



### 저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원 저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리와 책임은 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)



이학석사 학위논문

Empirical Bayes Estimation of Negative  
Binomial Models And Its Applications

음이항 분포 모형의  
경험적 베이즈 추정과 그 응용

2019년 2월

서울대학교 대학원

통계학과

최재철

## **Abstract**

# **Empirical Bayes Estimation of Negative Binomial Models And Its Applications**

최재철(CHOI JAECHÉOL)

통계학과(Department of Statistics)

The Graduate School

Seoul National University

Data such as the number of accidents in a location and the number of children per women is characterized by the facts that one or a small number of observations are available in each subject and the observation values are nonnegative integers. When statistical inference is performed based on these types of data, empirical Bayes estimation methods based on Poisson distribution have been commonly used. However, the Poisson distribution has the limitation that the mean and the variance are the same. The negative binomial distribution and the zero-inflated Poisson distribution have the same support as that of the Poisson distribution, but with two parameters, a more flexible model fit is possible. In this thesis, we explore and compare various methods that fit the negative binomial distribution model and zero-inflated Poisson model in a situation where there is only one datum per a subject. We also applied them to some case studies including automobile insurance claims data and fatal traffic accidents data.

---

**Keywords:** Poisson Compound Decision Problem, Empirical Bayes Estimation, Negative Binomial Distribution, Zero-Inflated Poisson Distribution, Unbiased Risk Estimator

**Student Number:** 2017-28258

# Contents

<b>1</b>	<b>Introduction</b>	<b>5</b>
<b>2</b>	<b>Models</b>	<b>7</b>
2.1	Poisson Model . . . . .	7
2.1.1	Properties of Poisson Model . . . . .	8
2.1.2	Estimation Strategies for Poisson Model . . . . .	8
2.1.3	Limitations of Poisson Model . . . . .	9
2.2	Negative Binomial Model . . . . .	10
2.2.1	Constant Shape Parameter Method . . . . .	10
2.2.2	Estimation Strategies for The Constant Shape Negative Bi-nomial Model . . . . .	12
2.2.3	Conditional Prior Method . . . . .	13
2.2.4	Estimation Stratigy for NB Prior Models . . . . .	15
2.3	Zero-Inflated Poisson Modeling . . . . .	15
2.3.1	Zero-Inflated Poisson density . . . . .	15
2.3.2	Estimation Strategy for Zero-Inflated Poisson Model . . . . .	16
<b>3</b>	<b>Simulations</b>	<b>17</b>

3.1	Estimation and Inference	19
3.1.1	Beta Negative Binomial	19
3.1.2	Gamma Poisson	20
3.1.3	Zero-Infaltered Poisson	21
3.1.4	Uniform Poisson	22
3.1.5	Known Uniform Poisson	22
<b>4</b>	<b>Case Studies</b>	<b>24</b>
4.1	Automobile Insurance Claims Data	24
4.2	Fatal Traffic Accidents Data	26
<b>5</b>	<b>Conclusion</b>	<b>28</b>

# List of Tables

2.1	Table of Conjugate Distributions . . . . .	12
3.1	Beta Negative Binomial Estimated parameters . . . . .	19
3.2	Beta Negative Binomial PMSE . . . . .	20
3.3	Gamma Poisson Estimated parameters . . . . .	20
3.4	Gamma Poisson PMSE . . . . .	20
3.5	Zero-Inflated Poisson Estimated parameters . . . . .	21
3.6	Zero-Inflated Poisson PMSE . . . . .	21
3.7	Uniform Poisson Estimated parameters . . . . .	22
3.8	Uniform Poisson PMSE . . . . .	22
3.9	Known Uniform Poisson Estimated Parameters . . . . .	23
3.10	Known Uniform Poisson PMSE . . . . .	23
4.1	Autumobile Insurance Claims Data . . . . .	24
4.2	MSD in Claims Data Case . . . . .	25
4.3	Claims Data Estimated parameters . . . . .	25
4.4	Claims Predictive Values . . . . .	25
4.5	Fatal Traffic Accidents Counts . . . . .	26
4.6	Accidents Data Estimated parameters . . . . .	27

4.7	Accidents Poisson PMSE . . . . .	27
4.8	Accidents Predictive Values . . . . .	27

# Chapter 1

## Introduction

This thesis is focused on the comparison of various empirical Bayes models for count data. This study has applications in three areas. First, data such as the number of children per women is observed only once per women, inferences of this type of data directly depends on the statistical models in this thesis. Secondly, we consider the number of accidents in a location per year which can be observed with repetitions, in the first year of the observation. Lastly, as the number of categorical covariates is increasing nowadays, the numbers of the observations for specific cases are often sparse even for large data. Then the models in this thesis will be applied to each specific case.

The goals are twofold. First, it is of great interest to estimate hyperparameters of prior distributions. In some cases, prior distribution itself gives intuition or comprehension of a phenomenon. Second, we often focus on predictions.

Three popular models for count data are Poisson, negative binomial, and zero-inflated Poisson models. Since the Poisson model has only one parameter, it is popular for data without repetitions. With the advantage of the em-

pirical Bayes model, two parameter models, the negative binomial model and the zero-inflated Poisson model, can be applied for data without repetition. Properties of each model will be studied in this thesis.

Although this thesis does not consider the case for data with repetitions or predictors, model improvements for such situations are expected to be analogous to those of normal empirical Bayes models.

Robbins (1985) first proposed a nonparametric empirical Bayes model for prediction replacing the prior distribution to the marginal distribution in the Poisson model. However, this nonparametric estimation has the weakness if observations are rare. Efron and Hastie (2016) proposed gamma prior empirical Bayes methods for the Poisson model. Leon-Novelo et al. (2015) proposed *F*-distribution prior empirical Bayes model for negative binomial model and applied it to RNA-Seq data. Xie et al. (2016) generalized empirical Bayes estimation for models with quadratic variance using URE estimation with the assumption of known variances. We generalized *F*-distribution prior in Leon-Novelo et al. (2015) to Beta Prime distribution, known variances in Xie et al. (2016) to unknown variance, and developed some expanded models.

This thesis is organized as follows. In Chapter 2, We first explore the basic Poisson model and the limitations of it. Next, as an alternative, we introduce definitions and lemmas related to the negative binomial model, the zero-inflated Poisson model, and their transforms. We also propose parameters estimation methods for each models. In Chapter 3, the performances of each model will be compared by simulations. In Chapter 4, We apply the proposed methods to automobile insurance claims data and fatal traffic accidents data. Conclusions can be found in Chapter 5.

# Chapter 2

## Models

### 2.1 Poisson Model

We first define three notations. First, we denote the distribution of  $X$  is  $gamma(\alpha, \beta)$  when  $pdf(X = x | \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$ ,  $x \geq 0$ . Second, we denote  $\binom{x}{k} \equiv \frac{x(x-1)\cdots(x-k+1)}{k!} = \frac{\Gamma(x+1)}{\Gamma(k+1)\Gamma(x-k+1)}$  for  $x \in \mathbb{R}$ ,  $k \in \{0, 1, 2, \dots\}$ . That is, the support of  $x$  is expanded to real numbers. For the third notation, we reparametrize the negative binomial distribution. Commonly, we say  $X | p, k$  follows a negative binomial distribution when  $P(X = x | p, k) = \binom{x+k-1}{x} p^x (1-p)^k$ ,  $x=0, 1, \dots$  where  $p$  is the probability of success and  $k$  is the number of failures. That is,  $P(X = x | p, k)$  is the probability that there are  $x$  successes before  $k$ -th failure. Note that given the parameters  $p$ ,  $k$ , the mean  $\frac{pk}{1-p}$  is always smaller than the variance  $\frac{pk}{(1-p)^2}$ . As the third notation, we denote  $X | \mu, \alpha$  follows negative binomial distribution  $NB(\mu, \alpha)$  when  $P(X = x | \mu, \alpha) = \binom{x+\frac{1}{\alpha}-1}{x} \left(\frac{\mu\alpha}{\mu\alpha+1}\right)^x \left(\frac{1}{\mu\alpha+1}\right)^{\frac{1}{\alpha}}$ ,  $x=0, 1, \dots$ . Here,  $\alpha = \frac{1}{k}$  is the shape parameter and  $\mu = \frac{pk}{1-p}$  is the expected value of the number of success before  $k$ -th failure.

Then we get the fact that given the parameters  $\mu$  and  $\alpha$ , the mean is  $\mu$  and the variance is  $\mu+\mu^2\alpha$ .

### 2.1.1 Properties of Poisson Model

Suppose  $\theta$  follows  $\text{gamma}(\alpha, \beta)$  and  $X | \theta$  follows the Poisson distribution with mean  $\theta$ . Then the marginal density is

$$f_{\alpha, \beta}(x) = \int_0^\infty P(x | \theta) \pi(\theta) d\theta = \int_0^\infty \frac{\theta^x e^{-\theta}}{x!} \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} e^{-\beta\theta} d\theta = \frac{1}{x!} \frac{\Gamma(x + \alpha) \beta^\alpha}{\Gamma(\alpha) (\beta + 1)^{x+\alpha}}$$

Since  $P(\theta | x)$  is proportional to  $\theta^{\alpha-1} e^{-\beta\theta} \theta^x e^{-\theta}$ , the posterior density of  $\theta | X = x$  is  $\text{gamma}(x + \alpha, \beta + 1)$ . The predictive distribution is

$$\begin{aligned} P(Y = y | x) &= \int_0^\infty P(Y = y | \theta) \pi(\theta | x) d\theta = \int_0^\infty \frac{\theta^y e^{-\theta}}{y!} \frac{(\beta + 1)^{x+\alpha}}{\Gamma(x + \alpha)} \theta^{x+\alpha-1} e^{-(\beta+1)\theta} d\theta \\ &= \frac{1}{y!} \frac{\Gamma(y + x + \alpha) (\beta + 1)^{x+\alpha}}{\Gamma(x + \alpha) (\beta + 2)^{x+y+\alpha}} = \binom{y + x + \alpha - 1}{y} \left( \frac{1}{\beta + 2} \right)^y \left( \frac{\beta + 1}{\beta + 2} \right)^{x+\alpha} \end{aligned}$$

### 2.1.2 Estimation Strategies for Poisson Model

Suppose we observed  $X_1, \dots, X_N$  for  $\theta_1, \dots, \theta_N$  respectively and we assume each  $\theta_i | \alpha, \beta$  follows identically and independently  $\text{gamma}(\alpha, \beta)$

#### MLE based on marginal density

Using the marginal density, MLE can be calculated based on log-likelihood. Then, we get estimators  $(\hat{\alpha}, \hat{\beta}) = \text{argmax} \sum_{i=1}^N \log f_{\alpha, \beta}(x_i) = \text{argmax} \sum_{x=0}^{x_{\max}} y_x \log f_{\alpha, \beta}(x)$ . Here,  $x_{\max}$  is the maximum value of  $x_i$ 's observed and  $y_x$  is the number of  $x_i$ 's the same with  $x$ . We call it PoiMLE Estimator.

## URE Estimator

Unbiased Risk Estimator(URE) with respect to mean vector also can be applied. Using the posterior distribution, we define  $\hat{\theta}_i = E(\theta_i | X_i) = \frac{X_i + \alpha}{\beta + 1} = (1 - w)X_i + w\theta$  where  $w = \frac{\beta}{1 + \beta}$ ,  $\theta = \frac{\alpha}{\beta}$ . Let  $\hat{\Theta} = (\hat{\theta}_1, \dots, \hat{\theta}_N)$ .

We first show  $URE(\alpha, \beta) = \frac{1}{N} \sum_{i=1}^N \left[ \left( \frac{\beta}{1 + \beta} \right)^2 \left( X_i - \frac{\alpha}{\beta} \right)^2 + \left( 1 - 2 \frac{\beta}{1 + \beta} \right) X_i \right]$ . Note that

$$\begin{aligned} & \frac{1}{N} \sum_{i=1}^n E_{\theta_i} \left[ \left( \frac{\beta}{1 + \beta} \right)^2 \left( X_i - \frac{\alpha}{\beta} \right)^2 + \left( 1 - 2 \frac{\beta}{1 + \beta} \right) X_i \right] \\ &= \frac{1}{N} \sum_{i=1}^N E_{\theta_i} [w^2(X_i - \theta)^2 + (1 - 2w)X_i] \\ &= \frac{1}{N} \sum_{i=1}^N [w^2(Var_{\theta_i}(X_i) + (\theta_i - \theta)^2) + (1 - 2w)Var_{\theta_i}(X_i)] \\ &= \frac{1}{N} \sum_{i=1}^N [(1 - w)^2 Var_{\theta_i}(X_i) + w^2(\theta_i - \theta)^2] = \frac{1}{N} \sum_{i=1}^N E_{\mu_i} [\hat{\theta}_i - \theta_i]^2 = Risk(\hat{\theta}, \theta) \end{aligned}$$

Hence, we estimate parameters as  $(\hat{\alpha}, \hat{\beta}) = \text{argmin } URE(\alpha, \beta)$ . We call it PoiURE Estimator.

It is proved in Xie et al. (2012) that under regular conditions, URE estimator is not only consistent but also satisfies the risk optimality while MLE does not.

### 2.1.3 Limitations of Poisson Model

Recall that the mean and variance of the Poisson distribution is the same. However it is not uncommon to observe that the variance is bigger than the mean in the real analysis. Negative binomial distribution has always bigger variance than its mean, so it is usually considered as a good alternative to the

Poisson model in over-dispersion problems. In other case, data sometimes features inflated frequency on the zero value. zero-inflated Poisson model is usually considered as the solution to this problem. We will explore empirical Bayes versions of these models.

## 2.2 Negative Binomial Model

We begin with four notations. First, we denote  $B(\alpha, \beta) \equiv \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$ . Second, we denote  $X | \delta_1, \delta_2$  follows a beta distribution  $Beta(\delta_1, \delta_2)$  when  $pdf(X = x | \delta_1, \delta_2) = \frac{1}{B(\delta_1, \delta_2)} x^{\delta_1-1} (1-x)^{\delta_2-1}$  for  $0 \leq x \leq 1$ . Third, we denote  $X | \delta_1, \delta_2$  follows a beta prime distribution  $BP(\delta_1, \delta_2)$  when  $pdf(X = x | \delta_1, \delta_2) = \frac{1}{B(\delta_1, \delta_2)} \left(\frac{x}{x+1}\right)^{\delta_1-1} \left(\frac{1}{x+1}\right)^{\delta_2+1}$  for  $0 \leq x$ . Then the mean is  $\frac{\delta_1}{\delta_2+1}$  and the variance is  $\frac{\delta_1(\delta_1+\delta_2-1)}{(\delta_2+2)(\delta_2+1)^2}$ . we remark that  $X$  follows  $Beta(\delta_1, \delta_2)$  only and if only  $\frac{X}{1-X}$  follows  $BP(\delta_1, \delta_2)$ . Fourth, we denote  $X | k, \delta_1, \delta_2$  follows a beta negative binomial distribution  $BNB(k, \delta_1, \delta_2)$  when  $P(X = x | k, \delta_1, \delta_2) = \binom{x+k-1}{x} \frac{B(x+\delta_1, k+\delta_2)}{B(\delta_1, \delta_2)}$  for  $x=0, 1, 2, \dots$ .

### 2.2.1 Constant Shape Parameter Method

We first assume that  $\alpha = \frac{1}{k}$  is fixed(known constant). Then conjugate Prior of  $\mu\alpha$  is Beta prime Distribution i. e.  $\pi(\mu\alpha | \delta_1, \delta_2) = \frac{1}{B(\delta_1, \delta_2)} \left(\frac{\mu\alpha}{\mu\alpha+1}\right)^{\delta_1-1} \left(\frac{1}{\mu\alpha+1}\right)^{\delta_2+1}$  for  $\delta_1, \delta_2 > 0$  or equivalently,  $\pi(\mu | \delta_1, \delta_2) = \frac{\alpha}{B(\delta_1, \delta_2)} \left(\frac{\mu}{\mu+1}\right)^{\delta_1-1} \left(\frac{1}{\mu+1}\right)^{\delta_2+1}$  for  $\delta_1, \delta_2 > 0$ . It can be derived using the fact that if  $p \equiv \frac{\mu\alpha}{\mu\alpha+1}$  then  $p \sim Beta(\delta_1, \delta_2)$ .

The marginal density of X is  $\text{BNB}(\frac{1}{\alpha}, \delta_1, \delta_2)$  in that

$$\begin{aligned}
f_{\delta_1, \delta_2}(x) &= \int_0^\infty NB(x | \mu, \alpha) \times \pi(\mu\alpha | \delta_1, \delta_2) d\mu\alpha \\
&= \int_0^\infty \binom{x + \frac{1}{\alpha} - 1}{x} \left( \frac{\mu\alpha}{\mu\alpha + 1} \right)^x \left( \frac{1}{\mu\alpha + 1} \right)^{\frac{1}{\alpha}} \\
&\quad \times \frac{1}{B(\delta_1, \delta_2)} \left( \frac{\mu\alpha}{\mu\alpha + 1} \right)^{\delta_1 - 1} \left( \frac{1}{\mu\alpha + 1} \right)^{\delta_2 + 1} d\mu\alpha \\
&= \binom{x + \frac{1}{\alpha} - 1}{x} \frac{B(x + \delta_1, \frac{1}{\alpha} + \delta_2)}{B(\delta_1, \delta_2)} = d\text{BNB}\left(\frac{1}{\alpha}, \delta_1, \delta_2\right)
\end{aligned}$$

Since  $\pi(\mu\alpha | x) \propto \left( \frac{\mu\alpha}{\mu\alpha + 1} \right)^{\delta_1 - 1} \left( \frac{1}{\mu\alpha + 1} \right)^{\delta_2 + 1} \left( \frac{\mu\alpha}{\mu\alpha + 1} \right)^x \left( \frac{1}{\mu\alpha + 1} \right)^{\frac{1}{\alpha}}$ , the posterier distribution of  $\mu\alpha$  is  $\pi(\mu\alpha | x) = \frac{1}{B(x + \delta_1, \frac{1}{\alpha} + \delta_2)} \left( \frac{\mu\alpha}{\mu\alpha + 1} \right)^{x + \delta_1 - 1} \left( \frac{1}{\mu\alpha + 1} \right)^{\frac{1}{\alpha} + \delta_2 + 1}$   
 $= \text{dBP}(x + \delta_1, \frac{1}{\alpha} + \delta_2)$ . The predictive density is

$$\begin{aligned}
P(Y = y | x) &= \int_0^\infty P(Y = y | \mu\alpha = t) \pi(\mu\alpha = t | x) dt \\
&= \int_0^\infty \binom{y + \frac{1}{\alpha} - 1}{y} \left( \frac{t}{t + 1} \right)^y \left( \frac{1}{t + 1} \right)^{\frac{1}{\alpha}} \\
&\quad \times \frac{1}{B(x + \delta_1, \frac{1}{\alpha} + \delta_2)} \left( \frac{t}{t + 1} \right)^{x + \delta_1 - 1} \left( \frac{1}{t + 1} \right)^{\frac{1}{\alpha} + \delta_2 + 1} dt \\
&= \binom{y + \frac{1}{\alpha} - 1}{y} \frac{B(y + x + \delta_1, \frac{1}{\alpha} + \frac{1}{\alpha} + \delta_2)}{B(x + \delta_1, \frac{1}{\alpha} + \delta_2)} = d\text{BNB}\left(\frac{1}{\alpha}, x + \delta_1, \frac{1}{\alpha} + \delta_2\right)
\end{aligned}$$

Hence,  $E(Y | x) = \frac{x + \delta_1}{1 + \alpha(\delta_2 - 1)}$  if  $\frac{1}{\alpha} + \delta_2 > 1$ ,  $\infty$  otherwise.

The above results can be summarised as the table bellow.

Parameter	$p_i$	$\mu_i \alpha$
Prior	Beta( $\delta_1, \delta_2$ )	BP( $\delta_1, \delta_2$ )
Sampling	NB( $p_i, k$ )	NB( $\mu_i, \alpha$ )
Marginal	BNB( $k, \delta_1, \delta_2$ )	$BNB(\frac{1}{\alpha}, \delta_1, \delta_2)$
Posterior	Beta( $X_i + \delta_1, k + \delta_2$ )	$BP(X_i + \delta_1, \frac{1}{\alpha} + \delta_2)$
Predictive	$BNB(k, X_i + \delta_1, k + \delta_2)$	$BNB(\frac{1}{\alpha}, X_i + \delta_1, \frac{1}{\alpha} + \delta_2)$

Table 2.1: Table of Conjugate Distributions

## 2.2.2 Estimation Strategies for The Constant Shape Negative Binomial Model

As in 2.1.2, Suppose we observed  $X_1, \dots, X_N$  for  $\theta_1, \dots, \theta_N$  respectively and  $x_{max}$  is the maximum value of  $x_i$ s observed while  $y_x$  is the number of  $x_i$ s the same with x.

### MLE based on marginal density

Suppose  $X_i | \mu_i$  follows independently and identically  $NB(\mu_i, \alpha)$  for  $i=1, 2, \dots, N$ . Then using the marginal distribution, we get  $(\hat{\delta}_1, \hat{\delta}_2, \hat{\alpha}) = \text{argmax}_{x=0}^{x_{max}} y_x \log f_{\delta_1, \delta_2}(x)$ . We call it NBFixMLE.

### URE Estimator

We first show  $URE(\delta_1, \delta_2, \alpha) = \frac{1}{n} \sum_{i=1}^n \left[ \left( \frac{\alpha(\delta_2 - 1)}{1 + \alpha(\delta_2 - 1)} \right)^2 \left( X_i - \frac{\delta_1}{\alpha(\delta_2 - 1)} \right)^2 + \left( 1 - 2 \frac{\alpha(\delta_2 - 1)}{1 + \alpha(\delta_2 - 1)} \right) \frac{X_i + \alpha X_i^2}{1 + \alpha} \right]$ . Let  $w = \frac{\alpha(\delta_2 - 1)}{1 + \alpha(\delta_2 - 1)}$ ,  $\mu = \frac{\delta_1}{\alpha(\delta_2 - 1)}$ . Note that from the posterior distribution,

$$E(\mu_i | X_i) = \hat{\mu}_i = \frac{1}{1 + \alpha(\delta_2 - 1)} X_i + \frac{\alpha(\delta_2 - 1)}{1 + \alpha(\delta_2 - 1)} \frac{\delta_1}{\alpha(\delta_2 - 1)} = (1 - w) X_i + w \mu$$

and since  $E_{\mu_i}(X_i + \alpha X_i^2) = \mu_i + \alpha(\mu_i^2 + Var_{\mu_i}(X_i)) = (1 + \alpha)Var_{\mu_i}(X_i)$ , we get  $E_{\mu_i}\left(\frac{X_i + \alpha X_i^2}{1 + \alpha}\right) = Var_{\mu_i}(X_i)$ . Therefore,

$$\begin{aligned} & \frac{1}{n} \sum_{i=1}^n E_{\mu_i} \left[ \left( \frac{\alpha(\delta_2 - 1)}{1 + \alpha(\delta_2 - 1)} \right)^2 \left( X_i - \frac{\delta_1}{\alpha(\delta_2 - 1)} \right)^2 + \left( 1 - 2 \frac{\alpha(\delta_2 - 1)}{1 + \alpha(\delta_2 - 1)} \right) \frac{X_i + \alpha X_i^2}{1 + \alpha} \right] \\ &= \frac{1}{n} \sum_{i=1}^n E_{\mu_i} \left[ w^2 (X_i - \mu)^2 + (1 - 2w) \frac{X_i + \alpha X_i^2}{1 + \alpha} \right] \\ &= \frac{1}{n} \sum_{i=1}^n [w^2 (Var_{\mu_i}(X_i) + (\mu_i - \mu)^2) + (1 - 2w) Var_{\mu_i}(X_i)] \\ &= \frac{1}{n} \sum_{i=1}^n [(1 - w)^2 Var_{\mu_i}(X_i) + w^2 (\mu_i - \mu)^2] = \frac{1}{n} \sum_{i=1}^n E_{\mu_i} [\hat{\mu}_i - \mu_i]^2 = Risk(\hat{\mu}, \mu) \end{aligned}$$

Now we have  $(\hat{\delta}_1, \hat{\delta}_2, \hat{\alpha}) = \text{argmin } URE(\delta_1, \delta_2, \alpha)$ . We call it NBURE

### 2.2.3 Conditional Prior Method

As we showed in 2.2.1, note that conjugate prior of  $\mu$  given  $\alpha$  is  $\pi(\mu | \alpha, \delta_1, \delta_2) = \frac{\alpha}{B(\delta_1, \delta_2)} \left( \frac{\mu\alpha}{\mu\alpha+1} \right)^{\delta_1-1} \left( \frac{1}{\mu\alpha+1} \right)^{\delta_2+1}$ . This can be shown as

$$\begin{aligned} f_{\delta_1, \delta_2}(x | \alpha) &= \int_0^\infty P(x | \mu, \alpha) \pi(\mu | \alpha) d\mu \\ &= \int_0^\infty \frac{\alpha}{B(\delta_1, \delta_2)} \left( \frac{\mu\alpha}{\mu\alpha+1} \right)^{\delta_1-1} \left( \frac{1}{\mu\alpha+1} \right)^{\delta_2+1} \\ &\quad \times \binom{x + \frac{1}{\alpha} - 1}{x} \left( \frac{\mu\alpha}{\mu\alpha+1} \right)^x \left( \frac{1}{\mu\alpha+1} \right)^{\frac{1}{\alpha}} d\mu \\ &= \frac{1}{B(\delta_1, \delta_2)} \binom{x + \frac{1}{\alpha} - 1}{x} \int_0^\infty \alpha \left( \frac{\mu\alpha}{\mu\alpha+1} \right)^{x+\delta_1-1} \left( \frac{1}{\mu\alpha+1} \right)^{\frac{1}{\alpha}+\delta_2+1} d\mu \\ &= \binom{x + \frac{1}{\alpha} - 1}{x} \frac{B(x + \delta_1, \frac{1}{\alpha} + \delta_2)}{B(\delta_1, \delta_2)} = dBNB\left(\frac{1}{\alpha}, \delta_1, \delta_2\right) \end{aligned}$$

There is no known prior density of  $\alpha$  giving a closed-form of posterior density. However, since  $\alpha$  is a shape parameter, it is natural to assume that there is a prior density of  $\alpha$ . We tried Gamma, Beta, and Negative binomial distributions as prior distributions and calculated the posterior numerically. We assume  $1 \leq k$  for both the beta prior and the truncated negative binomial prior.

### **Gamma Prior Method**

Suppose  $\alpha \sim \text{gamma}(u_1, u_2)$  then

$$\begin{aligned} f_{\delta_1, \delta_2, u_1, u_2}(x) &= \int_0^\infty \int_0^\infty P(x | \mu, \alpha) \pi(\mu | \alpha) \pi(\alpha) d\mu d\alpha \\ &= \int_0^\infty \binom{x + \frac{1}{\alpha} - 1}{x} \frac{B(x + \delta_1, \frac{1}{\alpha} + \delta_2)}{B(\delta_1, \delta_2)} \frac{u_2^{u_1}}{\Gamma(u_1)} \alpha^{u_1-1} e^{-u_2\alpha} d\alpha \end{aligned}$$

### **Beta Prior Method**

Suppose  $\alpha \sim \text{Beta}(u_1, u_2)$  then

$$\begin{aligned} f_{\delta_1, \delta_2, u_1, u_2}(x) &= \int_0^1 \int_0^\infty P(x | \mu, \alpha) \pi(\mu | \alpha) \pi(\alpha) d\mu d\alpha \\ &= \int_0^1 \binom{x + \frac{1}{\alpha} - 1}{x} \frac{B(x + \delta_1, \frac{1}{\alpha} + \delta_2)}{B(\delta_1, \delta_2)} \frac{\alpha^{u_1-1} (1-\alpha)^{u_2-1}}{B(u_1, u_2)} d\alpha \end{aligned}$$

### **Truncated Negative Binomial Prior Method**

Suppose  $\frac{1}{\alpha} - 1 \sim NB(m, a)$  then

$$\begin{aligned} f_{\delta_1, \delta_2, m, a}(x) &= \sum_{\frac{1}{\alpha}=1}^{\infty} \int_0^\infty P(x | \mu, \alpha) \pi(\mu | \alpha) \pi(\alpha) d\mu \\ &= \sum_{k=1}^{\infty} \binom{x + k - 1}{x} \frac{B(x + \delta_1, k + \delta_2)}{B(\delta_1, \delta_2)} \binom{k - 1 + \frac{1}{a} - 1}{k - 1} \left( \frac{ma}{ma + 1} \right)^{k-1} \left( \frac{1}{ma + 1} \right)^{\frac{1}{a}} \end{aligned}$$

## 2.2.4 Estimation Stratigy for NB Prior Models

Parameters of each model can be estimated by MLE based on marginal density. We call these NBGamma, NBBeta, NBTNB, respectively. We did not consider URE estimator for above models.

## 2.3 Zero-Inflated Poisson Modeling

### 2.3.1 Zero-Inflated Poisson density

In this section, we consider  $\theta | w, \alpha, \beta$  follows  $w \times \delta_0 + (1-w) \times \text{gamma}(\alpha, \beta)$  where  $\delta_0$  is dirac delta. That is,  $\theta$  has a point mass at zero. Here,  $X | \theta$  follows  $Poi(\theta)$ . We assume when  $\theta=0$ ,  $X=0$  with probability 1. Then the marginal density is

$$f_{\alpha, \beta, w}(x) = \begin{cases} w + (1-w) \int_0^\infty P(x | \theta) \pi(\theta) d\theta, & x = 0 \\ (1-w) \int_0^\infty P(x | \theta) \pi(\theta) d\theta, & x \neq 0 \end{cases}$$

$$= w * \delta_0 + (1-w) \frac{1}{x!} \frac{\Gamma(x+\alpha) \beta^\alpha}{\Gamma(\alpha)(\beta+1)^{x+\alpha}}$$

The posterior density of  $\theta$  is  $w \times \delta_0 + (1-w) \times \text{gamma}(x+\alpha, \beta+1)$  and The predictive distribution is  $w * \delta_0 + (1-w) NB\left(\frac{x+\alpha}{\beta+1}, \frac{1}{x+\alpha}\right)$

### 2.3.2 Estimation Strategy for Zero-Inflated Poisson Model

#### MLE based on marginal density

Suppose  $X_i | \theta_i \stackrel{iid}{\sim} poi(\theta_i)$  for  $i=1, 2, \dots, n$ . Let  $y_x$  be the number of cases in which  $X_i = x$ . Then

$$(\hat{\alpha}, \hat{\beta}, \hat{w}) = argmax \left[ y_0 \times \log \left( w + (1-w) \left( \frac{\beta}{\beta+1} \right)^\alpha \right) + \sum_{x=1}^{x_{max}} y_x \left( \log(1-w) + \log \Gamma(x+\alpha) + \alpha \log \beta - \log \Gamma(\alpha) - (x+\alpha) \log(\beta+1) \right) \right]$$

we call it IP estimator. We did not consider URE estimator for the above model.

# Chapter 3

## Simulations

We compare the proposed estimators via numerical simulations. In each distribution, in each of 50 iterations, two observations were generated for N=5000 subjects. First 5000 observations were used for density estimation and prediction, and the other 5000 observations were used for validation of the prediction. Among models considered, Gamma prior method and Beta prior method were excluded because of their instability. the negative binomial model with fixed  $\alpha$  by URE also came out to be extremely unstable, so we used a truncated model instead. That is,  $UREP = \max(URE, 0)$  is used for optimization.

The results were summarised in two tables. The first table shows the parameter estimated for each model. The second table shows the predictive MSE. Among models considered, five stable models were used for prediction. Based on results in 2, predicted counts were calculated as below. Note that In Poisson

case, the expected number of claims in the next year is as well-known,

$$E(Y | x) = \frac{x + \delta_1}{1 + \delta_2}$$

In the Negative Binomial case, when  $\alpha$  is known, the expected number of claims in the next year is

$$E(Y | x) = \frac{x + \delta_1}{1 + \alpha(\delta_2 - 1)}$$

and when  $\alpha$  is unknown and has prior distribution  $\pi(\alpha)$ , then the expected number of claims in the next year is

$$E(Y | x) = \int_A \frac{x + \delta_1}{1 + \alpha(\delta_2 - 1)} \pi(\alpha) d\alpha$$

where  $A$  is the support of  $\alpha$ . Also, in zero-inflated poisson case,

$$E(Y | x) = (1 - w) \frac{x + \delta_1}{1 + \delta_2}$$

Using plug-in estimators, we get the predicted values of each model. Oracle estimators are calculated using the true parameters.

## 3.1 Estimation and Inference

### 3.1.1 Beta Negative Binomial

Data were generated in the assumptions as follow.

$$p_i \sim \text{beta}(2, 7), k = 2, X_i | p_i \sim NB(p_i, k) \Leftrightarrow \mu_i / 2 \sim BP(2, 7), \alpha = 1/2$$

The results are summarized as bellow.

	$\delta_1$	$\delta_2$	$p_1$	$p_2$
PoiMLE	0.691 (0.037)	1.036 (0.069)	NA (NA)	NA (NA)
PoiURE	0.589 (0.048)	0.882 (0.08)	NA (NA)	NA (NA)
NBFixMLE	1.985 (0.077)	6.908 (0.525)	1.985 (0.077)	NA (NA)
NBUREP	5 (0)	5 (0)	5 (0)	NA (NA)
NBTNB	65.422 (12.407)	112.083 (20.922)	0.145 (0.03)	69.982 (11.07)
IP	0.691 (0.037)	1.036 (0.069)	0 (0)	NA (NA)

Table 3.1: Beta Negative Binomial Estimated parameters

As explained in Chapter 2, The meaning of each parameter is different. For Poisson MLE(PoiMLE) and Poisson URE(PoiURE),  $\theta_i \sim \text{gamma}(\delta_1, \delta_2)$ . For Negative Binomial with Constant  $\alpha$  MLE(NBFixMLE),  $\mu_i \alpha \sim BP(\delta_1, \delta_2)$  and  $p_1 = k = \frac{1}{\alpha}$ . For Negative Binomial with Constant  $\alpha$  truncated URE(NBUREP),  $\mu_i \alpha \sim BP(\delta_1, \delta_2)$  and  $p_1 = \alpha$ . For Negative Binomial with Negative Binomial prior for  $\alpha$  (NBTNB),  $\mu_i \alpha | \alpha \sim BP(\delta_1, \delta_2)$  and  $\frac{1}{\alpha} - 1 \sim NB(p_1, p_2)$ . For zero-inflated Poisson(IP),  $\theta_i \sim w + \delta_0 + (1 - w)\text{gamma}(\delta_1, \delta_2)$  with  $w = p_1$ . Table 3.1 shows that NBFixMLE almost exactly estimated the true parameter value. Furthermore, as  $p_1=w$  is always estimated to be 0, PoiMLE and IP behaved in the same way in this case. Next, we compare the prediction performance.

	PoiMLE	PoiURE	NBFixMLE	NBTNB	IP	oracle
SSEPredMean	1.418	1.449	1.333	1.413	1.418	1.333
SSEPredSd	0.090	0.088	0.093	0.101	0.090	0.093

Table 3.2: Beta Negative Binomial PMSE

Table 3.2 shows that NBFixMLE performed well as the oracle estimator.

### 3.1.2 Gamma Poisson

Data were generated in the assumptions as follow.

$$\theta_i \sim \text{gamma}(2, 7), X_i | \theta_i \sim \text{Poi}(\theta_i)$$

The results are summarized as below.

	$\delta_1$	$\delta_2$	$p_1$	$p_2$
PoiMLE	3.03 (0.186)	2.013 (0.129)	NA (NA)	NA (NA)
PoiURE	3.028 (0.201)	2.012 (0.138)	NA (NA)	NA (NA)
NBFixMLE	4.805 (1.83)	1502.419 (2721.161)	727.517 (1333.231)	NA (NA)
NBUREP	5 (0)	5 (0)	5 (0)	NA (NA)
NBTNB	100 (0)	67.224 (0.884)	0 (0)	349.197 (64.778)
IP	3.166 (0.288)	2.084 (0.177)	0.008 (0.011)	NA (NA)

Table 3.3: Gamma Poisson Estimated parameters

Table 3.3 shows that PoiMLE, PoiURE and IP estimated the parameters well. NBUREP stopped at the initial value at every iterations. Next, we compare prediction performances.

	PoiMLE	PoiURE	NBFixMLE	NBTNB	IP	oracle
SSEPredMean	2.011	2.011	2.034	2.243	2.012	2.010
SSEPredSd	0.052	0.052	0.062	0.057	0.052	0.052

Table 3.4: Gamma Poisson PMSE

Table 3.4 shows that PoiMLE, PoiURE and IP performed well as the oracle estimator.

### 3.1.3 Zero-Infalted Poisson

Data were generated in the assumptions as follow.

$w = 0.1$ ,  $\theta_i \sim w\delta_0 + (1-w)gamma(6, 4)$ ,  $X_i | \theta_i \sim Poi(\theta_i)$ ,  $X_i = 0$  with probability 1 if  $\theta_i = 0$

The results are summarized as bellow.

	$\delta_1$	$\delta_2$	$p_1$	$p_2$
PoiMLE	3.118 (0.207)	2.308 (0.15)	NA (NA)	NA (NA)
PoiURE	3.374 (0.235)	2.497 (0.171)	NA (NA)	NA (NA)
NBFixMLE	5.757 (2.161)	4603.663 (7650.688)	1944.498 (3218.454)	NA (NA)
NBUREP	5 (0)	5 (0)	5 (0)	NA (NA)
NBTNB	100 (0)	74.831 (1.054)	0 (0)	279.777 (15.445)
IP	6.207 (1.258)	4.119 (0.765)	0.101 (0.016)	NA (NA)

Table 3.5: Zero-Inflated Poisson Estimated parameters

Table 3.5 shows that IP estimated the true parameters well. Next, we compare prediction performances.

	PoiMLE	PoiURE	NBFixMLE	NBTNB	IP	oracle
SSEPredMean	1.737	1.736	1.757	1.877	1.760	1.764
SSEPredSd	0.044	0.044	0.051	0.046	0.048	0.041

Table 3.6: Zero-Inflated Poisson PMSE

Table 3.6 shows that PoiMLE, PoiURE and NBFixMLE performed even better than the oracle estimator. More iterations seems to be needed. Here, only NBTNB performed bad.

### 3.1.4 Uniform Poisson

Data were generated in the assumptions as follow.

$\theta_i \sim unif(1,2)$ ,  $X_i | \theta_i \sim Poi(\theta_i)$  The results are summarized as bellow.

	$\delta_1$	$\delta_2$	$p_1$	$p_2$
PoiMLE	34.434 (16.344)	22.991 (10.98)	NA (NA)	NA (NA)
PoiURE	34.685 (16.388)	23.16 (11.009)	NA (NA)	NA (NA)
NBFixMLE	10.234 (0.045)	70 (0)	10.234 (0.045)	NA (NA)
NBUREP	5 (0)	5 (0)	5 (0)	NA (NA)
NBTNB	100 (0)	67.412 (0.597)	0 (0)	381.297 (5.988)
IP	45.209 (22.894)	29.973 (15.179)	0.006 (0.007)	NA (NA)

Table 3.7: Uniform Poisson Estimated parameters

All models considered cannot represent uniform distribution exactly. Therefore, we focus on the prediction performances.

	PoiMLE	PoiURE	NBFixMLE	NBTNB	IP	oracle
SSEPredMean	1.576	1.576	1.586	1.578	1.577	1.576
SSEPredSd	0.044	0.044	0.043	0.044	0.044	0.044

Table 3.8: Uniform Poisson PMSE

Table 3.8 shows that all models performed well.

### 3.1.5 Known Uniform Poisson

Data were generated in the assumptions as follow.

$X_i | \theta_i \sim Poi(\theta_i)$ , where  $\Theta = (\theta_1, \dots, \theta_{200})$ . It is,  $\Theta$  is the sequence from 5 to 15 with length of 200. The results are summarized as bellow.

	d1	d2	p1	p2
PoiMLE	12.086 (2.571)	1.213 (0.264)	NA (NA)	NA (NA)
PoiURE	12.576 (2.591)	1.262 (0.266)	NA (NA)	NA (NA)
NBFixMLE	26.628 (10.201)	68.361 (4.978)	27.466 (7.164)	NA (NA)
NBUREP	196.057 (8.488)	24.976 (2.933)	374.506 (6.862)	NA (NA)
NBTNB	100 (0)	10.596 (0.236)	0 (0)	279.306 (37.413)
IP	12.264 (2.653)	1.23 (0.271)	0.001 (0.002)	NA (NA)

Table 3.9: Known Uniform Poisson Estimated Parameters

All models considered cannot represent uniform distribution exactly. Therefore, we focus on the prediction performances.

	PoiMLE	PoiURE	NBFixMLE	NBTNB	IP	oracle
SSEPredMean	14.698	14.697	15.244	17.262	14.697	10.013
SSEPredSd	1.469	1.474	1.526	1.530	1.473	1.038

Table 3.10: Known Uniform Poisson PMSE

Unlike Table 3.8, Table 3.10 shows that performances were not as good as the oracle estimator. It is because  $\Theta$  is not random for the oracle estimator here. However, PoiMLE, PoiURE, IP showed similarly better performances than others.

From Table 3.2 to Table 3.10, we conclude no estimator which always performs the best. There is no situation where NBTBN performs best.

# Chapter 4

## Case Studies

### 4.1 Automobile Insurance Claims Data

In this session, we consider a European automobile insurance claims data in Efron and Hastie (2016). It is of great interest to predict the claims of policy holders so the insurance company calculates the insurance premium for each policy holder. We apply aforementioned models to the data. First, we will find the best fit for the data, and then predict the claims of each policyholder next year using the results. The data is given in Table 4.1 and the results are given as follows.

Claims	0	1	2	3	4	5	6	7
Counts	7840	1317	239	42	10	4	4	1

Table 4.1: Autumobile Insurance Claims Data

Since there is no extra data to validate in this case, we mainly focus on training error as Efron and Hastie (2016). The results of model fittings are given as

below.

	PoiMLE	PoiURE	NBFixMLE	NBURE	NBUREP	NBTNB	IP
SSE	0.142	0.634	0.050	463.499	40.327	0.016	0.142

Table 4.2: MSD in Claims Data Case

Table 4.2 shows training error of each models. In this case, NBTNB performs best and NBFixMLE performs better than PoiMLE. As we can see in Table 4.3, estimated weight is 0 so IP shows the same result as the original Poisson model does. Based on Table 4.2, we present prediction values using only PoiMLE, PoiURE, NBFixMLE, NBTNB.

	d1	d2	p1	p2
PoiMLE	0.70	3.27		
PoiURE	0.62	2.88		
NBFixMLE	1.86	17.07	1.86	
NBURE	3790.64	3.62	6774.93	
NBUREP	5.00	5.00	5.00	
NBTNB	18.73	97.56	0.12	54.09
IP	0.70	3.27	0.00	

Table 4.3: Claims Data Estimated parameters

	0	1	2	3	4	5	6	7
PoiMLE	0.164	0.398	0.632	0.866	1.100	1.334	1.568	1.802
PoiURE	0.159	0.417	0.675	0.933	1.191	1.449	1.707	1.965
NBFixMLE	0.192	0.296	0.399	0.503	0.606	0.710	0.813	0.917
NBTNB	0.213	0.224	0.236	0.247	0.258	0.270	0.281	0.293

Table 4.4: Claims Predictive Values

Table 4.4 shows the prediction of the next year. While NBTNB showed the best performance in MSD as in Table 4.2, we expect predictive values of NBFixMLE

are reliable most for bad performances of NBTNB in Chapter 3.

## 4.2 Fatal Traffic Accidents Data

In this section, we study an example based on fatal traffic accidents in each municipality of Korea in November 2017. The data is available at ([http://taas.koroad.or.kr/sta/acs/exs/typical.do?menuId=WEB\\_KMP\\_OVT\\_UAS\\_PDS](http://taas.koroad.or.kr/sta/acs/exs/typical.do?menuId=WEB_KMP_OVT_UAS_PDS)). We removed some municipalities with missing observations so 192 municipalities were analyzed. There were four Mondays, Tuesdays, and five Thursdays and Fridays in that month. There were no additional national holidays. For municipality  $i$ , let  $X_i$  be the total number of fatal traffic accidents in Mondays or Thursdays. Similarly, for municipality  $i$ , let  $Y_i$  be the sum of the number of fatal traffic accidents in Tuesdays or Fridays. We predicted  $Y_i$  based on  $X_i$  using models in Chapter 2. The summarized data is given in Table 4.5

	0	1	2	3	4
Mondays or Thursdays	119	51	18	2	2
Tuesdays or Fridays	112	59	16	4	1

Table 4.5: Fatal Traffic Accidents Counts

The results of model fitting are given as belows.

	d1	d2	p1	p2
PoiMLE	2.73	5.19		
PoiURE	2.82	5.35		
NBFixMLE	6.23	74.68	6.23	
NBURE	242.04	2.83	252.41	
NBUREP	5.00	5.00	5.00	
NBTNB	36.54	70.00	0.00	108.59
IP	5.53	9.18	0.13	

Table 4.6: Accidents Data Estimated parameters

Table 4.6 shows the estimated parameters. Note that  $p_1$  of NBTNB means  $k$  is estimated to be 1 with probability 1.  $p_1$  of IP is 0.13, which shows there are more zeros than usual.

	PoiMLE	PoiURE	NBFixMLE	NBTNB	IP
PMSE	119.099	118.933	116.485	115.579	116.746

Table 4.7: Accidents Poisson PMSE

Table 4.7 shows prediction MSE. NBTNB performed best and NBFixMLE IP also show improvements comparing the Poisson models.

	0	1	2	3	4
PoiMLE	0.441	0.603	0.764	0.926	1.087
PoiURE	0.443	0.601	0.758	0.915	1.073
NBFixMLE	0.485	0.563	0.641	0.719	0.797
NBTNB	0.603	0.712	0.821	0.930	1.039
IP	0.474	0.560	0.646	0.731	0.817

Table 4.8: Accidents Predictive Values

Table 4.8 shows prediction values of the models. We can see that MBFixMLE, NBTNB, IP tend to predict more for smaller values and less for bigger values.

# **Chapter 5**

## **Conclusion**

In this thesis, we generalized  $F$ -distribution prior in Leon-Novelo et al. (2015) to Beta Prime distribution, known variances in Xie et al. (2016) to unknown variance, and developed some expanded models for one observation per subject data. Through various simulation results and real data analyses, we conclude that all of the models are not robust, and have high performance based on their own assumptions. Therefore, detecting the situation is important and the analysis of it is needed. Also, models for repeated measurements or covariates are left for future work. Furthermore, nonparametric techniques are also applicable to this type of data.

# Bibliography

- Efron, B. and Hastie, T. (2016). *Computer age statistical inference*, volume 5. Cambridge University Press.
- Leon-Novelo, L., Fuentes, C., and Emerson, S. (2015). Bayesian estimation of negative binomial parameters with applications to rna-seq data. *arXiv preprint arXiv:1512.00475*.
- Robbins, H. (1985). An empirical bayes approach to statistics. In *Herbert Robbins Selected Papers*, pages 41–47. Springer.
- Xie, X., Kou, S., and Brown, L. D. (2012). Sure estimates for a heteroscedastic hierarchical model. *Journal of the American Statistical Association*, 107(500):1465–1479.
- Xie, X., Kou, S. C., and Brown, L. (2016). Optimal shrinkage estimation of mean parameters in family of distributions with quadratic variance. *Annals of statistics*, 44(2):564.

## 요약(국문초록)

특정 지역과 시간대에 일어난 사건의 수, 여성 1인당 자녀 수 등의 자료는 대상별로 관측치가 음 아닌 정수 형태로 하나밖에 없다는 특징이 있다. 이러한 자료를 바탕으로 통계적 추정을 할 때, 포아송 분포를 가정한 경험적 베이즈 추정 방식을 많이 사용해 왔다. 그러나 포아송분포는 평균과 분산이 항상 같다는 제약이 있다. 음이항분포와 0이 기대보다 많이 관측된 포아송분포는 모수가 두 개라는 점에서 보다 유연한 모형적합이 가능하다. 본 논문에서는 1인당 자료가 하나밖에 없는 상황에서 음이항분포 모형과 0이 기대보다 많이 관측된 포아송 분포 모형을 적합하는 다양한 방법을 탐구하고 비교해 보았다. 또한, 이를 보험 신고 자료와 교통사고 자료에 응용해 보았다.

---

**주요어:** 복합 결정 문제, 경험적 베이즈 추정, 음이항분포, 0이 기대보다 많이 관측된 포아송분포, 비편향 리스크 추정량

**학번:** 2017-28258