation of marginal probability of Y is demonstrated for

$Y = 0, 1, 2, 3$ in Appendix A.1, and one can still extend the mentioned procedure to get albeit complicated values for the probability that Y assumes any positive value. Besides, the derivation of the other conditional distribution of the bivariate pseudo-Poisson, i.e., $P(X = x|Y = y)$, has been included in Appendices A.1 Section.

3.3 GoF tests for the bivariate pseudo-Poisson

In the following section, we discuss GoF tests (I) based on unique characteristics (consistent tests), (II) based on the moments (non-consistent tests) and, (III) the simple classical χ^2 goodness of fit test.

3.3.1 New tests based on moments

In the following, we will be extending a univariate GoF test based on the Fisher index to the bivariate case. We know that for a multivariate distribution, the Fisher index of dispersion is not uniquely defined. However, in the following, we use the definition of the multivariate Fisher dispersion given by Kokonendji and Puig [23] in Section 3 as; for any d-dimensional discrete random variable $\mathbf{Z}$ with mean vector $E(\mathbf{Z})$ and covariance matrix $Cov(\mathbf{Z})$ the generalized dispersion index is

$$GDI(\mathbf{Z}) = \frac{\sqrt{\mathbf{E(Z)}}^{\mathbf{T}}\mathbf{Cov(Z)}\sqrt{\mathbf{E(Z)}}}{\mathbf{E(Z)^T E(Z)}}. \tag{3.3.1}$$

For the bivariate pseudo-Poisson model, define the random vector $\mathbf{Z} = (\mathbf{X}, \mathbf{Y})^{\mathbf{T}}$ for and the moments are (c.f. Arnold and Manjunath [4] page 2309–2310)

$$E(\mathbf{Z}) = (\lambda_1, \lambda_2 + \lambda_3\lambda_1)^T \tag{3.3.2}$$

$$cov(\mathbf{Z}) = \begin{bmatrix} \lambda_1 & \lambda_1\lambda_3 \\ \lambda_1\lambda_3 & \lambda_2 + \lambda_3\lambda_1 + \lambda_3^2\lambda_1 \end{bmatrix}.$$

Now, using the definition given in Kokonendji and Puig [23] page 183, the dispersion index for the bivariate pseudo-Poisson is

$$\begin{aligned}
GDI(\mathbf{Z}) &= \frac{\lambda_1^2 + 2\lambda_1^{\frac{3}{2}}\lambda_3\sqrt{\lambda_2+\lambda_3\lambda_1} + (\lambda_2+\lambda_3\lambda_1)(\lambda_2+\lambda_3\lambda_1+\lambda_3^2\lambda_1)}{\lambda_1^2+(\lambda_2+\lambda_3\lambda_1)^2} \\
&= 1 + \frac{2\lambda_1^{\frac{3}{2}}\lambda_3\sqrt{\lambda_2+\lambda_3\lambda_1} + (\lambda_2+\lambda_3\lambda_1)\lambda_3^2\lambda_1}{\lambda_1^2+(\lambda_2+\lambda_3\lambda_1)^2} > 1,
\end{aligned} \tag{3.3.3}$$

which indicates over-dispersion.

For the corresponding sample version, consider the n sample observations. If $\mathbf{Z}_1 = (X_1, Y_1)^T, \ldots, \mathbf{Z}_n = (X_n, Y_n)^T$ is a random sample of size n from a bivariate population. Denote $\overline{\mathbf{Z}}_n = \frac{1}{n}\sum_{i=1}^n \mathbf{Z}_i = (\overline{X}, \overline{Y})^T$ and $\widehat{cov(\mathbf{Z})} = \frac{1}{n-1}\sum_{i=1}^n \mathbf{Z}_i\mathbf{Z}_i^T - \overline{\mathbf{Z}}_n\overline{\mathbf{Z}}_n^T$ are the sample mean vector and the sample covariance matrix, respectively. Then the empirical bivariate dispersion index is

$$\widehat{GDI(\mathbf{Z})}_n = \frac{\sqrt{\overline{\mathbf{Z}}_n^T}\,\widehat{cov(\mathbf{Z})}\sqrt{\overline{\mathbf{Z}}_n}}{\overline{\mathbf{Z}}_n^T\overline{\mathbf{Z}}_n}. \tag{3.3.4}$$

According to Theorem 1 in Kokonendji and Puig [23] page 184, as $n \to \infty$, $\sqrt{n}\{\widehat{GDI(\mathbf{Z})}_n - GDI(\mathbf{Z})\} \sim N(0, \sigma_g^2)$, where $\sigma_g^2 = \Delta^T\Gamma\Delta$, where

$$\Gamma = \begin{bmatrix} \Sigma & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

and

$$\Sigma = \begin{bmatrix} var(X) & cov(X,Y) \\ cov(X,Y) & var(Y) \end{bmatrix}.$$

A new bivariate GoF test for the count data based on the Fisher dispersion index is

$$FI_n^{(.)} = \sqrt{n}\{\widehat{GDI(\mathbf{Z})}_n - GDI(\mathbf{Z})\} \tag{3.3.5}$$

and the null hypothesis is rejected for large values of $FI_n^{(.)}$. The asymptotic distribution of the test statistic is

$$\frac{\widehat{GDI(\mathbf{Z})}_n - GDI(\mathbf{Z})}{\frac{\sigma_g}{\sqrt{n}}} \sim^{asy.} N(0,1), \text{ as } n \to \infty. \tag{3.3.6}$$

For detailed proof, c.f. Theorem 1 in Kokonendji and Puig [23] page 184. However, for the two sub-models of the bivariate pseudo-Poisson model, i.e., when $\lambda_2 = \lambda_3$ is Sub-Model I and when $\lambda_2 = 0$ is Sub-Model II, the new test statistics are

$$FI_n^{(SI)} = \sqrt{n}\{\widehat{GDI(\mathbf{Z})}_n - GDI^{(SI)}(\mathbf{Z})\}, \tag{3.3.7}$$

and

$$FI_n^{(SII)} = \sqrt{n}\{\widehat{GDI(\mathbf{Z})}_n - GDI^{(SII)}(\mathbf{Z})\}, \tag{3.3.8}$$

where

$$GDI^{(SI)}(\mathbf{Z}) = 1 + \frac{2\lambda_1^{\frac{3}{2}}\lambda_3^{\frac{3}{2}}\sqrt{1+\lambda_1} + (1+\lambda_1)\lambda_3^3\lambda_1}{\lambda_1^2 + \lambda_3^2(1+\lambda_1)^2}, \tag{3.3.9}$$

and

$$GDI^{(SII)}(\mathbf{Z}) = 1 + \frac{2\lambda_1^{\frac{3}{2}}\lambda_3^{\frac{3}{2}}\sqrt{\lambda_1} + \lambda_3^3\lambda_1^2}{\lambda_1^2 + \lambda_3^2\lambda_1^2}. \tag{3.3.10}$$

One can derive test statistic $FI_n^{(SI)}$ and $FI_n^{(SII)}$. The estimated dispersion index can be obtained by plugging in the m.l.e estimates of $\lambda_i, i = 1,2,3$. Also, due to the invariance and asymptotic properties of the m.l.e estimates, the proposed test statistics are normally distributed (with appropriate scaling). For large sample sizes, the null hypothesis is rejected whenever the test statistic value exceeds the standard normal quantile value. In Section 3.4, we analyze the proposed test statistic's finite, large, and asymptotic behavior.

3.3.2 Tests based on unique characteristic

In the following, we consider a few test statistics for the full, sub-model I, and sub-model II.

Muñoz and Gamero (M&G) method

The GoF tests for a bivariate random variable based on the finite sample size are limited. This is due to difficulty deriving closed-form expression for the critical region with finite sample size. Yet, in the following, we construct the GoF test for the bivariate pseudo-Poisson distribution using Munoz and Gamero's[29] finite sample size test for the classical bivariate Poisson distribution. For a finite sample test based on the p.g.f to test GoF for the univariate Poisson, we refer to Unam and Cimat [37]. Furthermore, using a bootstrapping technique, the critical region for the test is simulated and illustrated with an example in Section 3.4.

Let (X, Y) be a bivariate random variable with p.g.f, $G(t_1, t_2; \lambda_1, \lambda_2, \lambda_3)$, $(t_1, t_2)^T \in [0, 1]^2$. For the given data set (X_i, Y_i), $i = 1, ..., n$, we denote by $G_n(t_1, t_2) = \frac{1}{n}\sum_{i=1}^{n} t_1^{X_i} t_2^{Y_i}$ an empirical counterpart of the bivariate p.g.f. According to Muñoz and Gamero [29], a reasonable test for testing the compatibility of the assumed density should reject the null hypothesis for large values of the given statistic

$$T_{P,n,w}^{(.)}(\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3) = \int_0^1 \int_0^1 g_n^2(t_1, t_2; \hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3) w(t_1, t_2) dt_1 dt_2 \tag{3.3.11}$$

where $\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3$ are consistent estimators of λ's and
$g_n(t_1, t_2; \hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3) = \sqrt{n}\{G_n(t_1, t_2) - G(t_1, t_2; \hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3)\}$ and also $w(t_1, t_2) \geq 0$ is a measurable weight function satisfying

$$\int_0^1 \int_0^1 w(t_1, t_2) dt_1 dt_2 < \infty. \tag{3.3.12}$$

The above condition implies that the test statistic $T^{(.)}_{P,n,w}(\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3)$ is finite for the fixed sample size n. Similarly, for the Sub-model, I & II with appreciate p.g.f. one can derive test statistic $T^{(SI)}_{P,n,w}$ and $T^{(SII)}_{P,n,w}$.

Due to the difficulty in obtaining an explicit expression for the critical region, it has been argued in Muñoz and Gamero [29] and in Muñoz [28], the rejection regions can be simulated using the bootstrapping method. The general procedure to identify an appropriate weight function is difficult to argue. One can consider the weight functions, which include a more prominent family of functions. A few weight functions are considered in Appendix A in Example 3 and derived from its test statistics. In Section 3.4, we analyzed the effects of weight functions and feasible parameter values on the critical region.

Kocherlakota and Kocherlakota(K&K) method

Let $\mathbf{Z_1}, ..., \mathbf{Z_n}$ be a random sample from the bivariate distribution $F(\mathbf{Z}; \grave{\_})$, where $\underline{\theta} = (\theta_1, ..., \theta_d)^T$ is the d-dimensional parameter vector. Let $G(t_1, t_2; \underline{\theta})$ be the p.g.f. of $\mathbf{Z} = (\mathbf{X}, \mathbf{Y})^{\mathbf{T}}$, $t_1, t_2 \in \mathbb{R}^2$ and parameter vector $\underline{\theta}$ is estimated by the maximum likelihood estimation (m.l.e) method and the estimator we denote by $\widehat{\underline{\theta}}$. Let $G_n(t_1, t_2) = \frac{1}{n}\sum_{i=1}^{n} t_1^{X_i} t_2^{Y_i}$, $t \in \mathbb{R}$ be the empirical p.g.f. (e.p.g.f.) then the test statistic

$$T_N(t_1, t_2) = \frac{G_n(t_1, t_2) - G(t_1, t_2; \widehat{\underline{\theta}})}{\sigma}, \ |t_1| < 1; |t_2| < 1 \tag{3.3.13}$$

is asymptotically follows the standard normal distribution, where $\sigma^2 = \frac{1}{n}[G(t_1^2, t_2^2; \underline{\theta}) - G^2(t_1, t_2; \underline{\theta}))] - \sum_{i=1}^{d}\sum_{j=1}^{d} \sigma_{i,j} \frac{\partial G(t_1,t_2;\underline{\theta})}{\partial \theta_i} \frac{\partial G(t_1,t_2;\underline{\theta})}{\partial \theta_j}$, $((\sigma_{i,j}))$ is the inverse of the Fisher information matrix and σ can be estimated by plugging in the m.l.e of $\underline{\theta}$. We refer to Kocherlakota and Kocherlakota(K&K) [22] for the asymptotic distribution of the test statistic.

Now, for the sub-model I, the Fisher information matrix is

$$I^{(SI)}(\lambda_1, \lambda_3) = n \begin{bmatrix} E\left(\frac{X}{\lambda_1^2}\right) & 0 \\ 0 & E\left(\frac{Y}{\lambda_3^2}\right) \end{bmatrix} = \begin{bmatrix} \frac{n}{\lambda_1} & 0 \\ 0 & \frac{n(1+\lambda_1)}{\lambda_3} \end{bmatrix}.$$

Similarly, for the sub-model II, the Fisher information matrix is

$$I^{(SII)}(\lambda_1,\lambda_3) = n\begin{bmatrix} E\left(\frac{X}{\lambda_1^2}\right) & 0 \\ 0 & E\left(\frac{Y}{\lambda_3^2}\right)\end{bmatrix} = \begin{bmatrix} \frac{n}{\lambda_1} & 0 \\ 0 & \frac{n\lambda_1}{\lambda_3}\end{bmatrix}.$$

The GoF test statistic under sub-model I is

$$T_{PN}^{(SI)}(t_1,t_2) = \frac{G_n(t_1,t_2) - G_I(t_1,t_2;\hat{\lambda}_1,\hat{\lambda}_3)}{\sigma^{(SI)}},\ |t_1|<1, |t_2|<1 \tag{3.3.14}$$

where $G_n(.)$ is empirical p.g.f. and $G_I(t_1,t_2;\hat{\lambda}_1,\hat{\lambda}_3)$ is estimated p.g.f. of the sub-model I and

$$\begin{aligned}\sigma^{2(SI)} &= \frac{1}{n}[G_I(t_1^2,t_2^2;\lambda_1,\lambda_3) - G_I^2(t_1,t_2;\lambda_1,\lambda_3)] - \frac{\lambda_1}{n}\frac{\partial^2 G_I(t_1,t_2;\lambda_1,\lambda_3)}{\partial\lambda_1^2} \\ &\quad - \frac{\lambda_3}{n(\lambda_1+1)}\frac{\partial^2 G_I(t_1,t_2;\lambda_1,\lambda_3)}{\partial\lambda_3^2}.\end{aligned} \tag{3.3.15}$$

Similarly, for the sub-model II, the GoF test statistic will be

$$T_{PN}^{(SII)} = \frac{G_n(t_1,t_2) - G_{II}(t_1,t_2;\hat{\lambda}_1,\hat{\lambda}_3)}{\sigma^{(SII)}},\ |t_1|<1, |t_2|<1. \tag{3.3.16}$$

where $G_n(.)$ is empirical p.g.f. and $G_{II}(t_1,t_2;\hat{\lambda}_1,\hat{\lambda}_3)$ is estimated p.g.f. of the sub-model II and

$$\begin{aligned}\sigma^{2(SII)}(t_1,t_2) &= \frac{1}{n}[G(t_1^2,t_2^2;\lambda_1,\lambda_3) - G_{(II)}^2(t_1,t_2;\lambda_1,\lambda_3)] - \frac{\lambda_1}{n}\frac{\partial^2 G_{(II)}(t_1,t_2;\lambda_1,\lambda_3)}{\partial\lambda_1^2} \\ &\quad - \frac{\lambda_3}{n\lambda_1}\frac{\partial^2 G_{II}(t_1,t_2;\lambda_1,\lambda_3)}{\partial\lambda_3^2}.\end{aligned} \tag{3.3.17}$$

The bootstrapped finite sample and asymptotic distributions of the GoF test statistic of $T_{PN}^{(.)}$ are studied in Section 3.4.

In the following, we propose a test procedure that will be supremum on the absolute value of the K&K test statistic with (t_1, t_2) over $(-1, 1) \times (-1, 1)$. The reason behind proposing such a test is exemplified in Section 3.4. The mentioned GoF testing procedure for the K&K method is originally discussed in Feiyan Chen [8] for the univariate and bivariate geometric models. Besides, Feiyan Chen (2013) also discusses the K&K method for the multiple t-values for the GoF test for geometric models, c.f. page 12 of Chen[8] . However, in the present note, we are interested in proposing tests free from the choices of t-values; hence the advantages or disadvantages of considering multiple t-values are not discussed or illustrated in this note.

The GoF test statistic is

$$T_{SPN}^{(.)} = \sup_{(t_1,t_2)\in\{(-1,1)\times(-1,1)\}} \left| \frac{G_n(t_1,t_2) - G_.(t_1,t_2;\hat{\lambda_1},\hat{\lambda_3})}{\sigma^{(.)}} \right| \tag{3.3.18}$$

where $G_n(.)$, $G_.$ and $\sigma^{(.)}$ are defined in Section 3.3.2. Also, note that deriving the asymptotic distribution of the statistic $T_{SPN}^{(.)}$ is theoretically ambiguous. Hence, in Section 3.4 the finite sample distribution of the test statistic $T_{SPN}^{(.)}$ is analyzed.

3.3.3 GoF test free from alternative

In the class of distribution-free tests, the χ^2 test is commonly used even when there is no specific alternative hypothesis. However, this also needs to be improved in assessing the power of the test.

χ^2 GoF

In the following, we are using the classical χ^2 GoF test, and cell probabilities are computed up to k. The cell probability matrix is given by

$X - Y$	0	1	2	3	...	k+
0	p_{00}	p_{01}	p_{02}	p_{03}	...	$P(X=0)-\sum_{j=0}^{k-1} p_{0j}$
1	p_{10}	p_{11}	p_{12}	p_{13}	...	$P(X=1)-\sum_{j=0}^{k-1} p_{1j}$
2	p_{20}	p_{21}	p_{22}	p_{23}	...	$P(X=2)-\sum_{j=0}^{k-1} p_{2j}$
3	p_{30}	p_{31}	p_{32}	p_{33}	...	$P(X=3)-\sum_{j=0}^{k-1} p_{3j}$
...	...	...	...	...	...	...
k+	$P(Y=0)-\sum_{i=0}^{k-1} p_{i0}$	$P(Y=1)-\sum_{i=0}^{k-1} p_{i1}$	$P(Y=2)-\sum_{i=0}^{k-1} p_{i2}$	$P(Y=3)-\sum_{i=0}^{k-1} p_{i3}$	...	$1-\sum_{i=k}^{\infty}\sum_{j=k}^{\infty} p_{ij}$

where $p_{ij} = P(X = i, Y = j)$. The test statistic is

$$T_{\chi^2} = \sum_{i=0}^{k}\sum_{j=0}^{k} \frac{O_{i,j} - E_{i,j}}{E_{i,j}} \tag{3.3.19}$$

where k is the truncation point, $O_{i,j}$ is frequency of (i, j) observation in the data of size n and $E_{i,j} = nP(X = i, Y = j)$. Hence, with Pearson theorem T_{χ^2} follows a χ^2 distribution with $[(k + 1) \times (k + 1) - 1 - 3]$ degrees of freedom.

Similarly, the above two tests for the Sub-Models I & II can be derived with appropriate cell probabilities $p_{ij} = P(X = i, Y = j)$. In Section 3.4, we analyze a finite sample and large sample behavior of the above two test statistics.

3.4 Examples

3.4.1 Simulation

In the following, we give a general procedure to analyze the finite sample distribution of the GoF test statistics with bootstrapping technique.

Step 1 Simulate n observations from the bivariate pseudo-Poisson with fixed parameter values. Otherwise, estimate parameters by moment or m.l.e. method, say $\hat{\lambda}_i$. Then compute GoF test statistics, say T_{obs}.

Step 2 Fix the number of bootstrapping samples, say B (ideal size is 5,000,10,000) and sample $m(< n)$ observation from the above sample. , repeat Step 1 and compute $T^b_{m,obs}$ for $b \in \{1, 2, ..., B\}$.

Step 3 From the frequency distribution of $T^b_{m,obs}$ obtain the quantile values and the empirical p-value is $\frac{1}{B}\{\text{Total no. of } T^b_{m,obs} \text{ greater than } T_{obs}\}$.

K&K Method

In the following, we discuss the finite, large, and asymptotic distribution of the test statistics $T^{(.)}_{PN}(t_1, t_2)$ and $T^{(.)}_{SPN}$ (c.f. Section 3.3.2). Here, we limit our analysis to sub-models of the bivariate pseudo-Poisson model, and statistical inference or parameter estimation is well defined (see Section 7 in Arnold and Manjunath [4]). However, in sub-models I and II, the method of moments and the maximum likelihood estimators coincide. Hence, due to the invariance property of the maximum likelihood estimator, the defined test statistic asymptotically follows standard normal with a variance that will be inverse of the Fisher information matrix.

Now, we consider bootstrapping size of $B = 5,000$ with varying sample size of $n = 20, 30, 50, 100, 500$ at different $t_i = \pm 0.01, \pm 0.5, \pm 0.9$, $i = 1, 2$.

Sub-Model I: (i.e. $\lambda_2 = \lambda_3$) The corresponding quantile values and density plots refer to Table 3.1 and Figure 3.1, respectively.

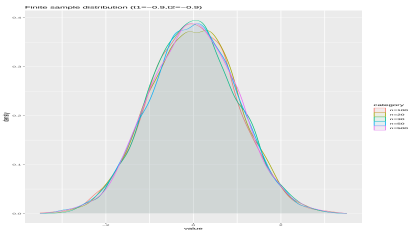
Sub-Model II: (i.e. $\lambda_2 = 0$) The corresponding quantile values and density plots refer to Table 3.2 and Figure 3.2, respectively.

According to the simulation study, it has been observed that whenever t is closer to zero, the empirical critical points are closer to the standard normal quantile values. It has been recommended that the t values be chosen either in the neighbourhood of zero or well-spanned in the interval $(-1, 1)$ to have consistency in the tests.

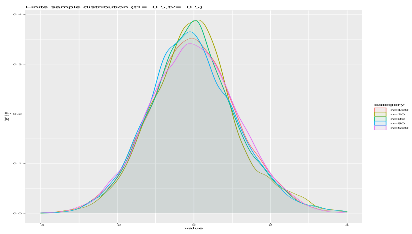
Note that from Table 3.1 & 3.2 and also from Figure 3.1 & 3.2 K&K-method finite sample distribution depends on the selected values for t. In particular, at $t_1 = -0.5(0.5)$ and $t_2 = -0.5(0.5)$ K & K statistic distributions are inconsistent. Hence, we consider the test statistic $T^{(.)}_{SPN}$(defined in (3.3.18)) such that the test support completely depends on the complete span of t-values. For an illustration of the proposed test, we are analyzing the finite sample distribution of the test statistic, which is computed with varying t_1 and t_2 from -0.99 to 0.99 at an increment of 0.01.

Finally, it has been argued in Feiyan Chen [8] that such tests are robust to the choice of alternatives and that the performance of the test is better than the K&K test because it also spanned the entire interval of $(-1, 1) \times (-1, 1)$.

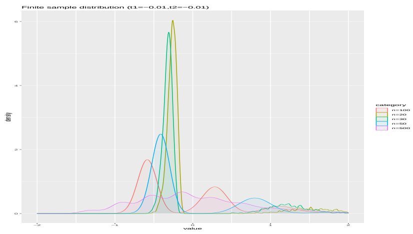
We refer to Table 3.3 and Figure 3.3 for the quantile values and frequency distribution of the test statistic, respectively. The test statistic's behavior is more stable and consistent for small and moderately large samples.



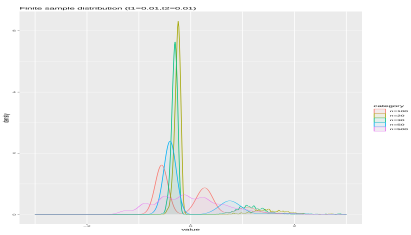
(a) $t_1 = -0.9, t_2 = -0.9$



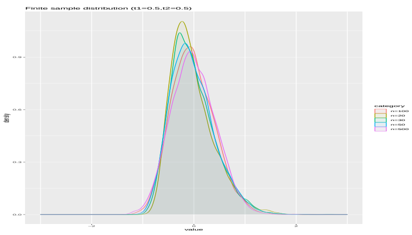
(b) $t_1 = -0.5, t_2 = -0.5$



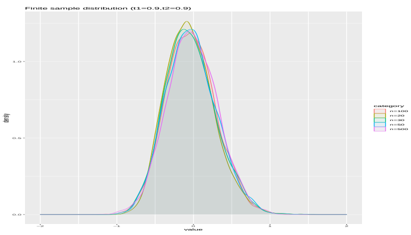
(c) $t_1 = -0.01, t_2 = -0.01$



(d) $t_1 = 0.01, t_2 = 0.01$

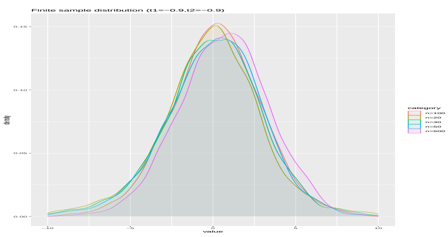


(e) $t_1 = 0.5, t_2 = 0.5$

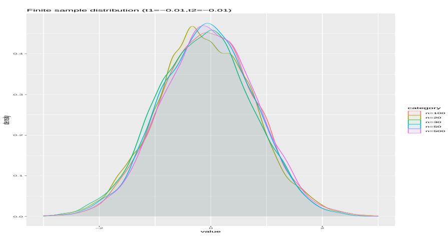


(f) $t_1 = 0.9, t_2 = 0.9$

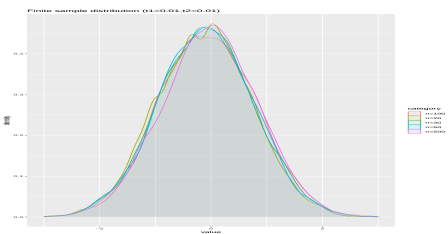
Figure 3.1: Finite sample distribution of $T_{PN}^{(SI)}$ for the Sub-Model I



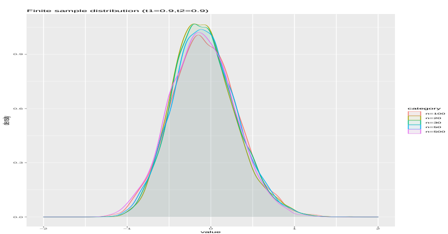
(a) $t_1 = -0.9, t_2 = -0.9$



(b) $t_1 = -0.01, t_2 = -0.01$

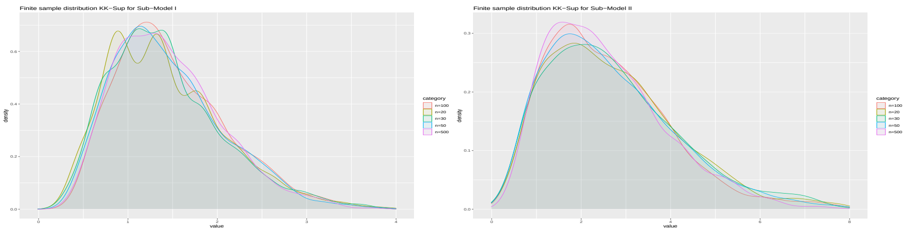


(c) $t_1 = 0.01, t_2 = 0.01$



(d) $t_1 = 0.9, t_2 = 0.9$

Figure 3.2: Finite sample distribution of $T_{PN}^{(SII)}$ for the Sub-Model II



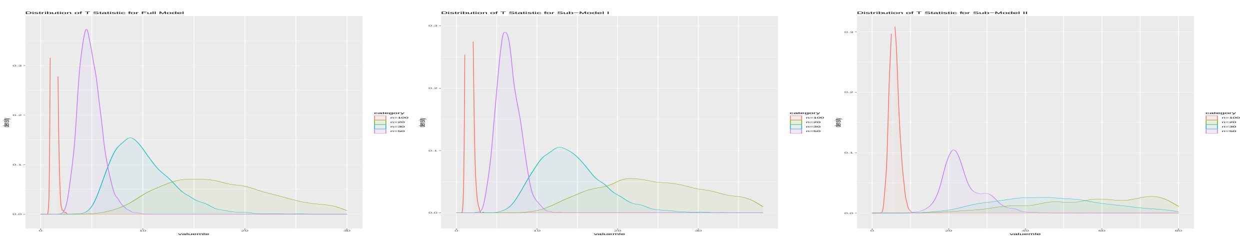
(a) $T^{(SI)}_{SPN}$ (b) $T^{(SII)}_{SPN}$

Figure 3.3: Finite sample distribution of the supremum of absolute deviation for the K&K-test statistic.

Muñoz and Gamero(M&G) method

Now, we consider GoF using p.g.f. (c.f. Muñoz and Gamero [29]) $T^{(.)}_{P,n,w}$ with depends on the underlying weight functions. We refer to 3.5, 3.6, 3.6 and Figure 3.4, 3.5, 3.6, 3.7, 3.8 for small and large sample distribution of the test statistic and its quantile values for the full and its sub-models.

To better understand the behavior of the test statistic, we examined the impact of different weights at $a1 = -0.9, -0.5, -0.01, 0.5, 3$ and $a2 = -0.9, -0.5, -0.01, 0.5, 5$ on the test statistic. According to the simulation study, the test is consistent and stable for moderately large sample sizes, irrespective of the weight chosen. Also, note that for the increasing sample size, the test statistic distribution is less variant and is shown to be consistent.



(a) (b) (c)

Figure 3.4: Example 1

Table 3.1: $T_{PN}^{(SI)}$ distribution for the Sub-Model I

		Sample size (0.5%, 2.5%, 5%; 95%, 97.5%, 99.5%)				
		$n = 20$	$n = 30$	$n = 50$	$n = 100$	$n = 500$
$T_{PN}^{(SI)}$	$t_1 = -0.9, t_2 = -0.9$	$(-2.503, -1.953, -1.652; 1.660, 1.912, 2.522)$	$(-2.447, -1.886, -1.626; 1.637, 1.981, 2.528)$	$(-2.582, -1.975, -1.669; 1.619, 1.928, 2.548)$	$(-2,622, -2.011, -1.674; 1.620, 1.912, 2.543)$	$(-2.582, -1.910, -1.612; 1.634, 1.948, 2.560)$
	$t_1 = -0.5, t_2 = -0.5$	$(-2.839, -2.187, -1.811; 1.953, 2.458, 3.196)$	$(-2.986, -2.168, -1.825; 2.031, 2.523, 3.390)$	$(-2.917, -2.189, -1.825; 1.927, 2.412, 3.342)$	$(-2.822, -2.198, -1.872; 1.988, 2.364, 3.106)$	$(-3.021, -2.156, -1.845; 1.947, 2.270, 3.078)$
	$t_1 = -0.01, t_2 = -0.01$	$(-0.464, -0.402, -0.376; 1.701, 1.976, 2.981)$	$(-0.519, -0.456, -0.428; 1.454, 1.688, 2.912)$	$(-0.593, -0.535, -0.510; 1.407, 2.036, 2.936)$	$(-0.747, -0.695, -0.668; 1.332, 1.734, 2.453)$	$(-1.388, -1.263, -0.997; 1.228, 1.553, 2.220)$
	$t_1 = 0.01, t_2 = 0.01$	$(-0.452, -0.392, -0.362; 1.620, 1.870, 2.456)$	$(-0.492, -0.440, -0.415; 1.352, 1.612, 2.853)$	$(-0.574, -0.520, -0.492; 1.231, 1.867, 2.456)$	$(-0.713, -0.665, -0.641; 1.240, 1.645, 2.347)$	$(-1.310, -1.216, -0.958; 1.223, 1.529, 2.049)$
	$t_1 = 0.5, t_2 = 0.5$	$(-0.751, -0.640, -0.581; 0.771, 0.977, 1.423)$	$(-0.821, -0.692, -0.619; 0.760, 0.966, 1.281)$	$(-0.863, -0.725, -0.634; 0.753, 0.9358, 1.304)$	$(-0.944, -0.753, -0.654; 0.710, 0.892, 1.238)$	$(-1.008, -0.792, -0.683; 0.707, 0.853, 1.111)$
	$t_1 = 0.9, t_2 = 0.9$	$(-0.750, -0.624, -0.541; 0.498, 0.625, 0.880)$	$(-0.733, -0.622, -0.541; 0.516, 0.630, 0.854)$	$(-0.776, -0.625, -0.543; 0.545, 0.643, 0.884)$	$(-0.833, -0.657, -0.550; 0.539, 0.654, 0.867)$	$(-0.829, -0.648, -0.552; 0.554, 0.669, 0.842)$

Table 3.2: $T_{PN}^{(SII)}$ distribution for the Sub-Model II

		Sample size (0.5%, 2.5%, 5%; 95%, 97.5%, 99.5%)				
		$n = 20$	$n = 30$	$n = 50$	$n = 100$	$n = 500$
$T_{PN}^{(SII)}$	$t_1 = -0.9, t_2 = -0.9$	$(-17.388, -8.192, -5.956; 5.168, 7.046, 13.129)$	$(-13.272, -6.980, -5.286; 4.862, 5.972, 10.505)$	$(-9.952, -6.382, -5.020; 4.653, 5.737, 7.871)$	$(-7.590, -5.342, -4.345; 4.612, 5.568, 7.522)$	$(-6.667, -4.788, -3.7800; 5.200, 5.930, 7.723)$
	$t_1 = -0.01, t_2 = -0.01$	$(-2.306, -1.700, -1.557; 1.347, 1.608, 2.216)$	$(-2.222, -1.803, -1.507; 1.303, 1.587, 2.116)$	$(-2.241, -1.726, -1.393; 1.264, 1.517, 2.023)$	$(-2.174, -1.679, -1.441; 1.341, 1.624, 2.230)$	$(-2.203, -1.690, -1.390; 1.362, 1.588, 2.173)$
	$t_1 = 0.01, t_2 = 0.01$	$(-2.301, -1.740, -1.432; 1.203, 1.549, 2.007)$	$(-2.215, -1.802, -1.440; 1.297, 1.558, 1.965)$	$(-2.176, -1.733, -1.480; 1.259, 1.568, 2.123)$	$(-2.261, -1.680, -1.434; 1.383, 1.618, 2.136)$	$(-2.219, -1.727, -1.436; 1.363, 1.624, 2.065)$
	$t_1 = 0.9, t_2 = 0.9$	$(-0.927, -0.737, -0.647; 0.580, 0.750, 1.014)$	$(-0.913, -0.755, -0.657; 0.572, 0.714, 0.986)$	$(-0.950, -0.784, -0.669; 0.590, 0.723, 0.981)$	$(-0.980, -0.825, -0.701; 0.594, 0.748, 1.012)$	$(-1.033, -0.853, -0.730; 0.549, 0.672, 0.906)$

Table 3.3: $T_{SPN}^{(SI)}$ distribution for the Sub-Model I

	Sample size (0.5%, 2.5%, 5%; 95%, 97.5%, 99.5%)				
	$n = 20$	$n = 30$	$n = 50$	$n = 100$	$n = 500$
$T_{SPN}^{(SI)}$	$(0.380, 0.450, 0.528; 2.647, 2.987, 3.005)$	$(0.337, 0.489, 0.586; 2.656, 3.012, 3.622)$	$(0.401, 0.511, 0.593; 2.623, 2.863, 3.567)$	$(0.423, 0.574, 0.643; 2.679, 3.013, 3.591)$	$(0.410, 0.550, 0.634; 2.558, 2.860, 3.279)$

Table 3.4: $T_{SPN}^{(SII)}$ distribution for the Sub-Model II

	Sample size (0.5%, 2.5%, 5%; 95%, 97.5%, 99.5%)				
	$n = 20$	$n = 30$	$n = 50$	$n = 100$	$n = 500$
$T_{SPN}^{(SII)}$	$(0.367, 0.676, 0.830; 6.868, 9.190, 24.875)$	$(0.422, 0.651, 0.771; 6.121, 7.016, 11.323)$	$(0.394, 0.620, 0.783; 5.559, 6.506, 9.039)$	$(0.423, 0.638, 0.793; 5.208, 6.221, 7.612)$	$(0.487, 0.701, 0.850; 5.113, 5.588, 6.720)$

Table 3.5: Example 1

		Sample size (0.5%, 2.5%, 5%; 95%, 97.5%, 99.5%)			
		$n = 20$	$n = 30$	$n = 50$	$n = 100$
$T_{P,n,w}$	Full Model	(7.085, 8.664, 9.632; 30.382, 33.138, 38.773)	(5.117, 5.793, 6.190; 15.599, 16.820, 20.041)	(2.666, 3.002, 3.230; 6.908, 7.490, 8.600)	(0.885, 0.960, 1.005; 1.635, 1.720, 1.911)
	Sub Model I	(10.808, 12.772, 13.980; 40.614, 42.949, 49, 598)	(6.560, 7.770, 8.455; 21.579, 23.463, 28.388)	(3.535, 4.008, 4.346; 8.914, 9.507, 10.722)	(0.958, 1.07, 1.126; 2.093, 2.217, 2.513)
	Sub Model II	(21.270, 29.847, 35.064; 147.186, 164.792, 197.586)	(16.463, 21.968, 24.682; 78.691, 86.521, 103.263)	(12.906, 15.543, 16.812; 33.672, 36.303, 41.303)	(3.191, 3.582, 3.951; 7.981, 8.418, 9.304)

Table 3.6: Example 2

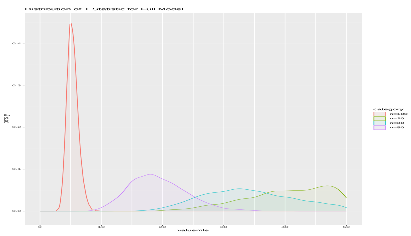
		Sample size (0.5%, 2.5%, 5%; 95%, 97.5%, 99.5%)			
		$n = 20$	$n = 30$	$n = 50$	$n = 100$
$T_{P,n,w}$	Full Model	(24.003, 29.695, 32.916; 109.461, 120.825, 143.666)	(18.688, 21.358, 23.045; 56.043, 60.840, 72.928)	(9.814, 11.380, 12.424; 18.230, 27.670, 33.967)	(3.340, 3.758, 3.964; 6.958, 7.362, 8.115)
	Sub Model I	(35.178, 41.764, 46.663; 153.635, 166.724, 205.881)	(21.543, 26.233, 28.611; 81.378, 90.566, 111.364)	(12.925, 14.971, 16.330; 36.434, 39.121, 44.605)	(3.735, 4.290, 4.546; 9.0899.648, 11.126)
	Sub Model II	(83.950, 120.662, 150.705; 750.712, 854.663, 1027.723)	(66.689, 94.637, 108.678; 398.689, 440.015, 540.523)	(51.414, 64.302, 72.087; 169.210, 184.953, 213.675)	(13.119, 15.157, 16.856; 38.858, 41.334, 46.675)

Table 3.7: $T_{P,n,w}^{(.)}$

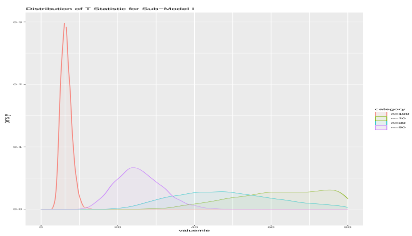
		Sample size (0.5%, 2.5%, 5%; 95%, 97.5%, 99.5%)			
		$n = 20$	$n = 30$	$n = 50$	$n = 100$
$T_{P,n,w}$	$(a_1 = -0.9, a_1 = -0.9)$	(30.918, 43.773, 54.367; 455.914, 527.123, 1023.235)	(25.182, 33.193, 38.630; 204.290, 414.503)	(13.956, 17.766, 20.112; 85.184, 103.305, 140.462)	(4.543, 5.389, 5.885; 18.364, 22.210, 29.855)
	$(a_1 = -0.01, a_1 = -0.01)$	(12.663, 18.047, 19.808; 69.013, 77.235, 97.099)	(9.193, 11.206, 12.488; 36.494, 39.859, 47.084)	(4.981, 5.839, 6.373; 15.345, 16.561, 19.420)	(1.774, 1.951, 2.060; 3.755, 3.994, 4.461)
	$(a_1 = 0.01, a_1 = 0.01)$	(10.923, 13.726, 15.786; 54.873, 61.512, 75.468)	(7.864, 9.473, 10.465; 29.302, 31.877, 37.684)	(4.216, 4.899, 5.315; 12.289, 13.215, 15.395)	(1.506, 1.645, 1.734; 3.037, 3.213, 3.570)
	$(a_1 = 0.5, a_1 = 0.5)$	(8.400, 10.410, 11.925; 40.708, 45.527, 54.879)	(5.935, 7.091, 7.752; 20.261, 21.926, 25.859)	(3.122, 3.598, 3.884; 8.470, 9.118, 10.522)	(1.120, 1.213, 1.271; 2.100, 2.210, 2.445)
	$(a_1 = 3, a_1 = 5)$	(3.066, 3.557, 3.851; 10.352, 11.398, 13.679)	(3.032, 3.353, 3.566; 6.451, 6.745, 7.330)	(1.480, 1.618, 1.677; 2.695, 2.832, 3.089)	(0.235, 0.259, 0.365; 0.533, 0.551, 0.584)
	$(a_1 = -0.9, a_1 = 5)$	(16.295, 26.760, 36.107; 235.847, 262.617, 326.237)	(14.523, 22.047, 27.541; 124.208, 137.808, 165.083)	(0.213, 12.853, 15.190; 52.653, 57.075, 67.911)	(8.825, 0.945, 1.028; 8.050, 9.502, 12.366)

Table 3.8: Distribution of the $FI_n^{(.)}$

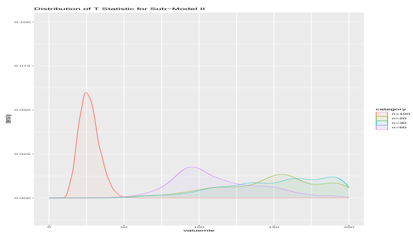
		Sample size (0.5%, 2.5%, 5%; 95%, 97.5%, 99.5%)				
		$n = 20$	$n = 30$	$n = 50$	$n = 100$	$n = 500$
$FI_n^{(.)}$	Full Model	(−7.818, −5.701, −4.665; 5.033, 6.900, 11.231)	(−7.760, −5.762, −4.698; 5.084, 6.703, 10.598)	(−7.166, −5.530, −4.748; 5.034, 6.544, 9.666)	(−7.167, −5.530, −4.748; 5.034, 6.544, 9.666)	(−6.897, −5.676, −4.841; 5.047, 6.063, 8.20)
	Sub Model I	(−9.537, −7.749, −6.766; 9.093, 11.852, 17.648)	(−9.947, −8.140, −7.131; 8.825, 11.500, 15.842)	(−10.477, −8.560, −7.505; 8.725, 10.887, 15.936)	(−11.018, −8.610, −7.507; 9.203, 11.223, 15.689)	(−11.759, −9.034, −7.802; 8.365, 10.293, 13.741)
	Sub Model II	(−15.205, −12.628, −11.181; 13.710, 18.740, 29.800)	(−15.917, −13.277, −11.695; 13.924, 17.255, 28.748)	(−17.191, −14.088, −12.162; 13.748, 17.947, 27.665)	(−18.434, −15.108, −13.076; 12.997, 15.976, 21.976)	(−20.647, −16.782, −14.478; 11.629, 14.497, 21.093)



(a)

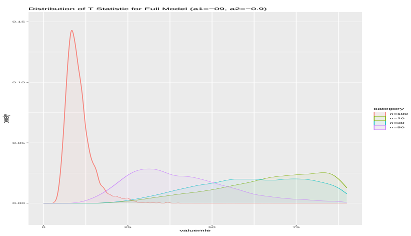


(b)

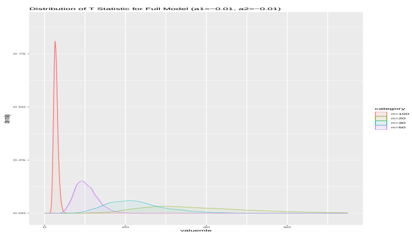


(c)

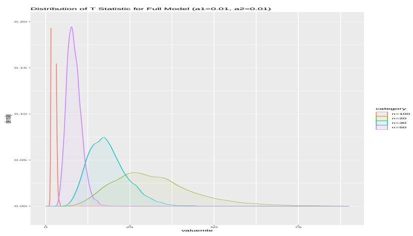
Figure 3.5: Example 2



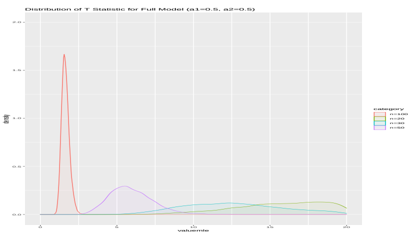
(a)



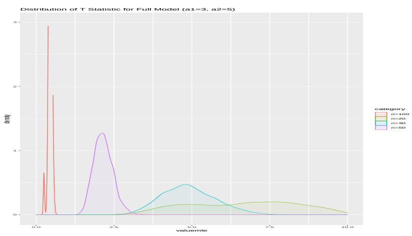
(b)



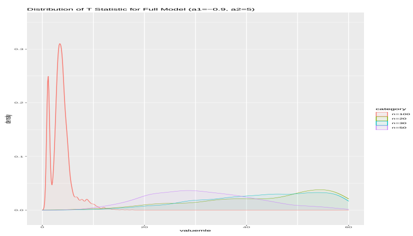
(c)



(d)

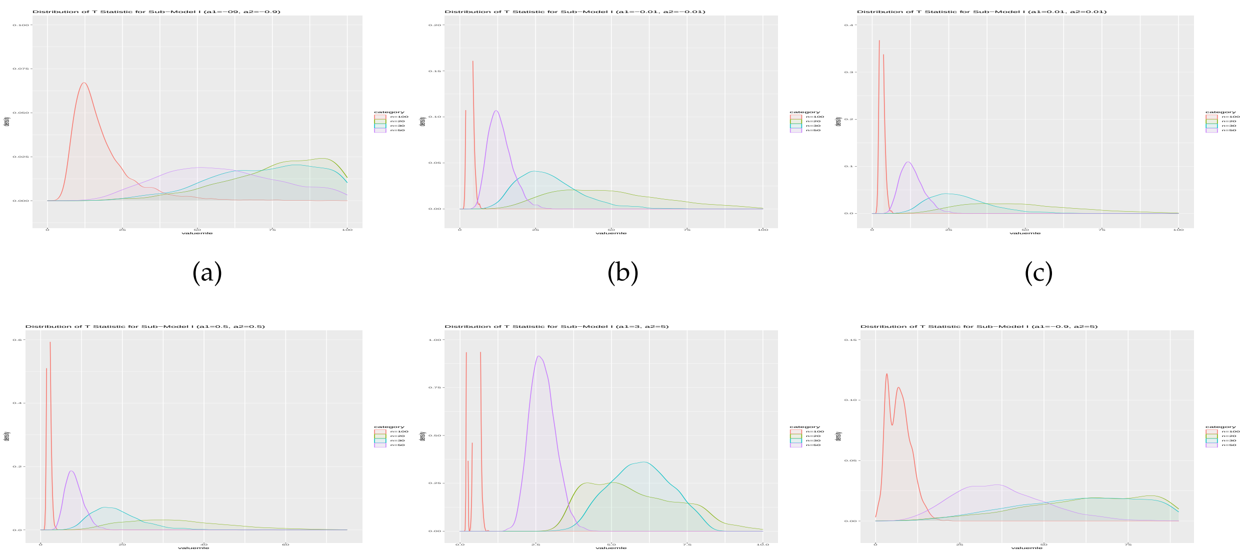


(e)



(f)

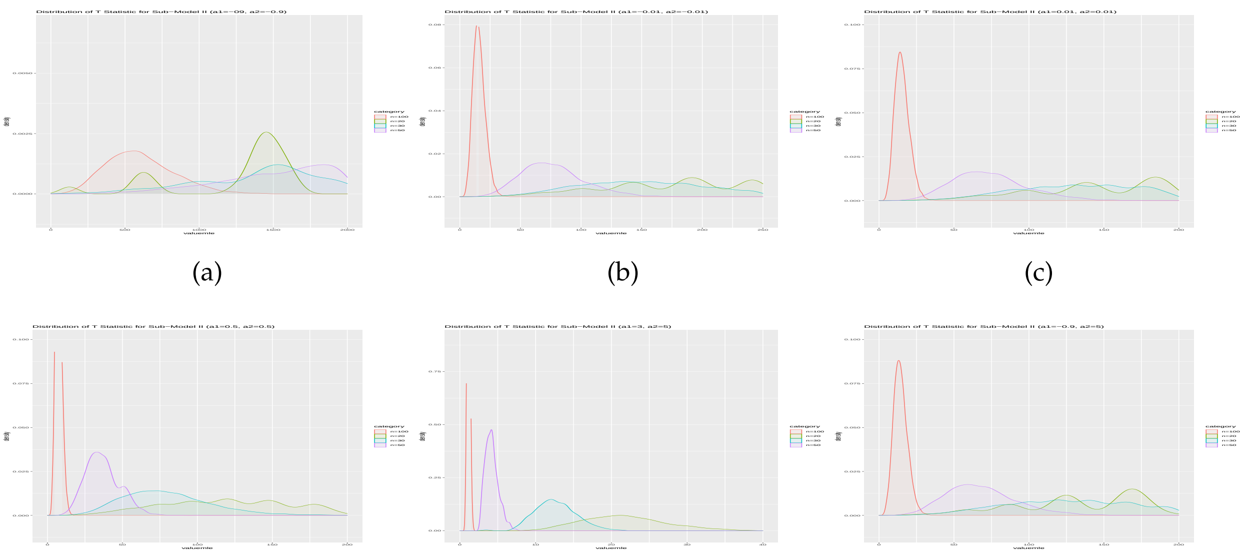
Figure 3.6: $T_{P,n,w}^{(.)}$



(a) (b) (c)

(d) (e) (f)

Figure 3.7: $T^{(SI)}_{P,n,w}$



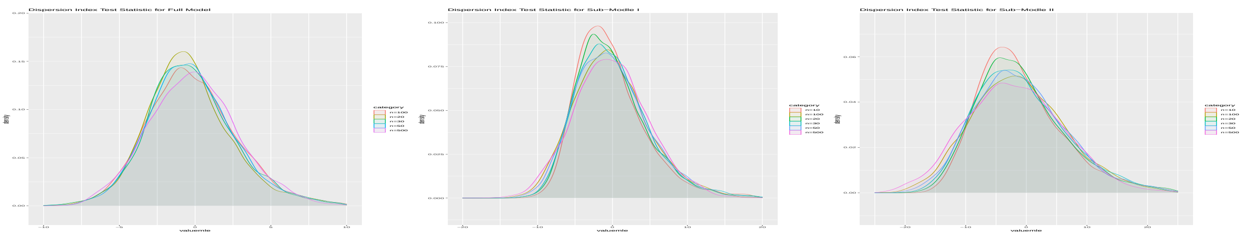
(a) (b) (c)

(d) (e) (f)

Figure 3.8: $T^{(SII)}_{P,n,w}$

Test based on moments

In the current section, we will analyze the non-consistent test defined in Section 3.3.1. The finite sample distribution of the $FI_n^{(.)}$, see Table 3.8 and Figure 3.9 for the distribution and its quantile values for the full and its sub-models. The simulation study clearly shows that the distribution of test statistics is shown to be standard normal behavior for increasing sample size. In addition, we note that for small and moderately large sample sizes, the test is conducted to be stable and consistent.

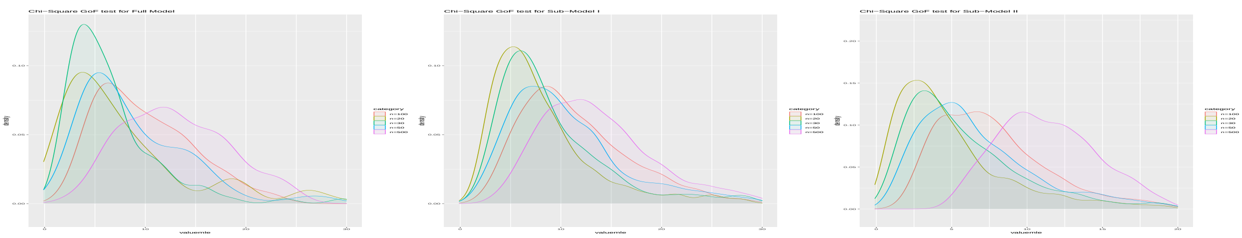


(a) Full Model (b) Sub-Model I (c) Sub-Model II

Figure 3.9: Distribution of the $FI_n^{(.)}$

GoF test free from alternative

The Chi-square GoF test statistic sample distribution for the full and its sub-models, see Figure 3.10. However, in the case when alternative just the negotiation of the null hypothesis distribution information, the Chi-square GoF test is recommended; otherwise, other tests which are mentioned perform better than the Chi-square. Also, the Chi-square test depends on the value of k chosen. For illustration, we have considered $K = 4$ and analyzed its finite and large sample distributions.



(a) Full Model (b) Sub-Model I (c) Sub-Model II

Figure 3.10: Chi-square GoF test for $k = 4$.

Power analysis

In the present section, we will be considering classical bivariate Poisson and bivariate Conway-Maxwell Poisson distributions as alternatives to analyze the power of each of the tests discussed above.

Hence, we simulate $n = 20, 30, 50, 100, 500$ samples from $Z_i \sim Poisson(\theta_i), i = 1, 2, 3$ and taking $U = Z_1 + Z_3$ and $V = Z_2 + Z_3$ the resultant joint random variable (U, V) will be n observations from the classical bivariate Poisson distribution. Nevertheless, to simulate $n = 20, 30, 50, 100, 500$ samples from the bivariate Conway-Maxwell Poisson, we begin with simulating an observation from the univariate Conway-Maxwell Poisson with parameter θ and ν, say N. Further, simulate N observations from the bivariate binomial distribution with specified cell probabilities, say $(W_{1i}, W_{2i}), i = 1, 2, ..., N$. Then, the random vector $(\sum_{i=1}^{N} W_{1i}, \sum_{i=1}^{N} W_{2i})$ will be an observation from the bivariate Conway-Maxwell Poisson distribution. For the desired sample size, repeat the above procedure for n times to have a specified sample size from the bivariate Conway-Maxwell Poisson distribution. We refer to Sellers et al. [33] for further discussion and an algorithm to simulate from the bivariate Conway-Maxwell Poisson using R software.

The empirical power computation is as follows

Step 1 Compute GoF test statistic value for the samples from alternative distribution, say T_{obs}.

Step 2 For the given bootstrapping size (say$B = 5000$), compute T_A^b for $b \in \{1, 2, ..., B\}$.

Step 3 Hence, $\frac{1}{B}\{$Total no. of $T_{m,obs}^b$ greater than $T_{obs}\}$ is an empirical power of the test.

We refer to Table 3.9 for the each of the tests empirical powers.

Table 3.9: Power (% of observations) under classical bivariate Poisson (BCBP(($\theta_1 = 1$, $\theta_2 = 3$, $\theta_3 = 4$))) and bivariate Con-Max Poisson (BCMP($\theta = 1$, $\nu = 5$, $\mu_1 = 0.1$, $oratio = \exp(1.5)$)) alternatives

		Sample size				
		$n = 20$	$n = 30$	$n = 50$	$n = 100$	$n = 500$
$T_{PN}^{(SII)}$	$t_1 = -0.9, t_2 = -0.9$	$(17.2, 8.8)$	$(21.6, 4.2)$	$(22.6, 2.6)$	$(59.2, 0.03)$	$(90.4, 0.01)$
	$t_1 = -0.5, t_2 = -0.5$	$(0.92, 82.4)$	$(0.89, 84.0)$	$(0.91, 92.4)$	$(0.93, 0.95.8)$	$(0.99, 0.97)$
	$t_1 = -0.01, t_2 = -0.01$	$(0.99, 0.91)$	$(0.97, 0.93)$	$(0.99, 0.99)$	$(0.99, 0.97)$	$(0.99, 0.99)$
	$t_1 = 0.01, t_2 = 0.01$	$(0.99, 0.99)$	$(0.99, 0.98)$	$(0.98, 0.97)$	$(0.99, 0.99)$	$(0.99, 0.98)$
	$t_1 = 0.5, t_2 = 0.5$	$(0.99, 0.99)$	$(0.91, 0.99)$	$(0.92, 0.99)$	$(0.89, 0.99)$	$(0.81, 0.99)$
	$t_1 = 0.9, t_2 = 0.9$	$(0.98, 0.99)$	$(0.92, 0.93)$	$(0.94, 0.96)$	$(0.99, 0.99)$	$(0.99, 0.99)$
$T_{SPN}^{(SII)}$	—	$(0.95, 0.92)$	$(0.99, 0.91)$	$(0.95, 0.98)$	$(0.99, 0.99)$	$(0.99, 0.97)$
$T_{P,n,w}^{(\cdot)}$	$a_1 = -0.9, a_2 = -0.9$	$(35.2, 67.8)$	$(56.9, 82.7)$	$(59.9, 12.6)$	$(35.0, 9.2)$	$--$
	$a_1 = -0.01, a_2 = -0.01$	$(6.0, 85.4)$	$(72.5, 27.4)$	$(24.5, 15.2)$	$(10.1, 50.3)$	$--$
	$a_1 = 0.01, a_2 = 0.01$	$(79.5, 2.6)$	$(20.3, 16.0)$	$(36.3, 34.2)$	$(6.1, 59.8)$	$--$
	$a_1 = 0.5, a_2 = 0.5$	$(19.2, 37.3)$	$(4.5, 27.0)$	$(26.1, 85.4)$	$(2.1, 55.5)$	$--$
	$a_1 = 3, a_2 = 5$	$(35.0, 74.0)$	$(72.7, 19.0)$	$(5.5, 25.5)$	$(1.2, 10.0)$	$--$
	$a_1 = -0.9, a_2 = 5$	$(12.0, 91.0)$	$(10.8, 72.0)$	$(4.7, 92.8)$	$(0.1, 43.6)$	$--$
$T_{P,n,w}^{(SI)}$	$a_1 = -0.9, a_2 = -0.9$	$(79.3, 96.5)$	$(12.3, 95.0)$	$(11.2, 21.8)$	$(13.4, 34.6)$	$--$
	$a_1 = -0.01, a_2 = -0.01$	$(73.5, 20.4)$	$(16.9, 9.5)$	$(0.9, 87.0)$	$(0.1, 13.2)$	$--$
	$a_1 = 0.01, a_2 = 0.01$	$(22.4, 3.8)$	$(15.9, 96.4)$	$(2.2, 93.8)$	$(13.8, 17.4)$	$--$
	$a_1 = 0.5, a_2 = 0.5$	$(43.2, 87.1)$	$(27.5, 4.8)$	$(21.3, 80.2)$	$(1.1, 69.4)$	$--$
	$a_1 = 3, a_2 = 5$	$(40.8, 43.2)$	$(36.0, 94.7)$	$(2.1, 75.8)$	$(1.3, 10.2)$	$--$
	$a_1 = -0.9, a_2 = 5$	$(54.4, 87.2)$	$(62.7, 72.0)$	$(4.7, 36.8)$	$(0.1, 24.4)$	$--$
$T_{P,n,w}^{(SII)}$	$a_1 = -0.9, a_2 = -0.9$	$(7.1, 12.9)$	$(41.9, 99.9)$	$(7.7, 26.2)$	$(60.71, 64.2)$	$--$
	$a_1 = -0.01, a_2 = -0.01$	$(71.4, 74.7)$	$(9.6, 56.6)$	$(2.9, 31.6)$	$(0.2, 36.5)$	$--$
	$a_1 = 0.01, a_2 = 0.01$	$(8.1, 36.1)$	$(2.5, 54.4)$	$(7.5, 67.1)$	$(0.1, 37.3)$	$--$
	$a_1 = 0.5, a_2 = 0.5$	$(38.9, 18.9)$	$(18.2, 99.6)$	$(5.4, 78.7)$	$(0.4, 96.2)$	$--$
	$a_1 = 3, a_2 = 5$	$(2.6, 70.0)$	$(2.5, 1.8)$	$(0.1, 82.3)$	$(0.0, 6.0)$	$--$
	$a_1 = -0.9, a_2 = 5$	$(50.6, 70.6)$	$(55.0, 36.4)$	$(4.4, 84.4)$	$(0.0, 13.2)$	$--$
$FI_n^{(\cdot)}$	—	$(58.4, 34.6)$	$(46.1, 86.7)$	$(56.0, 22.9)$	$(11.8, 42.4)$	$(63.5, 20.8)$
$FI_n^{(SI)}$	—	$(76.9, 13.5)$	$(6.3, 90.6)$	$(83.0, 40.6)$	$(41.7, 30.9)$	$(88.1, 73.8)$
$FI_n^{(SII)}$	—	$(69.9, 94.4)$	$(86.4, 70.2)$	$(95.0, 26.6)$	$(91.0, 46.7)$	$(100, 33.8)$

According to the power analysis, all tests are effective or significant in identifying the pseudo-Poisson and Conway-Maxwell Poisson distributions. When compared to the classical bivariate Poisson, tests are moderately consistent in detecting the true population. We conclude that one needs to think about altering the parameter values and conducting additional research on the same better to grasp the power of the classical bivariate Poisson alternative.

3.4.2 Real-life data

In the following section, we consider two data sets which are mentioned in Karlis and Tsiamyrtzis [21], Islam and Chowdhury [19], Leiter and Hamdani [24] and also in Arnold and Manjunath [4]. For empirical p-value computation, we have simulated 5000 observations from the pseudo-Poisson models with respective maximum likelihood values and compared them with the critical value of each test.

3.4.3 A particular data set I

We consider a data sets which is mentioned in Islam and Chowdhury [19] and also in Arnold and Manjunath [4]; the source of the data is from the tenth wave of the Health and Retirement Study (HRS). The data represents the number of conditions that one ever had (X) as mentioned by the doctors and, let (Y) denote the utilization of healthcare services (say, hospital, nursing home, doctor and home care). The Pearson correlation coefficient between X and Y is 0.063. The test for independence, classical inference (m.l.e and moment estimates), and AIC values for full model and its sub-models c.f. Arnold and Manjunath [4] page 16 and 18 (Table 10).

In the following, we will consider the full model and its sub-model II. The criteria for selecting below two models are discussed in Arnold, and Manjunath [4] on page 18 and Table 10. We refer to Table 3.10 for the critical values and its empirical p-values for the full model and sub-model II.

Table 3.10: Health and retirement study data (Full Model) and m.l.e. estimates Full model ($\hat{\lambda}_1 = 2.643, \hat{\lambda}_2 = 0.688, \hat{\lambda}_3 = 0.031$) for Sub-Model II ($\hat{\lambda}_3 = 0.031$)

		$n = 5567$	
		Test statistic value	p-value
$T_{PN}^{(SII)}$	$t_1 = -0.9, t_2 = -0.9$	151.734	0.025
	$t_1 = -0.5, t_2 = -0.5$	−2870.383	0.891
	$t_1 = -0.01, t_2 = -0.01$	755.821	0.901
	$t_1 = 0.01, t_2 = 0.01$	803.119	0.921
	$t_1 = 0.5, t_2 = 0.5$	1713.7	0.141
	$t_1 = 0.9, t_2 = 0.9$	3710.615	0.164
$T_{SPN}^{(SII)}$	—	12.740	0.097
$T_{PN}^{(.)}$	$a_1 = -0.9, a_2 = -0.9$	578.674	0.01
	$a_1 = -0.5, a_2 = -0.55$	117.940	0.99
	$a_1 = -0.01, a_2 = -0.01$	64.179	0.8
	$a_1 = 1, a_2 = 1$	67.564	0.09
	$a_1 = 3, a_2 = 5$	21.739	0.12
	$a_1 = -0.9, a_2 = 5$	23.830	0.02
$T_{PN}^{(SII)}$	$a_1 = -0.9, a_2 = -0.9$	659.816	0.99
	$a_1 = 1, a_2 = 1$	71.881	0.07
	$a_1 = -0.9, a_2 = 5$	24.465	0.02
$FI_n^{(.)}$	—	−13.532	0.987
$FI_n^{(SII)}$	—	−25.729	0.991
Chi-square $^{(.)}$	—	417.653	−−

The tests $T_{PN}^{(SII)}$ on neighbourhood of 0, $T^(SII)_{SPN}$, $T_{PN}^{(.)}$ (large than -0.9), $FI_n^{(.)}$ and $FI_n^{(SII)}$ are suggests that the Health and Retirement data fits bivariate pseudo-Poisson Full model and its Sub-Model II, which agree with the AIC values listed on pages 16 & 18 of Arnold and Manjunath's [4].

3.4.4 A particular data set II

Now, we consider a data set which is in Leiter and Hamdani [24]; the source of the data is a 50-mile stretch of Interstate 95 in Prince William, Stafford, and Spotsylvania counties in Eastern Virginia. The data represents the number of accidents categorized as fatal accidents, injury accidents, or property damage accidents (X), along with the corresponding number of fatalities and injuries (Y) for the period 1 January 1969 to 31 October 1970. For classical inference (m.l.e and moment estimates) and AIC values for full model and its sub-models c.f. Arnold and Manjunath [4] page 17 and 19 (Table 11). The criteria for selecting below two models are discussed in Arnold, and Manjunath [4] on page 19 and Table 11. It has been emphasized in Leiter and Hamdani [24] and Arnold and Manjunath [4] that mirrored sub-model II fits the data better than any other sub-models.

In the following, we will consider the two models. We refer to Table 3.11 for the critical values and its empirical p-values for the full model and Mirrored sub-model II.

Table 3.11: Accidents and fatalities (Full Model) and m.l.e. estimates Full model ($\hat{\lambda}_1 = 0.058, \hat{\lambda}_2 = 0.812, \hat{\lambda}_3 = 0.867$) and for mirrored Sub-Model II ($\hat{\lambda}_1 = 0.862, \hat{\lambda}_3 = 0.067$)

		$n = 639$	
		Test statistic value	p-value
Mirrored $T_{PN}^{(SII)}$	$t_1 = -0.9, t_2 = -0.9$	165.966	0.054
	$t_1 = -0.5, t_2 = -0.5$	-359.286	0.932
	$t_1 = -0.01, t_2 = -0.01$	-135.242	0.914
	$t_1 = 0.01, t_2 = 0.01$	-133.630	0.899
	$t_1 = 0.5, t_2 = 0.5$	-126.924	0.763
	$t_1 = 0.9, t_2 = 0.9$	-220.890	0.558
Mirrored $T_{SPN}^{(SII)}$	—	4.237	0.544
$T_{PN}^{(\cdot)}$	$a_1 = -0.9, a_2 = -0.9$	1057.191	0.99
	$a_1 = -0.5, a_2 = -0.55$	100.903	0.98
	$a_1 = -0.01, a_2 = -0.01$	24.906	0.87
	$a_1 = 1, a_2 = 1$	3.786	0.91
	$a_1 = 3, a_2 = 5$	1.178	0.01
	$a_1 = -0.9, a_2 = 5$	152.5798	0.007
Mirrored $T_{PN}^{(SII)}$	$a_1 = -0.9, a_2 = -0.9$	78.337	0.40
	$a_1 = 1, a_2 = 1$	4.049	0.91
	$a_1 = -0.9, a_2 = 5$	1.438	0.20
$FI_n^{(\cdot)}$	—	2.289	0.986
Mirrored $FI_n^{(SII)}$	—	3.443	0.3
Chi-square $^{(\cdot)}$	—	586	−−

The tests $T_{PN}^{(SII)}$ on a neighborhood of 0, $T_{SPN}^{(SII)}$, $T_{PN}^{(.)}$ (large than -0.9), $FI_n^{()}$ and $FI_n^{(SII)}$ suggests that the Accidents and Fatalities data fits well the bivariate pseudo-Poisson Full model and its mirrored Sub-Model II, which is in line with the AIC values listed on pages 16 & 18 of Arnold and Manjunath's [4]. A detailed discussion of the analysis is discussed in the conclusion chapter.

CHAPTER 4

Bayesian Inference for pseudo-Exponential data

4.1 Introduction

Bivariate conditionally specified models frequently offer valuable, adaptable models with a range of dependence patterns. In such cases, consideration should be given to what is recognized as pseudo-exponential models (According to Filus and Filus [12]-[14]). According to Cacoullos [7], the distributions with multiple parameters of the exponential family are characterized to obtain prior knowledge about the posterior densities. We create conjugate priors for discrete exponential families examined for the count data model in J. P. Chour [10]. The bivariate pseudo-exponential approaches have exponential first marginal and exponential conditional distributions for the second variable. We start by reviewing the pseudo-exponential development, focusing on a few simplified sub-models. The Bayesian inference for such sub-models is again explained in detail. For traditional inference issues and an illustration of uses of the bivariate pseudo-exponential distribution, we refer to Arnold and Arvanitis [2], and also we refer Arnold et al. [1].

4.2 The Bivariate Pseudo-Exponential Distributions

Let X and Y be a random variable with $X > 0$ and $Y > 0$ that could be used as a model for the lifespans of connected system components. Inside one method, the joint density is specified as a specific non-negative function on $(0, \infty) \times (0, \infty)$ that integrates to 1. In contrast, Filus and Filus[13] suggest using one marginal distribution (let's say X) and the family of conditional densities of the other variable (Y) given X=x, for every x, in various articles, such as Filus and Filus[12]. Therefore, the joint density of (X, Y) will have the following form, if h(x) denotes the density of X for each x, and $g_x(y)$ indicates the conditional density of Y given that X=x, then

$$f_{X,Y}(x, y) = h(x)g_x(y)I(x > 0, y > 0). \tag{4.2.1}$$

In general, the Pseudo-exponential distributions (inside the context of Filus and Filus) the equivalent to the case where h and g_x's have exponential densities. As a result, the joint density of a bivariate pseudo-Exponential distribution is one where

$$X \sim Exp(\theta_1), \tag{4.2.2}$$

and for each $x > 0$,

$$Y|X = x \sim Exp(\theta(x)), \tag{4.2.3}$$

where $\theta_1 > 0$ and $\theta(x) > 0 \; \forall$ x. Except for measurability, there are no constraints on the structure of the function $\theta(x)$. The corresponding joint density is given by

$$f_{X,Y}(x, y) = \theta_1 e^{-\theta_1 x} \theta(x) e^{-\theta(x) y} I(x > 0, y > 0). \tag{4.2.4}$$

where the positive function $\theta(x)$ is quite arbitrary. Suppose that $\theta(x) = \theta_2 + \theta_3 x$, then the corresponding joint p.d.f. is given by

$$f_{X,Y}(x, y) = \theta_1 e^{-\theta_1 x} (\theta_2 + \theta_3 x) e^{-(\theta_2 + \theta_3 x) y} I(x > 0, y > 0). \tag{4.2.5}$$

where $\theta_i > 0$, i=1,2,3.

4.2.1 Notes on priors for Pseudo-Exponential data

The joint density function for a bivariate Pseudo-Exponential distribution is of the form equation (4.2.5).

Consequently, the likelihood function for a sample of size n from this distribution is given by

$$L(\theta_1, \theta_2, \theta_3 | \underline{x}, \underline{y}) = \theta_1^n e^{-\theta_1 \sum_{i=1}^{n} x_i} (\prod_{i=1}^{n} (\theta_2 + \theta_3 x_i)) e^{-\theta_2 \sum_{i=1}^{n} y_i - \theta_3 \sum_{i=1}^{n} x_i y_i} \tag{4.2.6}$$

Note that this likelihood factors as follows:

$$L(\theta_1, \theta_2, \theta_3 | \underline{x}, \underline{y}) = \left\{ \theta_1^n e^{-\theta_1 \sum_{i=1}^{n} x_i} \right\} \left\{ (\prod_{i=1}^{n} (\theta_2 + \theta_3 x_i)) e^{-\theta_2 \sum_{i=1}^{n} y_i - \theta_3 \sum_{i=1}^{n} x_i y_i} \right\} \tag{4.2.7}$$

In the first factor only involves the parameter θ_1, while the second factor involves the parameters θ_2 and θ_3. We will call θ_1 the marginal parameter, and θ_2 and θ_3 will be called conditional parameters. This factorization will prove to be important in Bayesian inference for this model, as discuss below. Because of the factorization we will know that if a priori $\tilde{\theta_1}$ and $(\tilde{\theta_2}, \tilde{\theta_3})$ are independent, then they will be independent posteriori also.

Now, to the Pseudo-Exponential distribution defined in equation (4.2.5), it is simply assume a gamma prior for the marginal parameter $\tilde{\theta_1}$. i.e., a priori we assume that

$$\tilde{\theta_1} \sim \Gamma(\alpha_1, \beta_1), \tag{4.2.8}$$

where $0 < \alpha_1 < \infty$ and $0 < \beta_1 < \infty$.

Therefore, the priori density for $\tilde{\theta_1}$ is of the form

$$f_{\tilde{\theta_1}}(\theta_1) = \frac{\beta_1^{\alpha_1} \theta_1^{\alpha_1 - 1} e^{-\beta_1 \theta_1}}{\Gamma(\alpha_1)}; \qquad 0 < \theta_1 < \infty.$$

The choice of an a priori joint density of $(\tilde{\theta_2}, \tilde{\theta_3})$ that will be independent of $\tilde{\theta_1}$ is not so obvious. It is clear whether it is possible to choose such a prior that will be "Conjugate" with the second factor in equation (4.2.7). The joint posterior of $(\tilde{\theta_2}, \tilde{\theta_3})$ will need to be dealt with numerically.

4.3 Independent priors

Consider an a priori joint density in which all three parameters are independent, and each has a gamma distribution. Thus we have

$$f_{\tilde{\theta_1},\tilde{\theta_2},\tilde{\theta_3}}(\theta_1, \theta_2, \theta_3) = \prod_{i=1}^{3} \frac{\beta_i^{\alpha_i} \theta_i^{\alpha_i - 1} e^{-\beta_i \theta_i}}{\Gamma(\alpha_i)} \tag{4.3.1}$$

where $0 < \theta_i, \alpha_i, \beta_i < \infty$, for i=1,2,3.
The kernel of the posterior which is the product of the kernel of the factored likelihood (4.2.7) and the kernel of the prior (4.3.1) is of the form

$$\begin{aligned} ker(f_{\tilde{\theta_1},\tilde{\theta_2},\tilde{\theta_3}|\underline{X},\underline{Y}}(\theta_1, \theta_2, \theta_3|\underline{x}, \underline{y})) = &\left\{ \theta_1^{\alpha_1+n-1} e^{-\theta_1(\beta_1 + \sum_{i=1}^{n} x_i)} \right\} \\ &* \left\{ (\prod_{i=1}^{n} (\theta_2 + \theta_3 x_i)) \theta_2^{\alpha_2 - 1} \theta_3^{\alpha_3 - 1} e^{-\theta_2(\beta_2 + \sum_{i=1}^{n} y_i)} \right. \\ &\left. * \, e^{-\theta_3(\beta_3 + \sum_{i=1}^{n} x_i y_i)} \right\} \end{aligned} \tag{4.3.2}$$

From the first factor in equation (4.3.2) we recognise that a posteriori $\tilde{\theta_1}$ has a gamma distribution, i.e.,

$$\tilde{\theta_1}|\underline{X} = \underline{x}, \underline{Y} = \underline{y} \sim \Gamma(\alpha_1 + n, \beta_1 + \sum_{i=1}^{n} x_i) \tag{4.3.3}$$

The second factor is the kernel of the posterior density of $(\tilde{\theta_2}, \tilde{\theta_3})$. It will need to be dealt with numerically.

4.3.1 Sub-Model-I($\theta_2 = \theta_3$):

We first focus on the the sub-model of equation (4.2.5) obtained by equating θ_2 and θ_3. The model is given by

$$f_{X,Y}(x,y) = \theta_1 e^{-\theta_1 x}\theta_3(1+x)e^{-\theta_3(1+x)y}I(x>0,y>0). \tag{4.3.4}$$

The likelihood function for a sample of size n from this distribution is

$$L(\theta_1,\theta_3|\underline{x},\underline{y}) = \theta_1^n e^{-\theta_1\sum_{i=1}^n x_i}\theta_3^n(\prod_{i=1}^n(1+x_i))e^{-\theta_3(\sum_{i=1}^n y_i+\sum_{i=1}^n x_i y_i)}$$

The likelihood factors as follows:

$$L(\theta_1,\theta_3|\underline{x},\underline{y}) = \left\{\theta_1^n e^{-\theta_1\sum_{i=1}^n x_i}\right\}\left\{\theta_3^n(\prod_{i=1}^n(1+x_i))e^{-\theta_3(\sum_{i=1}^n y_i+\sum_{i=1}^n x_i y_i)}\right\} \tag{4.3.5}$$

For such a conjugate joint prior density for two parameters in the model, we can take one with independent gamma marginals. Thus

$$f_{\tilde{\theta_1},\tilde{\theta_3}}(\theta_1,\theta_3) = \frac{\beta_1^{\alpha_1}\theta_1^{\alpha_1-1}e^{-\beta_1\theta_1}}{\Gamma(\alpha_1)}.\frac{\beta_3^{\alpha_3}\theta_3^{\alpha_3-1}e^{-\beta_3\theta_3}}{\Gamma(\alpha_3)} \tag{4.3.6}$$

The kernel of the posterior density, which is the product of the kernel of the likelihood factors (4.3.5) and the kernel of the joint prior (4.3.6) is of the form

$$\begin{aligned} Ker(f_{\tilde{\theta_1},\tilde{\theta_3}|\underline{X},\underline{Y}}(\theta_1,\theta_3|\underline{x},\underline{y})) = &\left\{\theta_1^{\alpha_1+n-1}e^{-\theta_1(\beta_1+\sum_{i=1}^n x_i)}\right\} \\ &* \left\{\theta_3^{\alpha_3+n-1}e^{-\theta_3(\beta_3+\sum_{i=1}^n y_i+\sum_{i=1}^n x_i y_i)}\right\} \end{aligned} \tag{4.3.7}$$

A posteriori the two parameters have independent gamma distributions. Thus

$$\tilde{\theta_1}|\underline{X}=\underline{x},\underline{Y}=\underline{y} \sim \Gamma(\alpha_1+n,\beta_1+\sum_{i=1}^n x_i), \tag{4.3.8}$$

and, independently,

$$\tilde{\theta_3}|\underline{X} = \underline{x}, \underline{Y} = \underline{y} \sim \Gamma(\alpha_3 + n, \beta_3 + \sum_{i=1}^{n} y_i + \sum_{i=1}^{n} x_i y_i). \tag{4.3.9}$$

The squared error loss estimates of the parameters (the posterior means) are thus,

$$\hat{\theta_1}^{(B)} = \frac{\alpha_1 + n}{\beta_1 + \sum_{i=1}^{n} x_i}$$

and

$$\hat{\theta_3}^{(B)} = \frac{\alpha_3 + n}{\beta_3 + \sum_{i=1}^{n} y_i + \sum_{i=1}^{n} x_i y_i}$$

If we choose to an improper prior with $\alpha_1 = \alpha_3 = \beta_1 = \beta_3 = 0$, then the resulting Bayes estimates coincide with the corresponding maximum likelihood estimates(M.L.E's).

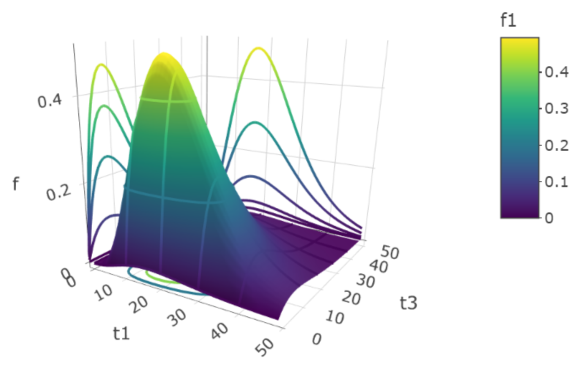


Figure 4.1: Density plot of independent gamma prior with hyper-parameter values are $\alpha_1 = 2, \alpha_3 = 4, \beta_1 = 2, \beta_3 = 3$ and parameter values $\theta_1 = 2, \theta_3 = 5$ with sample of size 30 of the sub-model-I.

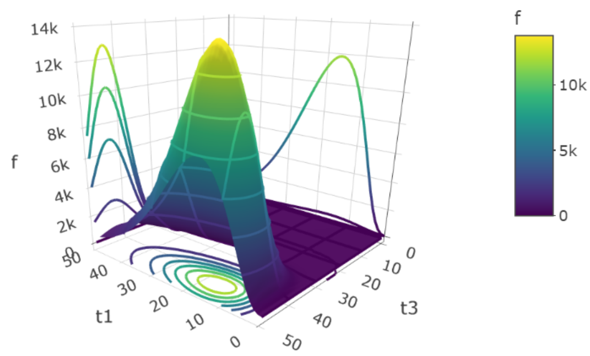


Figure 4.2: Density plot of posterior (independent gamma prior) with hyper-parameter values are $\alpha_1 = 2, \alpha_3 = 4, \beta_1 = 2, \beta_3 = 3$ and parameter values $\theta_1 = 2, \theta_3 = 5$ with sample of size 30 of the sub-model-I.

4.3.2 Sub-Model-II($\theta_2 = 0$):

Consider the sub-model-II obtained by setting $\theta_2 = 0$. The model is thus of the form

$$f_{X,Y}(x, y) = \theta_1 e^{-\theta_1 x} \theta_3 x e^{-\theta_3 xy} I(x > 0, y > 0). \tag{4.3.10}$$

The likelihood function for a sample of size n from this distribution is

$$L(\theta_1, \theta_3 | \underline{x}, \underline{y}) = \theta_1^n e^{-\theta_1 \sum_{i=1}^n x_i} \theta_3^n (\prod_{i=1}^n x_i) e^{-\theta_3 \sum_{i=1}^n x_i y_i}$$

The likelihood factors as follows:

$$L(\theta_1, \theta_3 | \underline{x}, \underline{y}) = \left\{ \theta_1^n e^{-\theta_1 \sum_{i=1}^n x_i} \right\} \left\{ \theta_3^n (\prod_{i=1}^n x_i) e^{-\theta_3 \sum_{i=1}^n x_i y_i} \right\} \tag{4.3.11}$$

If we use a prior with independent gamma marginals such as

$$f_{\tilde{\theta}_1, \tilde{\theta}_3}(\theta_1, \theta_3) = \frac{\beta_1^{\alpha_1} \theta_1^{\alpha_1 - 1} e^{-\beta_1 \theta_1}}{\Gamma(\alpha_1)} . \frac{\beta_3^{\alpha_3} \theta_3^{\alpha_3 - 1} e^{-\beta_3 \theta_3}}{\Gamma(\alpha_3)} \tag{4.3.12}$$

The kernel of the posterior density, which is the product of the kernel of the likelihood factors (4.3.11) and the kernel of the joint prior (4.3.12) is given by

$$\begin{aligned} Ker(f_{\tilde{\theta_1},\tilde{\theta_3}|\underline{X},\underline{Y}}(\theta_1,\theta_3|\underline{x},\underline{y})) = &\left\{\theta_1^{\alpha_1+n-1}e^{-\theta_1(\beta_1+\sum_{i=1}^n x_i)}\right\} \\ &* \left\{\theta_3^{\alpha_3+n-1}e^{-\theta_3(\beta_3+\sum_{i=1}^n x_i y_i)}\right\} \end{aligned} \tag{4.3.13}$$

A posteriori the two parameters have independent gamma distributions. Thus

$$\tilde{\theta_1}|\underline{X}=\underline{x},\underline{Y}=\underline{y} \sim \Gamma(\alpha_1+n,\beta_1+\sum_{i=1}^n x_i), \tag{4.3.14}$$

and, independently,

$$\tilde{\theta_3}|\underline{X}=\underline{x},\underline{Y}=\underline{y} \sim \Gamma(\alpha_3+n,\beta_3+\sum_{i=1}^n x_i y_i). \tag{4.3.15}$$

The squared error loss estimates of the parameters (the posterior means) are thus,

$$\hat{\theta_1}^{(B)} = \frac{\alpha_1+n}{\beta_1+\sum_{i=1}^n x_i}$$

and

$$\hat{\theta_3}^{(B)} = \frac{\alpha_3+n}{\beta_3+\sum_{i=1}^n x_i y_i}$$

If we choose to an improper prior with $\alpha_1 = \alpha_3 = \beta_1 = \beta_3 = 0$, then the resulting Bayes estimates coincide with the corresponding maximum likelihood estimates(M.L.E's).

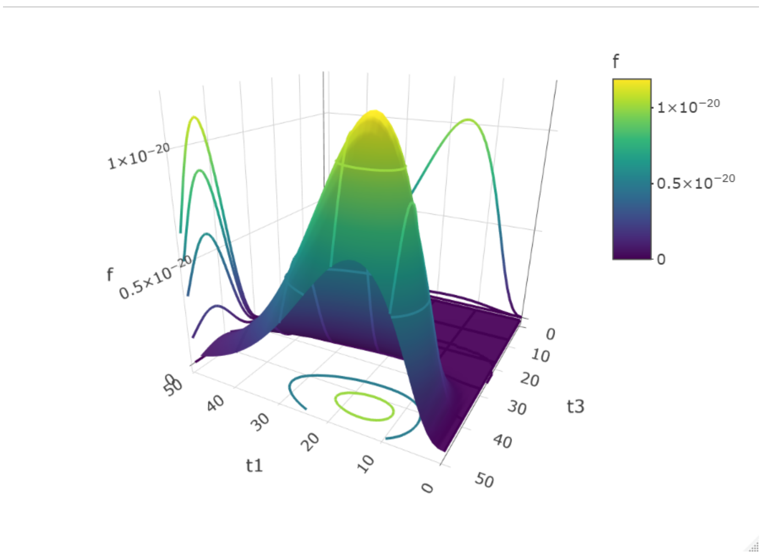


Figure 4.3: Density plot of posterior (independent gamma prior) with hyper-parameter values are $\alpha_1 = 2, \alpha_3 = 4, \beta_1 = 2, \beta_3 = 3$ and parameter values $\theta_1 = 2, \theta_3 = 5$ with sample of size 30 of the sub-model-II.

4.4 Pseudo-Gamma priors:-

In the following sections we will be discussing bivariate pseudo-gamma priors and their applications in more details for each of the sub-model I & II.

4.4.1 Sub-Model-I($\theta_2 = \theta_3$):

Consider the sub-model, specified in equation (4.3.1), the likelihood function for a sample of size n from this distribution is given by

$$L(\theta_1, \theta_3 | \underline{x}, \underline{y}) = \theta_1^n e^{-\theta_1 \sum_{i=1}^n x_i} \theta_3^n (\prod_{i=1}^n (1 + x_i)) e^{-\theta_3 (\sum_{i=1}^n y_i + \sum_{i=1}^n x_i y_i)} \tag{4.4.1}$$

This time we will consider a joint prior that is of the bivariate pseudo-gamma form. For it we assume that θ_3 has a $\Gamma(\tau_1, \psi_1)$ density, i.e.,

$$f(\theta_3) \propto \theta_3^{\tau_1 - 1} e^{-\theta_3 \psi_1} I(\theta_3 > 0), \quad 0 < \tau_1, \psi_1 < \infty.$$

and then for each value of θ_3, the conditional density θ_1 given θ_3 is assumed to be of the gamma form with an intensity parameter that is a linear function of θ_3. Thus

$$f(\theta_1|\theta_3) \propto (\psi_2 + \psi_3\theta_3)^{\tau_2}\theta_1^{\tau_2-1}e^{-(\psi_2+\psi_3\theta_3)\theta_1}I(\theta_1 > 0).$$

where $0 < \tau_2, \psi_3 < \infty$ and $0 \leq \psi_2 < \infty$.
The joint prior is thus of the form

$$f(\theta_1, \theta_3) \propto (\psi_2 + \psi_3\theta_3)^{\tau_2}\theta_1^{\tau_2-1}e^{-(\psi_2+\psi_3\theta_3)\theta_1}\theta_3^{\tau_1-1}e^{-\theta_3\psi_1}I(\theta_1 > 0, \theta_3 > 0). \tag{4.4.2}$$

Consider the simpler prior in which we assume that $\psi_2 = 0$. This prior density will be of the form

$$f_p(\theta_1, \theta_3) \propto \theta_1^{\tau_2-1}e^{-\theta_1\theta_3\psi_3}\theta_3^{\tau_1+\tau_2-1}e^{-\psi_1\theta_3}I(\theta_1 > 0, \theta_3 > 0). \tag{4.4.3}$$

The corresponding posterior density to a sample of size n from sub-model I will be

$$\begin{aligned} f(\theta_1, \theta_3|\underline{X} = \underline{x}, \underline{Y} = \underline{y}) &\propto (\theta_1^n e^{-\theta_1\sum_{i=1}^n x_i})(\theta_3^n e^{-\theta_3(\sum_{i=1}^n y_i + \sum_{i=1}^n x_i y_i})\\ &* (\theta_1^{\tau_2-1}e^{-\theta_1\theta_3\psi_3}\theta_3^{\tau_1+\tau_2-1}e^{-\psi_1\theta_3})\\ &\propto \theta_1^{\tau_2+n-1}e^{-\theta_1\sum_{i=1}^n x_i}\theta_3^{\tau_1+\tau_2-1}e^{-\theta_3(\psi_1+\sum_{i=1}^n y_i+\sum_{i=1}^n x_i y_i)}\\ &* e^{-\theta_1\theta_3\psi_3}\\ &\propto (\theta_3^{\tau_1+n-1}e^{-\theta_3(\psi_1+\sum_{i=1}^n y_i+\sum_{i=1}^n x_i y_i)})\\ &* (\theta_1^{\tau_2+n-1}e^{-\theta_1(\sum_{i=1}^n x_i+\psi_3\theta_3)}) \end{aligned} \tag{4.4.4}$$

The marginal posterior distributions of θ_1 and θ_3 are

$$f_{\theta_1}^p(\theta_1) \propto [\theta_1^{\tau_2+n-1}e^{-\theta_1\sum_{i=1}^n x_i}].\frac{\Gamma(\tau_1+n)}{(\psi_1+\sum_{i=1}^n y_i+\sum_{i=1}^n x_i y_i+\theta_1\psi_3)^{(\tau_1+n)}}, \tag{4.4.5}$$

and

$$f_{\theta_1}^p(\theta_1) \propto [\theta_3^{\tau_1+n-1}e^{-\theta_3(\psi_1+\sum_{i=1}^n y_i+\sum_{i=1}^n x_i y_i)}].\frac{\Gamma(\tau_2+n)}{(\sum_{i=1}^n x_i+\theta_3\psi_3)^{(\tau_2+n)}}. \tag{4.4.6}$$

For the plots of such prior and posterior densities, see Figures 4.4, 4.5, 4.6 (for priors) and Figures 4.7, 4.8, 4.9 (for posteriors).

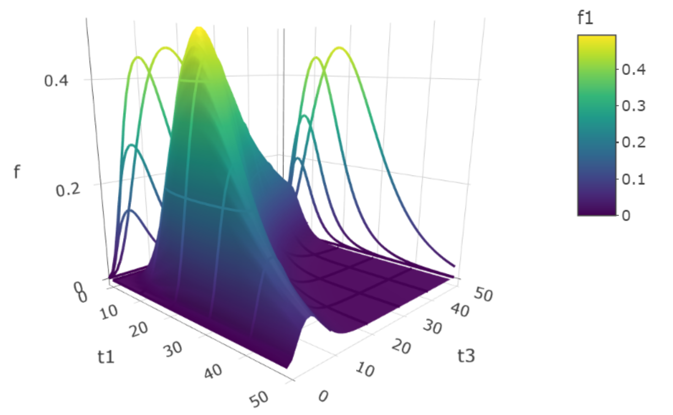


Figure 4.4: Density plot of pseudo-gamma prior with hyper-parameter values are $\tau_1 = 2, \tau_2 = 4, \psi_1 = 2, \psi_2 = 1(small), \psi_3 = 3$ and parameter values $\theta_1 = 2, \theta_3 = 5$ with sample of size 30 of the sub-model-I.

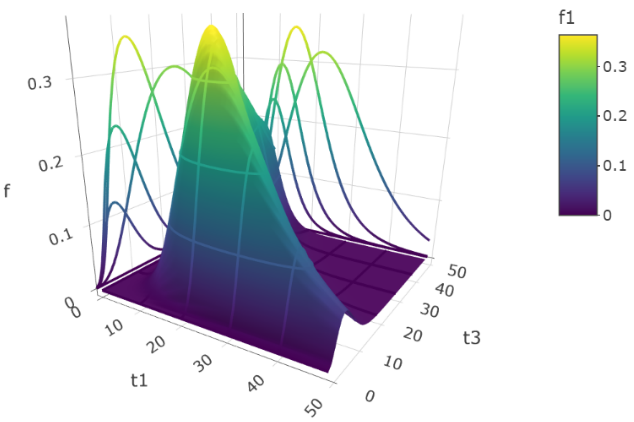


Figure 4.5: Density plot of pseudo-gamma prior with hyper-parameter values are $\tau_1 = 2, \tau_2 = 4, \psi_1 = 2, \psi_2 = 0(simple), \psi_3 = 3$ and parameter values $\theta_1 = 2, \theta_3 = 5$ with sample of size 30 of the sub-model-I.

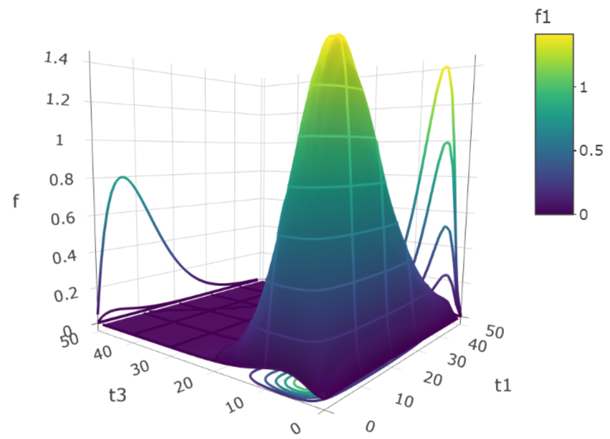


Figure 4.6: Density plot of pseudo-gamma prior with hyper-parameter values are $\tau_1 = 2, \tau_2 = 4, \psi_1 = 2, \psi_2 = 7(large), \psi_3 = 3$ and parameter values $\theta_1 = 2, \theta_3 = 5$ with sample of size 30 of the sub-model-I.

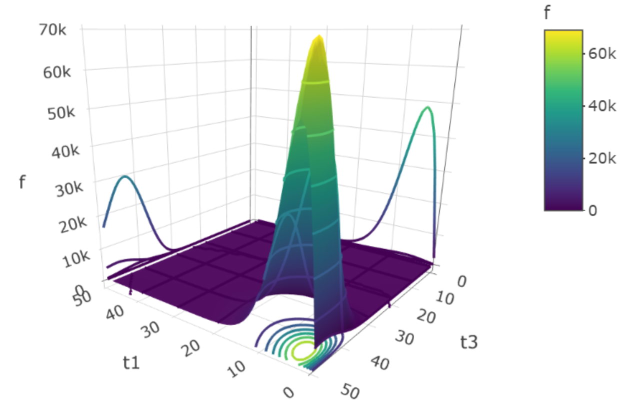


Figure 4.7: Density plot of posterior (pseudo-gamma prior) with hyper-parameter values are $\tau_1 = 2, \tau_2 = 4, \psi_1 = 2, \psi_2 = 1(small), \psi_3 = 3$ and parameter values $\theta_1 = 2, \theta_3 = 5$ with sample of size 30 of the sub-model-I.

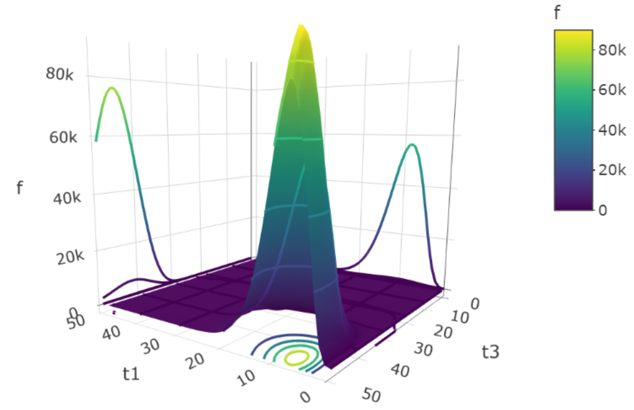


Figure 4.8: Density plot of posterior (pseudo-gamma prior) with hyper-parameter values are $\tau_1 = 2, \tau_2 = 4, \psi_1 = 2, \psi_2 = 0(simple), \psi_3 = 3$ and parameter values $\theta_1 = 2, \theta_3 = 5$ with sample of size 30 of the sub-model-I.

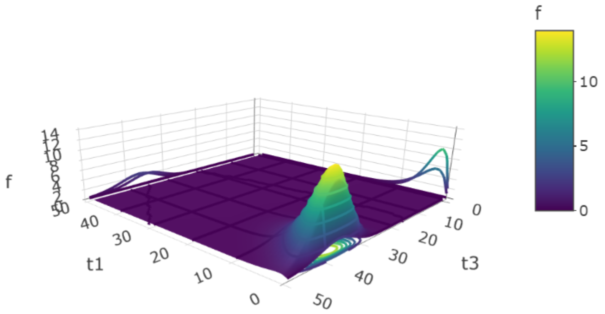


Figure 4.9: Density plot of posterior (pseudo-gamma prior) with hyper-parameter values are $\tau_1 = 2, \tau_2 = 4, \psi_1 = 2, \psi_2 = 7(large), \psi_3 = 3$ and parameter values $\theta_1 = 2, \theta_3 = 5$ with sample of size 30 of the sub-model-I.

4.4.2 Sub-Model-II ($\theta_2 = 0$):

Consider the sub-model obtained by setting $\theta_2 = 0$. The model is thus of the form

$$f_{X,Y}(x,y) = \theta_1 e^{-\theta_1 x}\theta_3 x e^{-\theta_3 xy} I(x > 0, y > 0).$$

The likelihood function for a sample of size n from this distribution is given by

$$L(\theta_1, \theta_3 | \underline{x}, \underline{y}) = \theta_1^n e^{-\theta_1 \sum_{i=1}^n x_i} \theta_3^n (\prod_{i=1}^n x_i) e^{-\theta_3 \sum_{i=1}^n x_i y_i}$$

the kernel of the likelihood function is given by

$$L(\theta_1, \theta_3 | \underline{x}, \underline{y}) \propto \theta_1^n e^{-\theta_1 \sum_{i=1}^n x_i} \theta_3^n e^{-\theta_3 \sum_{i=1}^n x_i y_i} \tag{4.4.7}$$

Consider the simpler joint pseudo-gamma prior in which we assume that $\psi_2 = 0$. This joint prior density will be

$$f_p(\theta_1, \theta_3) \propto \theta_1^{\tau_2 - 1} e^{-\theta_1 \theta_3 \psi_3} \theta_3^{\tau_1 + \tau_2 - 1} e^{-\psi_1 \theta_3} I(\theta_1 > 0, \theta_3 > 0). \tag{4.4.8}$$

The posterior density will be

$$
\begin{aligned}
f(\theta_1, \theta_3 | \underline{X} = \underline{x}, \underline{Y} = \underline{y}) &\propto (\theta_1^n e^{-\theta_1 \sum_{i=1}^n x_i})(\theta_3^n e^{-\theta_3 (\sum_{i=1}^n x_i y_i})\\
&* (\theta_1^{\tau_2 - 1} e^{-\theta_1 \theta_3 \psi_3} \theta_3^{\tau_1 + \tau_2 - 1} e^{-\psi_1 \theta_3})\\
&\propto \theta_1^{\tau_2 + n - 1} e^{-\theta_1 \sum_{i=1}^n x_i} \theta_3^{\tau_1 + \tau_2 - 1} e^{-\theta_3 (\psi_1 + \sum_{i=1}^n x_i y_i)}\\
&* e^{-\theta_1 \theta_3 \psi_3}\\
&\propto (\theta_3^{\tau_1 + n - 1} e^{-\theta_3 (\psi_1 + \sum_{i=1}^n x_i y_i)})\\
&* (\theta_1^{\tau_2 + n - 1} e^{-\theta_1 (\sum_{i=1}^n x_i + \psi_3 \theta_3)})
\end{aligned}
\tag{4.4.9}
$$

The marginal posterior distributions of θ_1 and θ_3 are

$$
f_{\theta_1}^p(\theta_1) \propto [\theta_1^{\tau_2 + n - 1} e^{-\theta_1 \sum_{i=1}^n x_i}] . \frac{\Gamma(\tau_1 + n)}{(\psi_1 + \sum_{i=1}^n x_i y_i + \theta_1 \psi_3)^{(\tau_1 + n)}}, \tag{4.4.10}
$$

and

$$
f_{\theta_1}^p(\theta_1) \propto [\theta_3^{\tau_1 + n - 1} e^{-\theta_3 (\psi_1 + \sum_{i=1}^n x_i y_i)}] . \frac{\Gamma(\tau_2 + n)}{(\sum_{i=1}^n x_i + \theta_3 \psi_3)^{(\tau_2 + n)}}. \tag{4.4.11}
$$

Note that the mean and variance of the marginals need to be dealt with numerically.

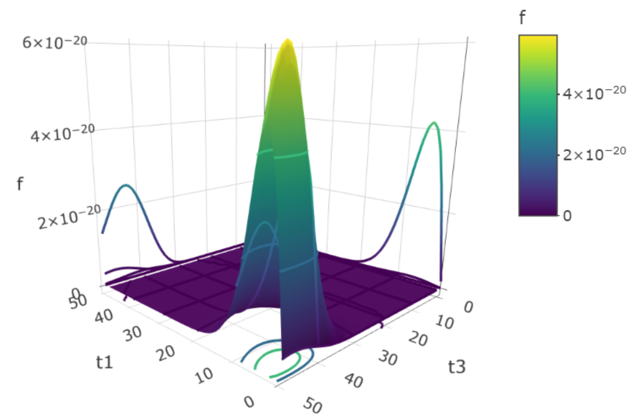


Figure 4.10: Density plot of posterior (pseudo-gamma prior) with hyper-parameter values are $\tau_1 = 2, \tau_2 = 4, \psi_1 = 2, \psi_2 = 1(small), \psi_3 = 3$ and parameter values $\theta_1 = 2, \theta_3 = 5$ with sample of size 30 of the sub-model-II.

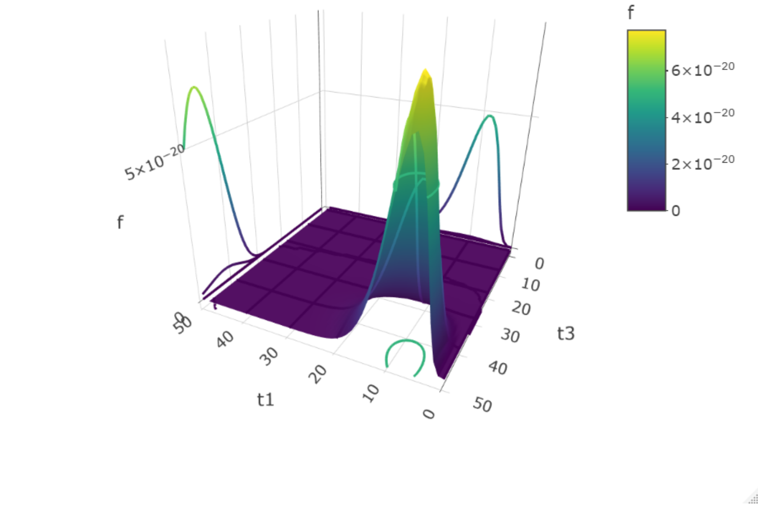


Figure 4.11: Density plot of posterior (pseudo-gamma prior) with hyper-parameter values are $\tau_1 = 2, \tau_2 = 4, \psi_1 = 2, \psi_2 = 0(simple), \psi_3 = 3$ and parameter values $\theta_1 = 2, \theta_3 = 5$ with sample of size 30 of the sub-model-II.

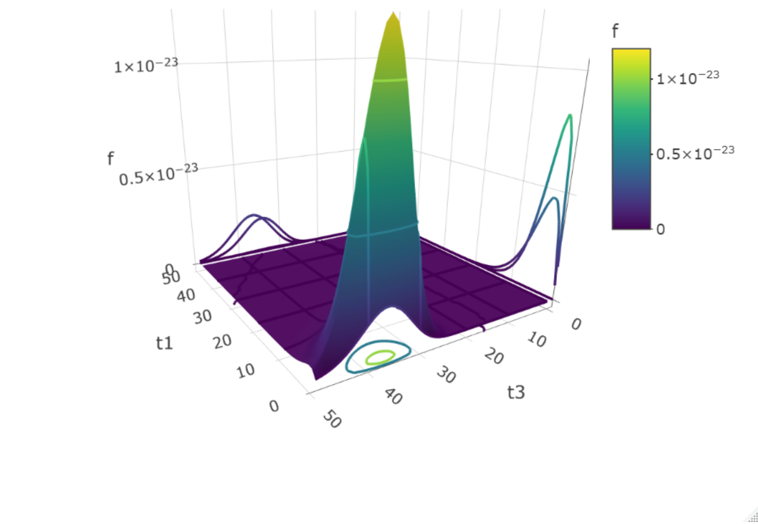


Figure 4.12: Density plot of posterior (pseudo-gamma prior) with hyper-parameter values are $\tau_1 = 2, \tau_2 = 4, \psi_1 = 2, \psi_2 = 7(large), \psi_3 = 3$ and parameter values $\theta_1 = 2, \theta_3 = 5$ with sample of size 30 of the sub-model-II.

4.5 Simulation study:

Due to the general marginal and conditional composition of exponential distributions, simulation can be performed in two steps: first, simulate x from exponential(θ_1) and then y from exponential($\theta_2 + \theta_3 x$). However, even with independent gamma priors, inference for the posterior distribution is difficult for the full model. We must rely on a numerical algorithm to compute marginal distributions and their moments. In this paper, we will use the Hit-And-Run Metropolis (HARM) algorithm to simulate posterior distributions from observations for all priors (improper, independent, and pseudo) and for full and sub-models. For the (HARM) algorithm implementation and comparison with Gibbs and Metropolis sampling, see Chen[9]. We refer to Hall[18] for current Bayesian simulation algorithms using R software.

We have simulated 10,000 data sets using the HARM algorithm with thinning at 10th sample from the posterior distributions of θ's with varying sample sizes of n = 20, 30, 50, 100, 200, 500: Mean and posterior 95% confidence intervals are mentioned in Table 4.1 and 4.2 are computed from 10000 iterations of the preceding procedure.

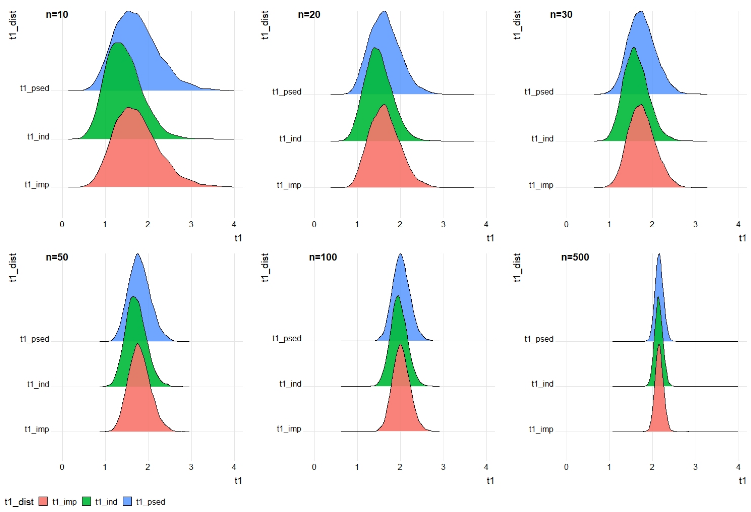


Figure 4.13: Posterior density plot of θ_1 (independent gamma, improper and pseudo-gamma priors) with hyper-parameters $\alpha_1 = 2, \alpha_2 = 3, \alpha_3 = 4, \beta_1 = 2, \beta_2 = 1, \beta_3 = 5$ and parameter values $\theta_1 = 2, \theta_2 = 1, \theta_3 = 3$ with sample of size(n=10, 20, 30, 50, 100, 500).

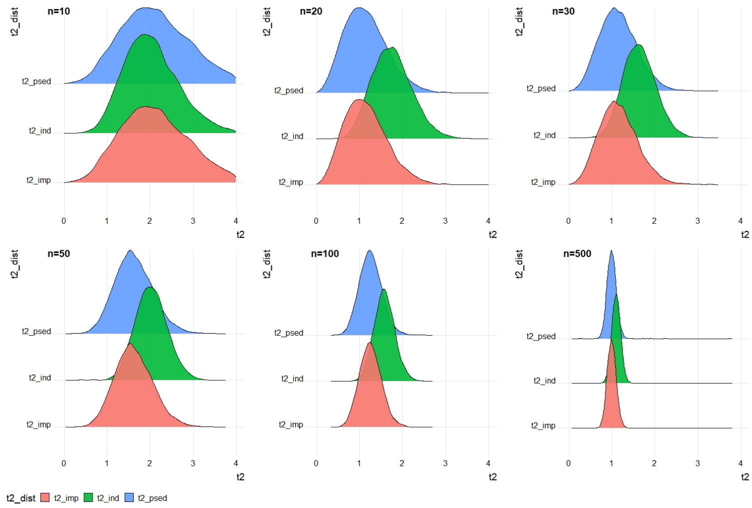


Figure 4.14: Posterior density plot of θ_2 (independent gamma, improper and pseudo-gamma priors) with hyper-parameters $\alpha_1 = 2, \alpha_2 = 3, \alpha_3 = 4, \beta_1 = 2, \beta_2 = 1, \beta_3 = 5$ and parameter values $\theta_1 = 2, \theta_2 = 1, \theta_3 = 3$ with sample of size(n=10, 20, 30, 50, 100, 500).

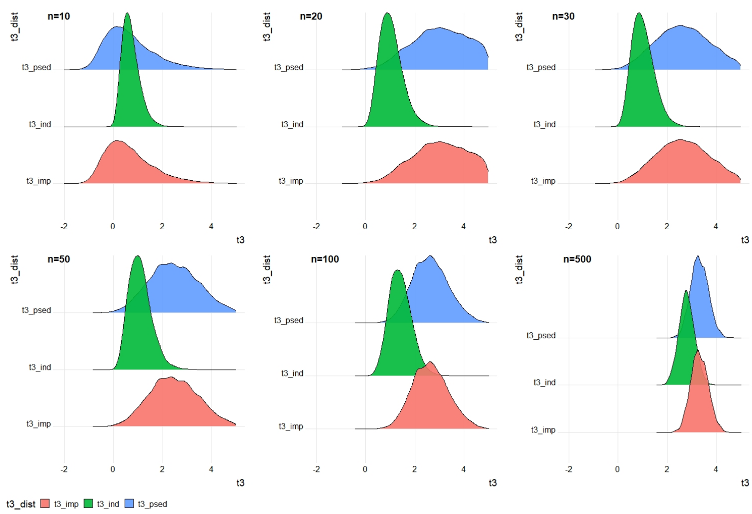


Figure 4.15: Posterior density plot of θ_3 (independent gamma, improper and pseudo-gamma priors) with hyper-parameters $\alpha_1 = 2, \alpha_2 = 3, \alpha_3 = 4, \beta_1 = 2, \beta_2 = 1, \beta_3 = 5$ and parameter values $\theta_1 = 2, \theta_2 = 1, \theta_3 = 3$ with sample of size(n=10, 20, 30, 50, 100, 500).

Table 4.1: Simulation (Sub-model I)

Sample size (n), parameters (P), posterior mean by independent gamma prior (IGP(1)), posterior mean by improper prior (ImP(2)), posterior mean by pseudo-gamma prior for $\psi_2 = 1$ (PGP1(3)), posterior mean by pseudo-gamma prior for $\psi_2 = 0$ (PGP2(4)), posterior mean by pseudo-gamma prior for $\psi_2 = 7$ (PGP3(5)), 95% confidence interval of IGP(1) (CI(1)), 95% confidence interval of ImP(2) (CI(2)), 95% confidence interval of PGP1(3) (CI(3)), 95% confidence interval of PGP2(4) (CI(4)), 95% confidence interval of PGP3(5) (CI(5))

n	P	SIP$_1$	SIP$_2$	SIP$_3$	SIP$_4$	SIP$_5$	CISIP$_1$	CISIP$_2$	CISIP$_3$	CISIP$_4$	CISIP$_5$
20	θ_1	1.483	1.624	0.998	1.071	0.815	(0.924 2.141)	(1.015 2.371)	(0.504 1.550)	(0.608 1.608)	(0.516 1.172)
	θ_3	2.304	3.306	2.239	2.075	1.984	(1.350 3.291)	(2.028 4.965)	(1.257 9.249)	(1.258 2.935)	(1.259 2.936)
30	θ_1	1.729	1.953	1.242	1.287	1.016	(1.152 2.365)	(1.257 2.648)	(0.836 1.707)	(0.853 1.769)	(0.702 1.401)
	θ_3	2.184	3.159	1.895	1.884	1.929	(1.486 3.041)	(2.011 4.498)	(1.282 2.643)	(1.283 2.616)	(1.319 2.625)
50	θ_1	2.294	2.450	1.677	1.746	1.404	(1.723 2.926)	(1.813 3.127)	(1.240 2.179)	(1.273 2.261)	(0.984 1.808)
	θ_3	2.488	3.091	2.121	2.030	2.309	(1.845 3.193)	(2.248 4.013)	(1.517 2.765)	(1.482 2.621)	(1.547 2.868)
100	θ_1	2.465	2.918	1.986	2.032	1.782	(2.033 2.981)	(2.129 3.265)	(1.625 2.399)	(1.648 2.479)	(1.449 2.128)
	θ_3	2.830	3.175	2.422	2.417	2.450	(2.278 3.435)	(2.565 3.811)	(1.970 2.892)	(1.971 2.942)	(1.961 2.973)
200	θ_1	2.039	2.066	1.823	1.845	1.801	(1.765 2.324)	(1.776 2.362)	(1.582 2.088)	(1.588 2.132)	(1.503 2.020)
	θ_3	2.850	3.028	2.650	2.813	2.663	(2.474 3.258)	(2.625 3.425)	(2.281 3.010)	(2.303 3.180)	(2.279 3.051)
500	θ_1	1.912	1.944	1.828	2.245	1.781	(1.747 2.096)	(1.761 2.085)	(1.673 1.996)	(1.669 2.042)	(1.639 1.934)
	θ_3	3.026	3.104	2.944	2.906	2.934	(2.776 3.262)	(2.829 3.384)	(2.676 3.206)	(2.614 3.190)	(2.669 3.204)

Table 4.2: Simulation (Sub-model II)

Sample size (n), parameters (P), posterior mean by independent gamma prior (IGP(1)), posterior mean by improper prior (ImP(2)), posterior mean by pseudo-gamma prior for $\psi_2 = 1$ (PGP1(3)), posterior mean by pseudo-gamma prior for $\psi_2 = 0$ (PGP2(4)), posterior mean by pseudo-gamma prior for $\psi_2 = 7$ (PGP3(5)), 95% confidence interval of IGP(1) (CI(1)), 95% confidence interval of ImP(2) (CI(2)), 95% confidence interval of PGP1(3) (CI(3)), 95% confidence interval of PGP2(4) (CI(4)), 95% confidence interval of PGP3(5) (CI(5))

n	P	SIP$_1$	SIP$_2$	SIP$_3$	SIP$_4$	SIP$_5$	CISIP$_1$	CISIP$_2$	CISIP$_3$	CISIP$_4$	CISIP$_5$
20	θ_1	1.473	1.643	0.792	0.822	0.629	(0.900 2.177)	(1.040 2.394)	(0.481 1.192)	(0.469 1.321)	(0.383 0.913)
	θ_3	3.744	11.814	3.344	3.417	3.739	(2.297 5.415)	(7.516 16.266)	(2.065 4.996)	(2.045 5.184)	(2.227 5.714)
30	θ_1	1.757	1.909	0.969	1.001	1.103	(1.222 2.399)	(1.268 2.617)	(0.633 1.399)	(0.667 1.444)	(0.540 1.400)
	θ_3	4.207	9.799	3.679	3.565	4.066	(2.954 5.829)	(6.403 14.006)	(2.452 5.282)	(2.430 4.943)	(2.538 6.711)
50	θ_1	1.824	1.821	0.988	1.011	0.860	(1.295 2.251)	(1.337 2.247)	(0.718 1.295)	(0.738 1.329)	(0.644 1.070)
	θ_3	6.041	12.405	5.126	5.070	5.388	(4.452 7.781)	(9.154 14.984)	(3.653 6.581)	(3.689 6.629)	(4.063 7.842)
100	θ_1	2.057	2.285	1.316	1.351	1.210	(1.680 2.481)	(1.711 2.646)	(1.052 1.604)	(1.087 1.649)	(0.982 1.478)
	θ_3	7.628	11.732	6.154	6.064	6.306	(6.319 9.046)	(9.091 14.179)	(4.912 7.569)	(4.872 7.438)	(4.984 7.669)
200	θ_1	1.983	2.005	1.466	1.482	1.398	(1.709 2.265)	(1.730 2.290)	(1.271 1.694)	(1.276 1.690)	(1.208 1.592)
	θ_3	8.995	11.257	7.583	7.565	7.661	(7.880 10.261)	(9.636 12.890)	(6.516 8.689)	(6.486 8.579)	(6.631 8.792)
500	θ_1	2.212	2.048	1.731	1.732	1.694	(1.866 2.270)	(1.876 2.211)	(1.589 1.880)	(1.558 1.894)	(1.543 1.848)
	θ_3	10.185	11.098	9.047	9.253	9.083	(9.129 11.412)	(10.163 12.109)	(8.268 9.915)	(8.177 10.141)	(8.301 10.008)

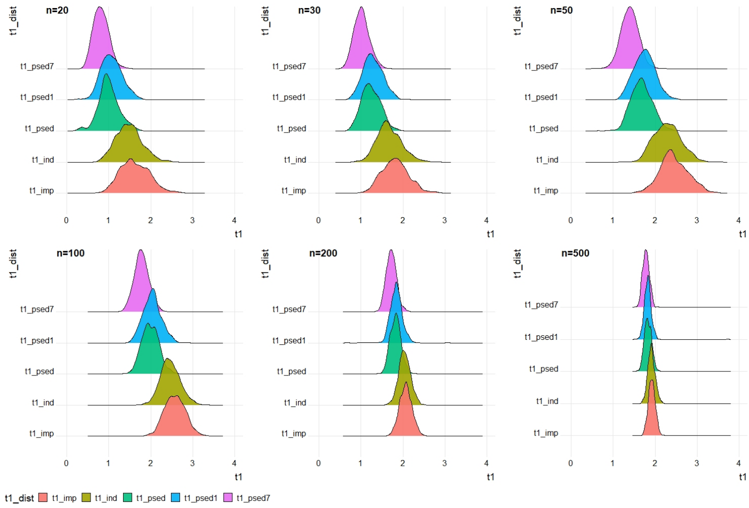


Figure 4.16: Posterior density plot of θ_1 with hyper-parameters $\alpha_1 = 2, \alpha_3 = 4, \beta_1 = 2, \beta_3 = 5$ and parameter values $\theta_1 = 2, \theta_3 = 3$ of sub-model-I.

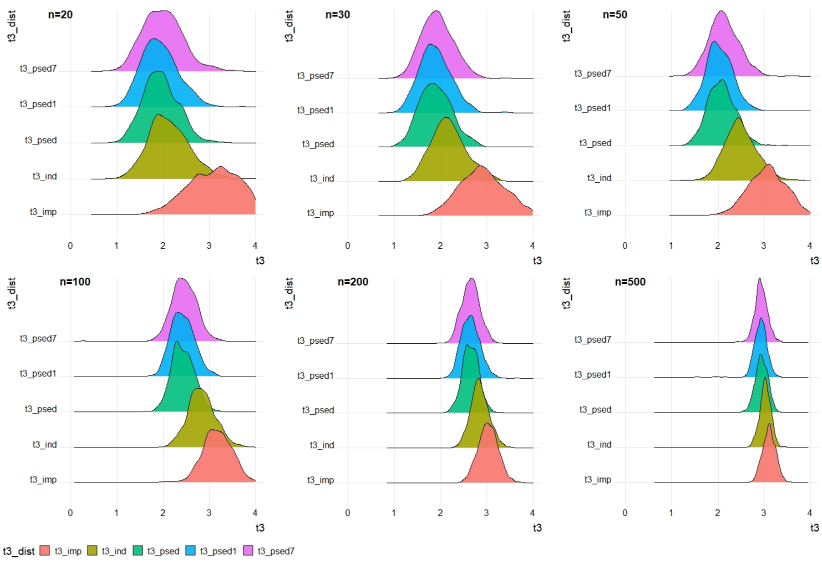


Figure 4.17: Posterior density plot of θ_3 with hyper-parameters $\alpha_1 = 2, \alpha_3 = 4, \beta_1 = 2, \beta_3 = 5$ and parameter values $\theta_1 = 2, \theta_3 = 3$ of sub-model-I.

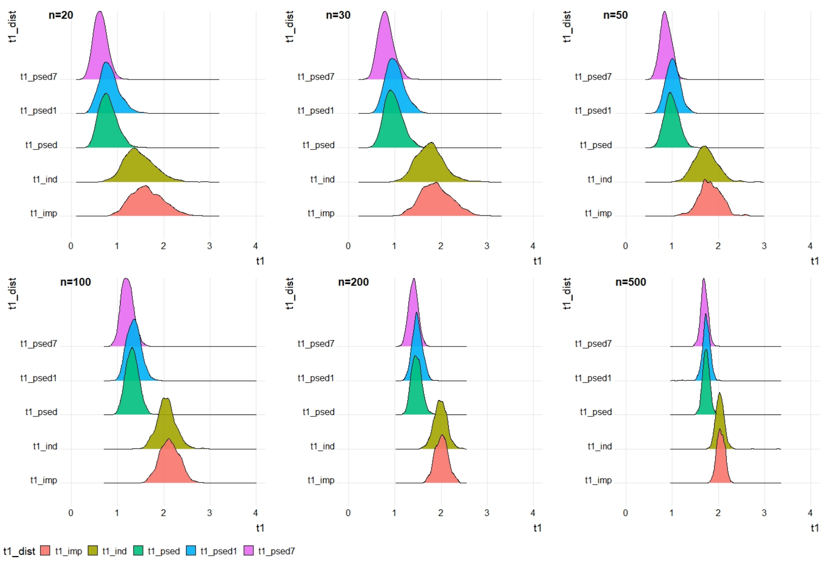


Figure 4.18: Posterior density plot of θ_1 with hyper-parameters $\alpha_1 = 2, \alpha_3 = 4, \beta_1 = 2, \beta_3 = 5$ and parameter values $\theta_1 = 2, \theta_3 = 3$ of sub-model-II.

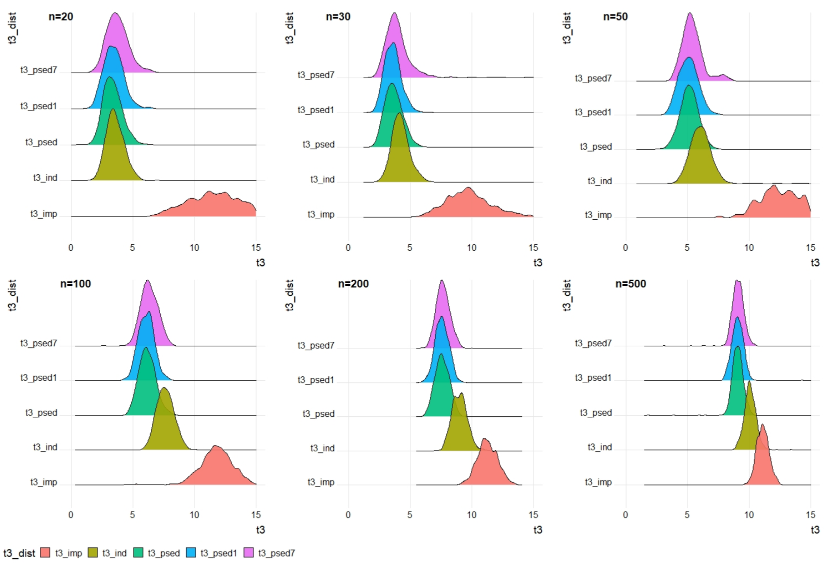


Figure 4.19: Posterior density plot of θ_1 with hyper-parameters $\alpha_1 = 2, \alpha_3 = 4, \beta_1 = 2, \beta_3 = 5$ and parameter values $\theta_1 = 2, \theta_3 = 3$ of sub-model-II.

Table 4.3: per Capita GDP and Infant Mortality rate dataset (full-model)

n	P	IGP(1)	ImP(2)	PGP1(3)	CI(1)	CI(2)	CI(3)
	θ_1	0.209	0.113	0.112	(0.041 0.056)	(0.031 0.170)	(0.040 0.321)
225	θ_2	0.026	0.032	2.234	(0.011 0.056)	(0.007 0.033)	(0.009 35.763)
	θ_3	0.004	0.098	0.022	(0.002 0.005)	(0.003 0.005)	(0.001 0.016)

Table 4.4: per Capita GDP and Infant Mortality rate dataset (sub-model-I)

n	P	IGP(1)	ImP(2)	PGP1(3)	CI(1)	CI(2)	CI(3)
225	θ_1	0.04103	0.04739	0.05478	(0.04906, 0.04113)	(0.04729, 0.04773)	(0.05410, 0.05768)
	θ_3	0.00468	0.00451	0.00420	(0.00447, 0.00498)	(0.00441, 0.00490)	(0.00400, 0.00464)

4.6 Applications:

One example from the social sciences where an inverse dependence relationship is expected is infant mortality and GDP. Both of these indicators are roughly exponentially distributed for countries and regions all over the world. The C.I.A. generates estimates of values based on massive databases of global affairs information. In 2022, 225 countries and regions will provide data on infant mortality as deaths per 1000 live births (Y) and per capita GDP in thousands of dollars (X).

Table 4.5: per Capita GDP and Infant Mortality rate dataset (sub-model-II)

n	P	IGP(1)	ImP(2)	PGP1(3)	CI(1)	CI(2)	CI(3)
225	θ_1	0.04131	0.05280	0.04991	(0.04125, 0.04142)	(0.04013, 0.06238)	(0.04768, 0.05091)
	θ_3	0.00506	0.00484	0.00508	(0.00484, 0.00536)	(0.00395, 0.00588)	(0.00481, 0.00550)

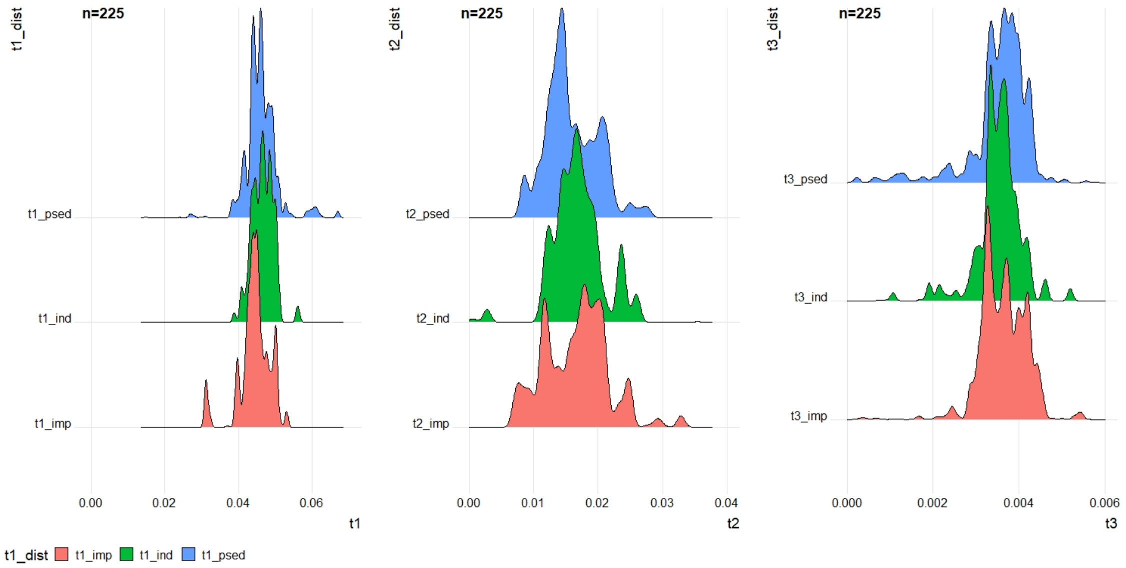


Figure 4.20: Posterior density plots for full-model

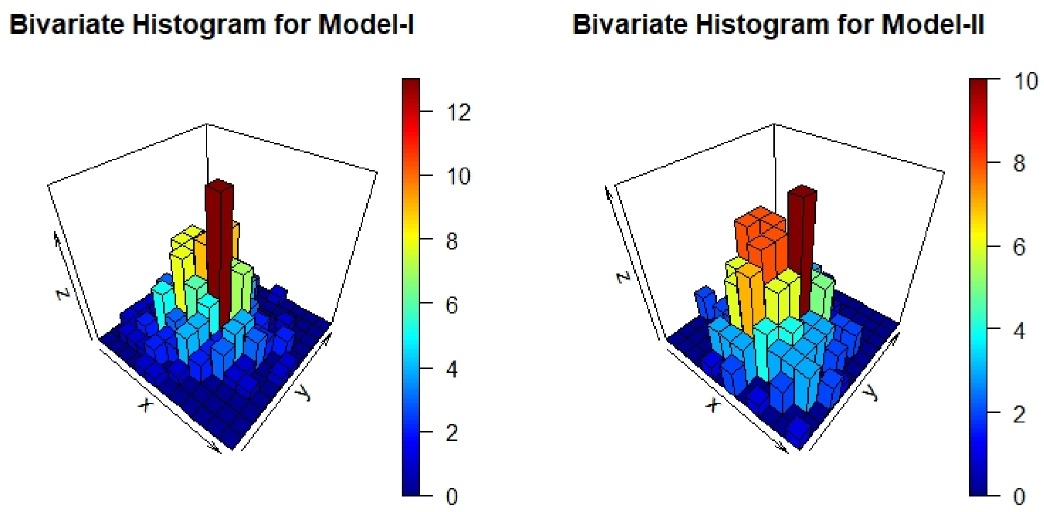


Figure 4.21: Histograms for the bivariate data of the per Capita GDP and Infant Mortality rate dataset for sub-model-I & II

CHAPTER 5

Conclusion

It has been advocated strongly by Arnold and Manjunath [4] that the pseudo-Poisson distribution might be considered the first choice in modeling bivariate count data whenever one marginal is equi-dispersed and another marginal is over-dispersed. Also, due to the simplicity of its structure, the pseudo-Poisson allows simple simulation and straightforward computation and Bayesian inference. Nevertheless, when external information indicates dependence between the parameters, there are very few bivariate priors in the literature that can accommodate such dependence. The pseudo-gamma prior, which was introduced in Chapter 2, can be considered a possible choice for Bayesian analysis with dependence on the prior. Through a simulation study and a real-life data analysis, it has been highlighted that the pseudo-Poisson with a pseudo-gamma prior performs better than an improper or an independent prior when the sample size is moderately large. The pseudo-Poisson with pseudo-gamma prior can be easily extended to any higher dimension, and one extension has been discussed in Section 6 of Chapter 2. Finally, we recommend pseudo-Poisson with pseudo-gamma prior for Bayesian analysis of count data whenever there is additional information about the dependence in the prior.

The GoF tests for the bivariate pseudo-Poisson and its sub-models were the main emphasis of the current note. We proposed a few GoF tests based on p.g.f., moments, and Chi-square tests. The supremum of the absolute difference between the calculated p.g.f. and its empirical equivalent, the new GoF test we proposed and analyzed. The bivariate Fisher index of the dispersion-based GoF test has also been added. Additionally, we took into account a few existing tests that depend on the estimated p.g.f and its empirical results, such as K&K, Munoz, and Gamero approaches. Finally, the Chi-square GoF test results for the pseudo-Poisson data were also examined. A finite sample, a fairly large sample, and asymptotic distributions of test statistics are examined for each of the tests discussed. In addition, we looked at the power and efficacy of each statistical test using the bivariate Com-Max-Poisson and the bivariate Classical Bivariate (BCP) as alternative distributions. It has been demonstrated that a test based on the supremum and index of dispersion is reliable, consistent, and satisfying. Particularly, the supremum-based test proved to be more robust to the choice of alternative distribution. Additionally, we suggest utilizing the Muñoz and Gamero (M&G) test for moderately small samples and the supremum (robust) and dispersion tests for moderately large samples. Due to the asymptotic distribution of the test statistic, we also recommend K&K and dispersion tests for sufficiently large data sets. Also, due to its robust property, we suggest considering the supremum and Chi-square GoF tests if there are no reasonable alternatives to the hypothesis. The suggested procedures and algorithms will add another tool for analyzing count data. The developed procedures will merit a spot in the toolkit of contemporary modelers because of their simple structure and fast computation.

Appendices

A.1 Examples

Consider the following examples:

Example 2. *define* $w_1(t_1,t_2) = c_1 + c_2 t_1 t_2 + c_3 t_1^2 t_2^2$, $(t_1,t_2)^T \in [0,1]^2$, $c_1, c_2, c_3 \in \mathbb{R}$ *and* $T_{n,w_1}(\hat\lambda_1, \hat\lambda_2, \hat\lambda_3)$ *is*

$$
\begin{aligned}
& c_1 \Bigg\{ \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \left(\frac{1}{(X_i + X_j + 1)(Y_i + Y_j + 1)} \right) + \\
+ \quad & \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \Bigg(\mathscr{P}(k; \hat\lambda_1) \mathscr{P}(l; \hat\lambda_2 + k\hat\lambda_3) \mathscr{P}(m; \hat\lambda_1) \mathscr{P}(n; \hat\lambda_2 + m\hat\lambda_3) \\
& \int_0^1 \int_0^1 t_1^{k+m} t_2^{l+n} dt_1 dt_2 \Bigg) - \\
& -2 \sum_{i=1}^{n} \sum_{x=0}^{\infty} \sum_{y=0}^{\infty} \left(\mathscr{P}(x; \hat\lambda_1) \mathscr{P}(y; \hat\lambda_2 + x\hat\lambda_3) \int_0^1 \int_0^1 t_1^{x+X_i} t_2^{y+Y_i} dt_1 dt_2 \right) \Bigg\} + \\
& c_2 \Bigg\{ \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \left(\frac{1}{(X_i + X_j + 2)(Y_i + Y_j + 2)} \right) + \\
+ \quad & \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \Bigg(\mathscr{P}(k; \hat\lambda_1) \mathscr{P}(l; \hat\lambda_2 + k\hat\lambda_3) \mathscr{P}(m; \hat\lambda_1) \mathscr{P}(n; \hat\lambda_2 + m\hat\lambda_3) \\
& \int_0^1 \int_0^1 t_1^{k+m+1} t_2^{l+n+1} dt_1 dt_2 \Bigg) - \\
& -2 \sum_{i=1}^{n} \sum_{x=0}^{\infty} \sum_{y=0}^{\infty} \left(\mathscr{P}(x; \hat\lambda_1) \mathscr{P}(y; \hat\lambda_2 + x\hat\lambda_3) \int_0^1 \int_0^1 t_1^{x+X_i+1} t_2^{y+Y_i+1} dt_1 dt_2 \right) \Bigg\} + \\
& c_3 \Bigg\{ \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \left(\frac{1}{(X_i + X_j + 3)(Y_i + Y_j + 3)} \right) + \\
+ \quad & \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \Bigg(\mathscr{P}(k; \hat\lambda_1) \mathscr{P}(l; \hat\lambda_2 + k\hat\lambda_3) \mathscr{P}(m; \hat\lambda_1) \mathscr{P}(n; \hat\lambda_2 + m\hat\lambda_3)
\end{aligned}
$$

ne more

$$
\begin{aligned}
& \int_0^1 \int_0^1 t_1^{k+m+2} t_2^{l+n+2} dt_1 dt_2 \Bigg) - \\
& -2 \sum_{i=1}^{n} \sum_{x=0}^{\infty} \sum_{y=0}^{\infty} \left(\mathscr{P}(x; \hat\lambda_1) \mathscr{P}(y; \hat\lambda_2 + x\hat\lambda_3) \int_0^1 \int_0^1 t_1^{x+X_i+2} t_2^{y+Y_i+2} dt_1 dt_2 \right) \Bigg\}
\end{aligned}
\tag{A.1.1}
$$

where $\mathscr{P}(i;.\hat{\lambda})$ is a Poisson probability at i for the estimated parameter $\hat{\lambda}$. Further simplication gives us

$$
\begin{aligned}
T_{n,w_1}(\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3) \quad = \quad & \frac{c_1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\frac{1}{(X_i+X_j+1)(Y_i+Y_j+1)}\right)+ \\
& \frac{c_2}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\frac{1}{(X_i+X_j+2)(Y_i+Y_j+2)}\right)+ \\
& \frac{c_3}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\frac{1}{(X_i+X_j+3)(Y_i+Y_j+3)}\right)+ \\
& c_1\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\left(\frac{\mathscr{P}(k;\hat{\lambda}_1)\mathscr{P}(l;\hat{\lambda}_2+k\hat{\lambda}_3)\mathscr{P}(m;\hat{\lambda}_1)\mathscr{P}(n;\hat{\lambda}_2+m\hat{\lambda}_3)}{(k+m+1)(l+n+1)}\right)+ \\
& c_2\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\left(\frac{\mathscr{P}(k;\hat{\lambda}_1)\mathscr{P}(l;\hat{\lambda}_2+k\hat{\lambda}_3)\mathscr{P}(m;\hat{\lambda}_1)\mathscr{P}(n;\hat{\lambda}_2+m\hat{\lambda}_3)}{(k+m+2)(l+n+2)}\right)+ \\
& c_3\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\left(\frac{\mathscr{P}(k;\hat{\lambda}_1)\mathscr{P}(l;\hat{\lambda}_2+k\hat{\lambda}_3)\mathscr{P}(m;\hat{\lambda}_1)\mathscr{P}(n;\hat{\lambda}_2+m\hat{\lambda}_3)}{(k+m+3)(l+n+3)}\right)+ \\
& -2c_1\sum_{i=1}^{n}\sum_{x=0}^{\infty}\sum_{y=0}^{\infty}\left(\frac{\mathscr{P}(x;\hat{\lambda}_1)\mathscr{P}(y;\hat{\lambda}_2+x\hat{\lambda}_3)}{(x+X_i+1)(y+Y_i+1)}\right) \\
& -2c_2\sum_{i=1}^{n}\sum_{x=0}^{\infty}\sum_{y=0}^{\infty}\left(\frac{\mathscr{P}(x;\hat{\lambda}_1)\mathscr{P}(y;\hat{\lambda}_2+x\hat{\lambda}_3)}{(x+X_i+2)(y+Y_i+2)}\right) \\
& -2c_3\sum_{i=1}^{n}\sum_{x=0}^{\infty}\sum_{y=0}^{\infty}\left(\frac{\mathscr{P}(x;\hat{\lambda}_1)\mathscr{P}(y;\hat{\lambda}_2+x\hat{\lambda}_3)}{(x+X_i+3)(y+Y_i+3)}\right). \qquad \text{(A.1.2)}
\end{aligned}
$$

We refer to Table 3.5 and Figuare 3.4 for the quantile values and frequency distribution for $a_1 = 1$ and $a_2 = 1$, respectively.

Example 3. *For a general form of $w(.,.)$, consider $w_2(t_1, t_2) = t_1^{a_1} t_2^{a_2}$, $(t_1, t_2)^T \in [0, 1]^2$, $a_1, a_2 \in (-1, \infty)$, which allows us to include a negative powers as well, then the T_{n,w_2} is*

$$
\begin{aligned}
& \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\frac{1}{(X_i+X_j+a_1+1)(Y_i+Y_j+a_2+1)}\right)+ \\
+ \quad & \sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\Bigg(\mathscr{P}(k;\hat{\lambda}_1)\mathscr{P}(l;\hat{\lambda}_2+k\hat{\lambda}_3)\mathscr{P}(m;\hat{\lambda}_1)\mathscr{P}(n;\hat{\lambda}_2+m\hat{\lambda}_3) \\
& \int_0^1\int_0^1 t_1^{k+m+a_1} t_2^{l+n+a_2} dt_1 dt_2\Bigg) -
\end{aligned}
$$

$$-2\sum_{i=1}^{n}\sum_{x=0}^{\infty}\sum_{y=0}^{\infty}\left(\mathscr{P}(x;\hat{\lambda}_1)\mathscr{P}(y;\hat{\lambda}_2+x\hat{\lambda}_3)\int_0^1\int_0^1 t_1^{x+X_i+a_1}t_2^{y+Y_i+a_2}dt_1dt_2\right).$$

Now, further simplification will give us closed form expression for the statistic

$$\begin{aligned} T_{n,w_2}(\hat{\lambda}_1,\hat{\lambda}_2\hat{\lambda}_3) &= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\frac{1}{(X_i+X_j+a_1+1)(Y_i+Y_j+a_2+1)}\right)+ \\ &+ \sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\left(\frac{\mathscr{P}(k;\hat{\lambda}_1)\mathscr{P}(l;\hat{\lambda}_2+k\hat{\lambda}_3)\mathscr{P}(m;\hat{\lambda}_1)\mathscr{P}(n\mathscr{P}(x;\hat{\lambda}_1)}{(k+m+a_1+1)(l+n+a_2+1)}\right)- \\ &\quad -2\sum_{i=1}^{n}\sum_{x=0}^{\infty}\sum_{y=0}^{\infty}\left(\frac{\mathscr{P}(x;\hat{\lambda}_1)\mathscr{P}(y;\hat{\lambda}_2+x\hat{\lambda}_3)}{(x+X_i+a_1+1)(y+Y_i+a_2+1)}\right). \end{aligned} \tag{A.1.3}$$

We refer to Table 3.6 and Figuare 3.5 for the quantile values and frequency distribution for $a_1 = 1$ and $a_2 = 1$, respectively. Also, with varying a_1 and a_2 values finite sample distribution of the statistics, see Table 3.7 and Figuares 3.6, 3.7 and 3.8.

ne more

Bibliography

[1] Arnold, B.C., Manjunath, B.G., Shobha, B., Veeranna, B.: *On Bayesian inference for Pseudo-Poisson data*. Journal of computational statistics and simulation (2022) 50, 75

[2] Arnold, B.C., Arvanitis, M.A.: *On bivariate pseudo-exponential distributions*. Journal of Applied Statistics (2019) 75

[3] Arnold, B.C., Castillo, E., Sarabia, J.M.: *Conditional Specification of Statistical Models*. Springer, New York (1999) 2, 11

[4] Arnold, B.C., Manjunath, B.G.: *Statistical inference for distributions with one Poisson conditional*. Journal of Applied Statistics **48**, 2306–2325 (2021) 3, 4, 5, 9, 13, 40, 41, 42, 43, 50, 52, 60, 70, 72, 74, 99

[5] Arnold, B.C., Manjunath, B.G.: *Pseudo-Poisson Distributions with Concomitant Variables*. Mathematical and Computational Applications **1**, 1–14 (2023) 10

[6] Balakrishnan, N., Lai, C.D.: *Continuous bivariate distributions*. Springer Science & Business Media (2009) 22

[7] Cacoullos, T.: *Characterizing priors by posterior expectations in multi-parameter exponential families*. Ann. Inst. Statist. Math. **39(A)**, 399–405 (1987) 75

[8] Chen, Feiyan: *The Goodness-of-fit tests for geometric models*. Dissertations. 350 58, 60

[9] Chen, M.H., Schmeiser, B.: *Performance of the Gibbs, Hit-and-Run, and Metropolis Samplers*. J. Comput. And Graph. Statist. **2(3)**, 251–272 (1993) 30, 91

[10] Chour, J.P.: *Characterizations of conjugate priors for discrete exponential families*. Stat. Sinica **11**, 409–418 (2001) 75

[11] Dubey, S.: *Compound gamma, beta and F distributions*. Metrika: International Journal for Theoretical and Applied Statistics **16(1)**, 27–31 (1970) 22

[12] Filus, J.K., Filus, L.Z.: *On some new classes of multivariate probability distributions*. Pakistan J. Statist. **22**, 21–42 (2006) 75, 76

[13] Filus, J.K., Filus, L.Z.: *On new multivariate probability distributions and stochastic processes with system reliability and maintenance applications*. Methodol. Comput. Appl. Probab. **9**, 425–446 (2007) 76

[14] Filus, J.K., Filus, L.Z.: *A method for multivariate probability distributions construction via parameter dependence*. comm. Statist. Theory Methods **42**, 716–721 (2013) 75

[15] Filus, J.K., Filus, L.Z., Arnold, B.C.: *Families of multivariate distributions involving "Triangular" transformations*. Comm. in Statistics-Theory and Methods **39**, 107–116 (2009) 13

[16] Ghosh, I., Marques, F., Chakraborty, S.: *A new bivariate Poisson distribution via conditional specification: properties and applications*. Journal of Applied Statistics **48(16)** (2021) 2, 13

[17] Gleeson, A.C., Douglas, J.B.: *Quadrate sampling and the estimation of Neyman Type A and Thomas distributional parameters*. Austral. J. Statist. **17(2)**, 103–113 (1975) 5, 6

[18] Hall, B.: *LaplacesDemon: An R Package for Bayesian Inference, https://cran.r-project.org/packages/LaplacesDemon/index.html.* . (2021) 30, 91

[19] Islam, M.A., Chowdhury, R.I.: *Analysis of Repeated Measures Data*. Springer Nature, Singapore (2017) 40, 70

[20] Johnson, N.L., Kemp, A.W., Kotz, S.: *Univariate Discrete Distributions*. John Wiley & Sons, New Jersey (2005) 5, 6

[21] Karlis, D., Tsiamyrtzis, P.: *Exact Bayesian modeling for bivariate Poisson data and extensions*. Statistics and Computing **18(1)**, 27–40 (2008) 13, 22, 40, 44, 70

[22] Kocherlakota, S., Kocherlakota, K.: *Univariate discrete distributions*. Marcel Dekker Inc., New York (2005) 50, 56

[23] Kokonendji, C.C., Puig, P.: *Fisher dispersion index for multivariate count distributions: A review and a new proposal*. J. Multivar. Anal. **165**, 180–193 (2018) 52, 53, 54

[24] Leiter, R.E., Hamdan, M.A.: *Some bivariate probability models applicable to traffic accidents and fatalities*. International Statistical Review **41**, 81–100 (1973) 4, 40, 42, 43, 70, 72

[25] Mardia, K.V.: *Multivariate Pareto distributions*. Ann. Math. Statist. **33:3**, 1008–1015 (1962) 1

[26] Meintanis, S.G.: *A review of testing procedures based on the empirical characteristic function*. The South African Stat. Journal **50**, 1–14 (2016) 49

[27] Mohsin, M., Kazianka, H., Pilz, J., Gebhardt, A.: *A new bivariate exponential distribution for modeling moderately negative dependence*. Stat. Methods. Appl. **23**, 123–148 (2014) 22

[28] Munoz, F.N.: *Goodness-of-fit tests for the bivariate Poisson distribution*. Communications in Statistics–Simulation and Computation **50:7**, 1998–2014 (2021) 50, 56

[29] Munoz, F.N., Jiménez-Gamero, M.D.: *Testing for the bivariate Poisson distribution*. Metrika **77**, 771–793 (2014) 50, 55, 56, 62

[30] Munoz, F.N., Jiménez-Gamero, M.D.: *goodness-of-fit test for the multivariate Poisson distribution*. SORT **40(1)**, 113–138 (2016) 50

[31] Nikitin, Y.Y.: *Tests based on characterizations and their efficiencies: A survey*. Acta et Commentationes Universitatis Tartuensis de Mathematica **21(1)**, 3–24 (2019) 49

[32] Obrechkoff, N.: *Theory of Probability*. Nauka i Izkustvo, Sofia (1963) 13

[33] Sellers, K.F., Morris, S.D., Balakrishnan, N.: *Bivariate Conway–Maxwell–Poisson distribution: Formulation, properties, and inference*. J. Multi. Anal. **150**, 152–168 (2016) 11, 68

[34] Sen, S., Lamichhane, R., Diawara, N.: *A bivariate distribution with conditional gamma and its multivariate form*. J. Mod. Appl. Stat. Methods **13(2)**, 169–184 (2014) 22

[35] Seshadri, V., Patil, G.P.: *A characterization of a bivariate distribution by the marginal and the conditional distributions of the same component*. Ann. Inst. Stat. Math. **15**, 215–221 (1964) 2

[36] Statisticat, L.L.C.: *LaplacesDemon: Complete Environment for Bayesian Inference. R package version 16.1.6* 31

[37] Unam, I., Cimat: *Goodness of fit for the Poisson distribution based on the probability generating function*. Commun. Statist.-Theory Meth. **20(10)**, 3093–3110 (1991) 55

Publications & Conferences

Publications in Journals

1. Barry C. Arnold, Banoth Veeranna, B. G. Manjunath and B. Shobha (2022), *"Bayesian inference for pseudo-Poisson data,"* Journal of Statistical Computation and Simulation.https://doi.org/10.1080/00949655.2022.2123918

2. Banoth Veeranna, B. G. Manjunath and B. Shobha (2022), *"Goodness of fit tests for the pseudo-Poisson distribution."* (Under revision)

3. Banoth Veeranna (2022), *"A note on Bayesian inference for the bivariate pseudo-exponential data."* (Submitted).

Presentation in Conferences

1. International Conference on Emerging Innovations in Statistics & Operations Research (EISOR-2018) in conjunction with 38^{th} Annual Convention of ISPS & 4^{th} Convention of IARS, organized by "Department of Statistics, M.D. University, Rohtak - 124001 (INDIA)". Title of the paper: *"Poisson Conditionals,"* held on $27^{th} - 30^{th}$ December, 2018.

2. International Virtual Conference on Prof. C.R. Rao's School of Thought on Statistical Sciences (ICON-CRRAO-STOSS 2020) organized by Department of Statistics, Pondicherry University, and Department of Mathematics and Statistics, and Mrs. A.V.N. College, Visakhapatnam, held on $21^{st} - 22^{nd}$ & $28^{th} - 29^{th}$ November, 2020. Title of the paper: *"Bayesian Analysis of the pseudo-Poisson distribution."*

3. International Conference (Virtual mode) on Emerging Trends in Statistics and Data Science in conjunction with 40^{th} Annual Convention of ISPS on $07^{th} - 10^{th}$ September, 2021. Jointly organized by the Department of Statistics of Cochin University of Science & Technology, Cochin, M.D. University, Rohtak, University of Kerala, Trivandrum, Bharathiar University, Coimbatore, and The Madura College(Autonomous), Madurai. Title of the paper: *"Goodness-of-Fit test for the bivariate pseudo-Poisson distribution."*

4. International Conference on Knowledge Discovers on Statistical Innovations & Recent Advances in Optimization (ICON-KSRAO-2022) organized by Department of Statistics, Andhra University, Visakhapatnam on 29^{th} & 30^{th} December, 2022. Title of the paper: *A Note on Goodness-of-Fit Tests for the Bivariate Pseudo-Poisson Model*

5. International Conference on Statistics, Probability, Data Science and Related Areas (ICSPDS2023) in conjunction with 42^{nd} Annual Convention of Indian Society for Probability and Statistics (ISPS), held during $04^{th} - 06^{th}$ January, 2023, organized jointly by the Department of Statistics, Cochin University of Science and Technology & Indian Society for Probability and Statistics (ISPS). Title of the paper: *"Some remarks on pseudo-Poisson and its goodness-of-fit."*

Workshops & Participation Certificates

Attended Workshops

1. National Workshop on ”Number Theory and Its Applications” (NWNTA-2018) conducted by School of Mathematics and Statistics, University of Hyderabad, Hyderabad during $19^{th} - 20^{th}$ March, 2018.

2. Workshop on ”Time Series and Financial Modelling” conducted by School of Mathematics and Statistics, University of Hyderabad, Hyderabad during March $23^{rd} - 24^{th}$, 2018.

3. Pre-Conference Workshop on “Modern Regression Methods Using R” organized by the P.G. Department of Statistics, Utkal University, Odisha, in conjunction with ISPS on 20^{th} December, 2019.

4. International Webinar on “Advances in Statistics and Data Science for Sustainable Human Development” organized by the Indian Society for Probability & Statistics (ISPS), held on 7^{th}–10^{th} September, 2020.

5. International Webinar on “Data Science Using PYTHON” organized by the Department of Statistics, Bharathiar University, Coimbatore, Tamil Nadu, India. From $10^{th} - 11^{th}$ September, 2020.

6. One-week online workshop on “Statistical Computing using R” held at DST-Centre for Interdisciplinary Mathematical Sciences, Institute of Science, Banaras Hindu University, from 11^{th}–16^{th} October, 2020.

7. Virtual workshop on “Response Surface and Mixture Experiment Methodologies for Agricultural and Fisheries Experimentation” organized by ICAR-CMFRI on 16^{th} December, 2020.

8. JIT Webinar topic titled “Design of Experiments” the “Live two days JIT Webinar series” organized by Jagannath Institute of Technology, on 22^{nd} & 23^{rd} January, 2021.

9. JIT Webinar topic titled “Data Mining in Health Care” the “Live two days JIT Webinar series” organized by Jagannath Institute of Technology, on 22^{nd} & 23^{rd} January, 2021.

10. Prof. C. R. Rao Birth Centenary Celebrations Two days webinar on "Awareness of Statistical Software's (R, Python, SAS, and SPSS) in conjunction with the Academic Competitions of ISPS" organized by Indian Society for Probability & Statistics (ISPS), held on $13^{th} - 14^{th}$ March, 2021.

11. AICTE Training and Learning (ATAL) Academy Online Elementary FDP on "Data Science with Statistical Methods Using R Programming-Basic Course" organized by Amity University, Uttar Pradesh, Lucknow Campus, on 21^{st}–25^{th} June, 2021.

12. Eighth International (Virtual) Workshop on Distribution Theory and its Applications - 2022 organized by the International Statistics Fraternity (ISF), School of Physical and Mathematical Sciences and department of Statistics, University of Kerala, Trivandrum during $13^{th} - 15^{th}$ December, 2022.

Participation Certificates

1. International Conference on "Recent Advances in Statistics and Data Science for Sustainable Development" organized by the Indian Society for Probability & Statistics (ISPS) at Utkal University, Odisha, held on 21^{st}–23^{rd} December, 2019.

2. Prof. C. R. Rao Birth Centenary Conference on "Statistics and Applications" organized by C R Rao AIMSCS and sponsored by DST-SERB at Hyderabad, India, on $6^{th} - 7^{th}$ February, 2020.

3. 13^{th} Plant Sciences Colloquium held at the University of Hyderabad, organized by Department of Plant Sciences, School of Life Sciences, UoH, Hyderabad, India. On 29^{th} October, 2021.

Awards & Achievements

Award

1. The Rajiv Gandhi National Fellowship Scheme For Higher Education of ST Students was awarded on April 2018. Award number: 201819-NFST-TEL-00347.

Achievements

1. Research Scholars School board member of the School of Mathematics and Statistics on August 2018 to September 2019.

2. The Hostel General Secretary of the J & K Men's hostel on September 2019 to September 2020.

Bivariate Pseudo Modeling

by Banoth Veeranna

Librarian
Indira Gandhi Memorial Library
UNIVERSITY OF HYDERABAD
Central University P.O.
HYDERABAD-500 046.

Submission date: 31-Jan-2023 02:16PM (UTC+0530)
Submission ID: 2003219722
File name: BANOTH_VEERANNA.pdf (6.89M)
Word count: 25155
Character count: 102754

Bivariate Pseudo Modeling

ORIGINALITY REPORT

31%	8%	29%	2%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Barry C. Arnold, Banoth Veeranna, B. G. Manjunath, B. Shobha. "Bayesian inference for pseudo-Poisson data", Journal of Statistical Computation and Simulation, 2022 Publication	22%
2	arxiv.org Internet Source	2%
3	Barry C. Arnold, Matthew A. Arvanitis. "On bivariate pseudo-exponential distributions", Journal of Applied Statistics, 2019 Publication	1%
4	Barry C. Arnold, Bangalore G. Manjunath. "Pseudo-Poisson Distributions with Concomitant Variables", Mathematical and Computational Applications, 2023 Publication	1%
5	"Conditional Specification of Statistical Models", Springer Science and Business Media LLC, 2004 Publication	<1%

Self paper (student's own work)

Dr. B.G. MANJUNATH
Assistant Professor
School of Mathematics & Statistics
University of Hyderabad
Hyderabad-500 046. T.S.

6	"International Encyclopedia of Statistical Science", Springer Science and Business Media LLC, 2011 Publication	<1%
7	edoc.pub Internet Source	<1%
8	Norman L. Johnson, Adrienne W. Kemp, Samuel Kotz. "Univariate Discrete Distributions", Wiley, 2005 Publication	<1%
9	Submitted to University of Newcastle upon Tyne Student Paper	<1%
10	uilspace.unilorin.edu.ng Internet Source	<1%
11	digitalcommons.njit.edu Internet Source	<1%
12	A. C. Gleeson, J. B. Douglas. "QUADRAT SAMPLING AND THE ESTIMATION OF NEYMAN TYPE A AND THOMAS DISTRIBUTIONAL PARAMETERS1", Australian Journal of Statistics, 1975 Publication	<1%
13	Submitted to University of Pretoria Student Paper	<1%

14	Barry C. Arnold, B. G. Manjunath. "Statistical inference for distributions with one Poisson conditional", Journal of Applied Statistics, 2021 Publication	<1 %
15	www.mdpi.com Internet Source	<1 %
16	ActEd Publication	<1 %
17	mafiadoc.com Internet Source	<1 %
18	vibdoc.com Internet Source	<1 %
19	N. Unnikrishnan Nair, P.G. Sankaran, N. Balakrishnan. "Reliability Theory", Elsevier BV, 2018 Publication	<1 %
20	Submitted to University of Mauritius Student Paper	<1 %
21	fatimahsity.blogspot.com Internet Source	<1 %
22	E. C. SKOWRONSKI. "Geometric criterion for the design of a non-oscillatory dynamical system†", International Journal of Control, 1972 Publication	<1 %

23	Fekadu Mihret Geremew, Talabattula Srinivas. "Reflective coupled microring resonators for reconfigurable photonic systems: Performance analysis", Results in Optics, 2021 Publication	<1%
24	groll.userweb.mwn.de Internet Source	<1%
25	proceedings.mlr.press Internet Source	<1%
26	"Applied Probability and Statistics", Springer Science and Business Media LLC, 2005 Publication	<1%
27	N. Balakrishnan, E. Chimitova, M. Vedernikova. "An Empirical Analysis of Some Nonparametric Goodness-of-Fit Tests for Censored Data", Communications in Statistics - Simulation and Computation, 2014 Publication	<1%
28	David D. Hanagal. "Modeling heterogeneity for bivariate survival data by the Weibull distribution", Statistical Papers, 12/05/2008 Publication	<1%
29	idoc.pub Internet Source	<1%
30	www.scribd.com Internet Source	<1%

31	personal.vu.nl Internet Source	<1 %
32	the-eye.eu Internet Source	<1 %
33	Saralees Nadarajah, Samuel Kotz. "Some bivariate beta distributions", Statistics, 2005 Publication	<1 %
34	repository.unmul.ac.id Internet Source	<1 %
35	cdn.cityandguilds.com Internet Source	<1 %
36	docplayer.com.br Internet Source	<1 %
37	dokumen.pub Internet Source	<1 %
38	repositorio.ufc.br Internet Source	<1 %
39	vdoc.pub Internet Source	<1 %
40	www.docme.ru Internet Source	<1 %
41	www.math.uiuc.edu Internet Source	<1 %
42	www.minmarks.com	

Internet Source

<1%

43	a-roshani.ir Internet Source	<1%
44	citeseerx.ist.psu.edu Internet Source	<1%
45	Li Zhu, Kimberly F. Sellers, Darcy Steeg Morris, Galit Shmueli. "Bridging the Gap: A Generalized Stochastic Process for Count Data", The American Statistician, 2017 Publication	<1%

Exclude quotes	On	Exclude matches	< 14 words
Exclude bibliography	On		