8-2019

Empirical Bayes Estimators and Borel-Tanner Distribution

Celestina Ruby Soltero
The University of Texas Rio Grande Valley

EMPIRICAL BAYES ESTIMATORS AND BOREL-TANNER DISTRIBUTION

A Thesis

by

CELESTINA RUBY SOLTERO

Submitted to the Graduate College of
The University of Texas Rio Grande Valley
In partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

August 2019

Major Subject: Mathematics

EMPIRICAL BAYES ESTIMATORS AND BOREL-TANNER DISTRIBUTION

A Thesis
by
CELESTINA RUBY SOLTERO

COMMITTEE MEMBERS

Dr. George P. Yanev
Chair of Committee

Dr. Hansapani Rodrigo
Committee Member

Dr. Tamer F. Oraby
Committee Member

Dr. Demba Fofana
Committee Member

August 2019

ABSTRACT

Soltero, Celestina R., Empirical Bayes Estimators and Borel-Tanner Distribution. Master of Science (MS), August, 2019, 32 pp., 3 tables, 4 figures, 31 references, 16 titles.

The motivation for this paper stems from the role Borel-Tanner (BT) distribution has as the distribution of the total outbreak number in epidemics modeled by branching processes. We briefly review Borel-Tanner distribution and its applications. In Chapter II we outline the Bayes decision problem, a construction for an Empirical Bayes (EB) estimator proposed by Liang [9] and discuss risk analysis. In Chapter III, the importance of randomization addressed and a classical construction of a monotonized EB estimator proposed by Houwalingen [14] is outlined. Lastly in Chapter IV we use R software to perform a Monte Carlo simulation and conduct a numerical study in which we construct data and estimators for the reproduction parameter of Borel-Tanner distribution. We implement a procedure oulined by Houwalingen [14] to obtain a monotonized version of the EB estimator poposed by Liang [9]. The estimators are assessed through risk analysis under squared error loss function and numerical study results are reviewed. The study suggests that the monotonized EB estimator outperforms the original EB estimator.

DEDICATION

To everyone who contributed to my academic success; my professors that made my education possible, as well as role models, colleagues, classmates and friends from whom I received emotional support and offered much to learn from throughout the years.

ACKNOWLEDGMENTS

I am a Science Technology and Engineering Partnership for Success (STEPS) Endowment and Nobel Laureate S. Chandraeskha Award recipient; throughout my graduate studies I was also a Graduate Teaching Assistant (GTA) for the Department of Mathematics and Statistical Sciences (SMSS) at UTRGV. I want to thank the corresponding foundations and the math department for partially funding my education and providing me with invaluable experience.

I am deeply grateful to my thesis director Dr. George Yanev for suggesting the area of interest and for his unconditional assistance and patience throughout the entire time we collaborated. I want to acknowledge Dr. Tamer Oraby for addressing my inquiries and providing insight each time I struggled with technicalities; I am also grateful to Dr. Hansapani Rodrigo and Dr. Demba Fofana for forming part of my thesis committee. I want to acknowledge Dr. Timothy Huber, whom I have known since I was an undergraduate, for encouraging me to further my education and always addressing each of my concerns with patience and reassurance. I am profoundly thankful to Dr. Hanan Amro for being an exceptional mentor in my life, always offering emotional as well as academic support and my choir teacher, Dr. Adrian Guerra who taught me from a young age to push my limits and persevere.

Lastly, I want to acknowledge my family. Malak, thank you for helping me in the little ways that you could. Maricruz, thank you for being so persistent towards the pursuit of my higher education even when I thought I was not ready and of course, for the dependable emotional support you have always offered. Gabriel, thank you for your time, the life lessons, your unconditional academic support and encouragement.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

1.1 Borel–Tanner Distribution

The Borel–Tanner distribution was originally derived as the distribution of the number of customers served in a busy period of a single server queuing process. It has probability mass function (pmf)

$$p_r(x \mid \theta) = c_r(x)\theta^{x-r}e^{-\theta x}\ , \tag{1.1}$$

where $0 < \theta < 1$, r is a positive integer, and $c_r(x) = \frac{rx^{x-r-1}}{(x-r)!}$ with mean $\frac{r}{1-\theta}$ and variance $\frac{r\theta}{(1-\theta)^3}$. It was introduced by French mathematician Emile Borel in 1942 for the case $r = 1$; thus, came to be known as the Borel distribution. In 1953, Tanner generalized the distribution [5] to any positive integer r. As a result, the former was then known as the Borel–Tanner (BT) distribution.

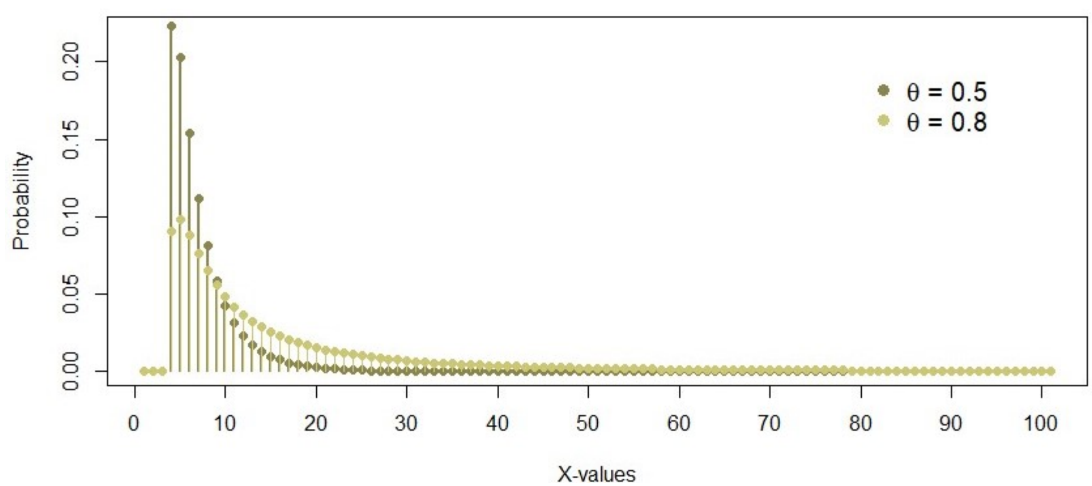

Figure 1.1: Borel–Tanner pmf with $r = 3$.

The BT distribution has surfaced in a variety of real-world phenomena. In queuing theory, (1.1) represents the probability that exactly x customers in a queue will be served before the first queue vanishes, beginning with r initial customers and traffic intensity θ, assuming Poisson arrivals and constant service time [5]. It has arisen in coalescence models [3], self propagating codes called worms which adversely impact the internet [12], herd size in finance modeling [11], cascading electrical outages [6] and highway traffic flows [8] in addressing the mean queue length behavior along a two-lane rural road where the presence of a queue in one lane prevents vehicles in the other lane from overtaking slower vehicles [4]. Our interest in BT distribution, however, stems from its role in modeling epidemics.

1.2 Branching Processes

This section is adapted from [7] and [2], unless otherwise stated. A system in which particles live for a random time and produce a random number of progenies is called a branching process. For an interesting historic overview on branching processes see [2]. Individuals from high social status were concerned about their noble family names ultimately becoming extinct, i.e., the number of a progeny (male individuals) may be zero. The oldest, simplest and best-known branching process is the Galton–Watson (GW) process also known as Bienaymé–Galton–Watson, since the oldest document found where the problem of extinction is considered from statistician Bienaymé dates back to 1845.

Branching processes are useful in many applications, e.g., describing higher organisms such as vertebrates or plants, biological cells, biomolecules and genes. In our study we apply it to epidemiology and consider the progeny to be the total number of infected individuals, i.e. the epidemic outbreak size. The GW branching process can be defined by the following recurrence formula

$$Z_{n+1} = \sum_{i=1}^{Z_n} X_{i,n}, \tag{1.2}$$

where $X_{i,n}$, $i,n = 1,2,\ldots$ are independent and identically distributed (iid) random variables (rv) that assume nonnegative integer values and $Z_0 = 1$. There are two basic assumptions (e.g. Yanev [15])

(i.) The number of offspring $X_{i,n}$ produced by a single parent particle is independent of the history of the process, and of other particles existing at the present.

(ii.) The offspring distribution is the same for all particles in all generations of the process.

The relationship between BT distribution and branching processes is in the event that the offspring distribution is $Poi(\theta)$, i.e., a GW process, then BT distribution gives the total number of individuals ever lived, that is, the total outbreak size.

CHAPTER II

BAYESIAN ESTIMATORS

The Bayesian estimation procedure can be summarized as follows. The prior distribution $G(\theta)$ is based on the belief of an observer and is formulated prior to seeing any actual data. It is a probability distribution which describes the variation of parameter θ. We have data x, taken from the population, indexed by θ. The sample has sampling distribution $p(x \mid \theta)$ which illustrates the observer's belief of where the data will be if θ is true. The prior is updated and is called the posterior distribution $G(\theta \mid x)$. This is done using and Bayes rule

$$G(\theta \mid x) = \frac{p(x \mid \theta)G(\theta)}{m(x)} \qquad \theta \in \Omega, \tag{2.1}$$

where $m(x)$ is the marginal distribution of X that is, $m(x) = \int_\Omega p(x,\theta)\mathrm{d}\theta$ and $p(x,\theta)$ is the joint probability mass function. The posterior distribution is now used to make inferences about θ.

2.1 Classical Bayes

A more detailed Bayes mathematical framework consists of the following elements (e.g. Stijnen [13]). An observation is taken from a random variable or vector X, the distribution of which depends on an unknown parameter θ. The problem is what decision to take concerning the true value of θ.

(i) A set S of observations, called sample space, equipped with a $\sigma-$algebra $\mathcal{S}$.

(ii) A collection $\mathcal{P}$ of probability measures on the space $(S,\mathcal{S})$. Usually, $\mathcal{P}$ is parametrized by some set suitable parameters $\mathcal{P} = \{P_\theta, \theta \in \Omega\}$.

(iii) A set A of possible actions which can be taken by the statistician upon observing some $x \in S$.

The set A, called the action space, is equipped with a $\sigma-$algebra $\mathcal{A}$.

(iv) A collection D of decision rules. A decision rule is defined to be a $\mathcal{S}-\mathcal{A}$ measurable map from S into A. A decision rule is defined to be a $\mathcal{S}-\mathcal{A}$ measurable map from S into A When using the decision rule $d \in D$, the statistician will take action $d(x) \in A$ upon observing $x \in S$.

(v) A loss function $L : \Omega \times A \longrightarrow \mathbb{R}$. For each $\theta \in \Omega$, the function $L(\theta, \cdot)$ must be $\mathcal{A}$ measurable and bounded from below on A. When taking $d(x) \in A$, if θ is the true parameter value, the statistician will incur a loss function $L(\theta, d(x))$.

(vi) A probability measure G (called the prior distribution) on Ω, which is equipped with the $\sigma-$ algebra $\mathcal{W}$.

Adopting the Bayesian framework, we define the Bayes estimator θ_G. Suppose $\theta \in \Omega$ is a realization of a rv Θ. Under the squared error loss function, with prior distribution G and Borel-Tanner pmf (1.1), it is well known that the Bayesian estimator θ_G for θ is the posterior mean

$$\theta_G(x) := E\left[\Theta \mid X = x\right] = \frac{\displaystyle\int_\Omega \theta^{x+1-r} e^{-x\theta} \mathrm{d}G(\theta)}{\displaystyle\int_\Omega \theta^{x-r} e^{-x\theta} \mathrm{d}G(\theta)}. \tag{2.2}$$

Consider a population parameter Θ that has significant physical interpretation, e.g., the reproduction number of a current outbreak modeled by a GW process. When sampling from a population whose distribution is given by $p(x \mid \theta)$, knowledge of θ provides knowledge over the entire population. In order to make reliable inferences about the population, it is of utmost importance to construct a quality estimator $\hat{\theta}$ for θ so that we can take measures in controlling the outbreak if necessary. If $\hat{\theta}$ is small enough, intervention might not even be necessary since, in that case, the epidemic will die out without affecting a significant population. On the other hand, if $\hat{\theta}$ is large enough, prevention methods might be needed to control the spread.

2.2 Measures for Estimators' Quality

Our aim is to obtain a good approximation for θ. A loss function L is used as a measure of discrepancies between a constructed estimator $\hat{\theta}$ and the true value of the parameter θ. If θ is real-valued parameter, a commonly used loss function is the squared error loss

$$L\big(\theta,\hat{\theta}\big) := \big(\hat{\theta}(x)-\theta\big)^2. \tag{2.3}$$

The quality of an estimator $\hat{\theta}$ is assessed by its risk function. At a point θ, the risk function is the expected loss that will be incurred if the estimator $\hat{\theta}$ is used. If the prior distribution is known, it is often possible to determine a decision rule with minimum Bayes risk.

Definition 1 (Bayes Risk)**.** Under loss function (2.3), the Bayes risk $r(G,\hat{\theta})$ of estimator $\hat{\theta}$ with respect to the prior distribution G is

$$\begin{aligned} r(G,\hat{\theta}) &= E_{(X,\Theta)}L\big(\theta,\hat{\theta}(x)\big) \\ &= E_{(X,\Theta)}\big(\hat{\theta}(x)-\theta\big)^2. \end{aligned} \tag{2.4}$$

By definition, the Bayesian estimator θ_G minimizes the Bayes risk (2.4) i.e.,

$$r(G,\theta_G) = min\; r(G,\hat{\theta}).$$

Definition 2 (Regret Risk)**.** The difference $R(\hat{\theta})$ between the Bayes risk and the minimum Bayes risk of any estimator $\hat{\theta}$ is called Regret Risk of $\hat{\theta}$,

$$R(\hat{\theta}) = r(G,\hat{\theta}) - r(G,\theta_G). \tag{2.5}$$

Remark. Since the Bayes risk is minimum when using estimator θ_G, the regret risk $R(\hat{\theta})$ of any estimator $\hat{\theta}$, is always greater then or equal to zero.

It is not difficult to see that $R(\hat{\theta}) = E_X\left[\hat{\theta}(x) - \theta_G(x)\right]^2$. Indeed,

$$
\begin{aligned}
R(\hat{\theta}) &= r(G,\hat{\theta}) - r(G,\theta_G) \\
&= E_{(x,\Theta)}\left[\hat{\theta}^2(x) - 2\hat{\theta}(x)\Theta + \Theta^2\right] - E_{(x,\Theta)}\left[\theta_G^2(x) - 2\theta_G(x)\Theta + \Theta^2\right] \\
&= E_X\Big[E_{(\Theta,X)}\left[\hat{\theta}^2(x) - 2\hat{\theta}(x)\Theta + 2\theta_G(x)\Theta - \theta_G^2(x)\right]\Big] \\
&= E_X\left[\hat{\theta}^2(x) - 2\hat{\theta}(x)\theta_G(x) + 2\theta_G^2(x) - \theta_G^2(x)\right] \\
&= E_X\left[\hat{\theta}^2(x) - 2\hat{\theta}(x)\theta_G(x) + \theta_G^2(x)\right] \\
&= E_X\left[\hat{\theta}(x) - \theta_G(x)\right]^2.
\end{aligned}
$$

In Bayesian theory, the Bayes estimator θ_G is considered the golden standard. Thus, we define the notion of "best" estimator $\hat{\theta}$ by that which is "closest" to θ_G, i.e., with minimum regret risk. When comparing estimators $\hat{\theta}_1$ and $\hat{\theta}_2$, if $R(\hat{\theta}_2) < R(\hat{\theta}_1)$, we consider $\hat{\theta}_2$ to be "better" than $\hat{\theta}_1$.

2.3 Empirical Bayes

It is often reasonable to assume that a prior distribution G exists, however is unknown. An Empirical Bayes (EB) approach is taken when we have "past" data parametrized by Θ which is usually unobservable to us and has a prior distribution G. This approach does not assume any specific prior, it simply restricts itself to this past data. In what follows, we adopt the Empirical Bayes approach, which relies on the assumption for existence of a prior G which, however, is unknown. In this setting, our investigation is that of one event in a sequence of similar independent events with same prior distribution G. The data of these preceding events can be used to estimate the prior G or the Bayes rule θ_G directly.

The parameter Θ can reasonably be considered a random variable with some prior distribution G. With the following scenario, Maritz [10] illustrates a situation in which the Empirical Bayes assumptions are fulfilled. Suppose prospective college students arrive sequentially and are subject to a college entrance exam. Based on their test score, a decision will be made about their admittance. It is reasonable to assume that each student has a predetermined potential, θ, which cannot be measured directly. However, the student's exam score X is a normal r.v. with mean θ and some known variance which is fixed for all students. A collection scores on a well designed exam can provide insight into the prior, G.

More precisely, in the empirical Bayes setup,consists of the following

(i) A sequence of independent and identically distributed (iid) copies

$$(X_1,\Theta_1),(X_2,\Theta_2),\ldots,(X_n,\Theta_n),\ldots$$

of the random pair (X,Θ) where Θ has a distribution G, and conditional on Θ, X has the Borel-Tanner distribution (1.1).

(ii) Assume X_i, $i=1,2,\ldots,n+1$ are observable and parametrized by Θ_i, $i=1,2,\ldots,n+1$.

(iii) Each Θ_i is unobservable and has unknown prior distribution G.

(iv) Let X_{n+1} stand for the present observation and $\underline{X}(n) := (X_1,\ldots,X_n)$ denote the n past observations.

The past data $\underline{X}$ can be used to gather information about the prior G. An EB estimator θ_n of the present parameter θ_{n+1} is a function of the currently observed value $X_{n+1}=x$ and the past data $\underline{X}$.

In case of Borel-Tanner, under squared error loss, Liang [9] successfully constructed an EB estimator θ_n for the Bayes estimator θ_G and studied its properties. The next definition is adapted from Liang [9].

Definition 3. For each positive integer $x=r,r+1,\ldots$, let

$$q_n(x) := \frac{1}{n}\sum_{j=1}^{n}\frac{I\{X_j=x\}}{c_r(x)} \quad \text{and} \quad \psi_n(x) := \frac{1}{n}\sum_{j=1}^{n}\frac{c_1(X_j-x)I\{X_j\geq x+1\}}{c_r(X_j)}. \tag{2.6}$$

With $q_n(x)\neq 0$, for each $x=r,r+1,\ldots$, the EB estimator θ_n is defined by

$$\theta_n(x) := \min\left\{\frac{\psi_n(x)}{q_n(x)},1\right\}. \tag{2.7}$$

The Bayes risk of the EB estimator $\theta_n(X)$ is

$$r(G,\theta_n) := E_n E_{(X_{n+1},\Theta_{n+1})}[\Theta_{n+1} - \theta_n(X_{n+1})]^2.$$

where E_X is the expectation with respect to $(X_1, X_2, \ldots, X_n)$. Using (2.5) for the regret risk of θ_n, we have

$$R(\theta_n) := r(G,\theta_n) - r(G,\theta_G).$$

In particular, θ_n is called asymptotically optimal for any prior G if $\lim_{n\to\infty} R(\theta_n) = 0$. In [9] Liang proves that θ_n given by (2.7) is asymptotically optimal and studies the $R(\theta_n)$ rate of convergence to zero.

CHAPTER III

MONOTONIZING THE EMPIRICAL BAYES ESTIMATOR

3.1 Randomization

Randomization reduces bias as much as possible; it is designed to "control" bias by all means. For a very basic and intuitive introduction to randomization see [1]. When a study is randomized it reduces or eliminates bias; thereby providing more reliable results and legitimacy to both the research and researchers as well.

Example 1 (Randomized Test)**.** Let X_1, X_2, X_3 be a sample from $Bin(1,\theta)$ where $0 \le \theta \le 1$ and θ is unknown. Let x be the number of successes in 3 independent trials. Consider $H_0 : \theta = \frac{1}{4}$ vs. $H_1 : \theta = \frac{3}{4}$ and let $\alpha = 0.05$. Then the probabilities are in Table 3.1. Clearly $P(X = 3)$ fully falls in

Table 3.1: Probability values for Randomization Example.

x	$P(X = x)$
0	$\frac{3!}{3!0!}\left(\frac{1}{4}\right)^0\left(\frac{3}{4}\right)^3 = \frac{27}{64} \approx 0.42$
1	$\frac{3!}{2!1!}\left(\frac{1}{4}\right)^1\left(\frac{3}{4}\right)^2 = \frac{27}{64} \approx 0.42$
2	$\frac{3!}{1!2!}\left(\frac{1}{4}\right)^2\left(\frac{3}{4}\right)^1 = \frac{9}{64} \approx 0.14$
3	$\frac{3!}{0!3!}\left(\frac{1}{4}\right)^3\left(\frac{3}{4}\right)^0 = \frac{1}{64} \approx 0.02$

the rejection region and $P(X = 2)$ does not. The problem here is that we are not using $\alpha = 0.05$ as our exact critical value; thus creating bias for the decision process. A way to fix this is to randomize the test. In this case we will add a weight c to $X = 2$, that is, we partially include the point $X = 2$ so that we obtain the exact critical value $\alpha = 0.05$,

$$\begin{aligned}
P(X=3)+cP(X=2) &= 0.05 \\
\frac{1}{64}+c\frac{9}{64} &= 0.05 \\
c &= \frac{0.05(64)-1}{9} \\
c &= \frac{2.2}{9}.
\end{aligned}$$

Thus, the optimal test of size $\alpha = 0.05$ is given by

$$\Phi_\theta(x) := \begin{cases} 0 & \text{if } x < 2 \\ \frac{2.2}{9} & \text{if } x = 2 \\ 1 & \text{if } x = 3. \end{cases}$$

Randomization assigns values by chance not by choice. In the above example we used a weight c to obtain the exact α value and eliminate bias. Randomization is a useful tool to reduce or completely eliminate bias from any experiment. There are several ways to randomize an experiment, in the section that follows we use a function, namely $D(a \mid x)$, to randomize the EB estimator θ_n.

3.2 A Monotonization Procedure

The EB estimator is not monotone with respect to x. We provide an illustration of θ_n in the Chapter IV numerical study. This is unwanted behavior for an estimator following BT distribution.

Proposition 1. *The BT distribution, (1.1) has monotone likelihood ratio (MLR), i.e.,*

$$q(x) = \frac{p_r(x \mid \theta_2)}{p_r(x \mid \theta_1)}$$

is increasing with respect to x whenever $0 < \theta_1 < \theta_2 < 1$.

Proof. Let g be the natural logarithm of the likelihood ratio q, $0 < \theta_1 < \theta_2 < 1$ and r a positive integer, i.e.,

$$
\begin{aligned}
g(x) &= \ln q(x) \\
&= \ln \frac{p_r(x \mid \theta_2)}{p_r(x \mid \theta_1)} \\
&= \ln \frac{\theta_2^{x-r} e^{-\theta_2 x}}{\theta_1^{x-r} e^{-\theta_1 x}} \\
&= \ln \left(\frac{\theta_2}{\theta_1}\right)^{x-r} + \ln e^{-x(\theta_2-\theta_1)} \\
&= (x-r)\ln\left(\frac{\theta_2}{\theta_1}\right) - x(\theta_2-\theta_1).
\end{aligned}
$$

Its derivative $g'$with respect to x is

$$
\begin{aligned}
g'(x) &= \ln\left(\frac{\theta_2}{\theta_1}\right) - (\theta_2-\theta_1) \\
&= \ln(\theta_2) - \theta_2 - (\ln(\theta_1) - \theta_1) \\
&= \ln(\theta_2 e^{-\theta_2}) - (\ln \theta_1 e^{-\theta_1}).
\end{aligned}
$$

Consider the function $h(\theta) = \ln(\theta e^{-\theta})$. Its derivative h' with respect to θ

$$
\begin{aligned}
h'(\theta) &= \left(\frac{1}{\theta e^{-\theta}}\right)(\theta e^{-\theta})' \\
&= \left(\frac{1}{\theta e^{-\theta}}\right)(e^{-\theta} - \theta e^{-\theta}) \\
&= \frac{1-\theta}{\theta}
\end{aligned}
$$

is always greater than zero for any $0 < \theta < 1$. Since

$$
g'(x) = h(\theta_2) - h(\theta_1) > 0
$$

whenever $0 < \theta_1 < \theta_2 < 1$, the function $g(x) = \ln q(x)$ is monotone increasing. Thus, $q(x)$ itself is monotone increasing. □

Due to this property of BT distribution, monotonicity is a desirable property for θ_n. However as Houwalingen [14] points out, this is not the case for the EB estimator; for this reason, he outlined a classical approach for monotonizing the EB estimator. In addition to monotonizing the θ_n^*, Houwalingen also shows that the monotonized EB estimator, θ_n^* has a smaller Regret risk than the EB estimator θ_n, i.e., θ_n^* is a "better" estimator than θ_n. A procedure for constructing a monotone estimator that dominates an EB estimator for distributions with MLR is given. In his paper, Houwalingen also provides examples of this estimator for the Geometric and Poisson distributions. In Chapter IV, we contribute yet another example to this classical construction by monotonizing the EB estimator for BT distribution.

Estimators for discrete distributions with MLR can be made monotone applying a procedure developed in [14] (see also [16]). Consider a simple randomized version of the estimator $\theta_n(x)$ represented by the following function $D(a \mid x)$ for $a \in (0,1)$:

$$D(a \mid x) := \begin{cases} 0 & \text{if } \theta_n(x) > a, \\ 1 & \text{if } \theta_n(x) \leq a. \end{cases}$$

The number $D(a \mid x)$ is the probability that an estimate $\theta_n(x)$ less than or equal to a is selected given $X = x$. Hence $D(a \mid x)$ is a cdf on the action space $(0,1)$ for every $X = x$. Define for $a \in (0,1)$

$$\alpha(a) := E(D(a \mid X)) = \sum_{\{x:\ \theta_n(x) \leq a\}} p_r(x \mid a). \tag{3.1}$$

Denote $F(x \mid \theta) := \sum_{k=r}^{x} p_r(k \mid \theta)$ for $x \geq r$ and $F(r-1 \mid \theta) = 0$. Now, we can construct a randomized estimator with $D^*(a \mid x)$ as follows

$$D^*(a \mid x) := \begin{cases} 0 & \text{if } \alpha(a) < F(x-1 \mid a) \\ \dfrac{\alpha(a) - F(x-1 \mid a)}{F(x \mid a) - F(x-1 \mid a)} & \text{if } F(x-1 \mid a) \leq \alpha(a) \leq F(x \mid a) \\ 1 & \text{if } F(x \mid a) < \alpha(a), \end{cases} \tag{3.2}$$

$D^*(1 \mid x) = 1$, and $D^*(0 \mid x) = \lim_{a \downarrow 0} D^*(a \mid x)$. Let $a \in (\theta_0, \theta_1)$ be fixed. It follows from the construction of D^*, that $E_a D^*(a \mid X) = E_a D(a \mid X)$.

The next proposition shows that, using the monotone estimator D^*, one can construct another (non-random) monotone estimator θ_n^*, say, with risk less than or equal to the risk of the θ_n.

Proposition 2. *Let $D^*(a \mid x)$ be the monotone estimator constructed in (3.2).Define*

$$\theta_n^*(x) := \int_0^1 a \, \mathrm{d}D^*(a \mid x). \tag{3.3}$$

Then the monotone non-random estimator $\theta_n^*(x)$ dominates $D^*(a \mid x)$, which itself dominates the initial estimator $D(a \mid x)$, i.e.,

$$R(\theta, \theta_n^*) \leq R(\theta, D^*) \leq R(\theta, D). \tag{3.4}$$

Proof. The proposition follows from the theorem in [14]. It suffices to verify that BT distribution satisfies all assumptions of the theorem. In particular, it has a MLR as it was shown in Proposition 1. Therefore, the second inequality in (3.4) follows as in [14]. That is, D^* represents a monotone estimator which dominates the initial estimator represented by D for all $\theta \in (0,1)$. It is not difficult to see that, under the squared error loss function, D^* itself is dominated by the non-random monotone estimator θ_n^*. Indeed, using Jensen's inequality, we have

$$\begin{aligned}
R(\theta, \theta_n^*(X)) &= E(\theta - \theta_n^*(X))^2 \\
&= E\left(\theta - \int_0^1 a \, dD^*(a \mid X)\right)^2 \\
&\leq E\left(\int_0^1 (\theta - a)^2 \, dD^*(a \mid X)\right) \\
&= R(\theta, D^*(a \mid X)).
\end{aligned}$$

□

CHAPTER IV

MONTE CARLO SIMULATION

It is our interest to construct quality estimators for Θ because this will allow us to take measures addressing an epidemic when necessary. In this chapter we present the results obtained from a Monte Carlo experiment using R software and interpret them as an epidemic size observation. Algorithms for the simulations are provided in this chapter and the code constructed in R software is given in its entirety in Appendix A. For simulation purposes, we use the following setting.

4.1 Numerical Study

Let X be a discrete random variable following BT distribution with a $Uni(0.5, 0.8)$ prior G for θ and let $r = 3$. Then, using (2.2), the Bayes estimator θ_G is given by

$$\theta_G(x) = \frac{\int_{0.5}^{0.8} (\theta^{x+1-3} e^{-x\theta}) \mathrm{d}\theta}{\int_{0.5}^{0.8} (\theta^{x-3} e^{-x\theta}) \mathrm{d}\theta}. \tag{4.1}$$

Also, calculating the maximum likelihood (ML) estimator θ_{ML} for the BT parameter θ we have

$$\begin{aligned} \ln p(x \mid \theta) &= \ln\left(c_r(x)\theta^{x-r} e^{-\theta x}\right) \\ &= \ln c_r(x) + (x-r)\ln\theta - \theta x. \end{aligned}$$

Taking the derivative with respect to θ,

$$\frac{\partial}{\partial\theta} \ln p(x|\theta) = \frac{x-r}{\theta} - x.$$

Setting it equal to zero we have,

$$\frac{x-r}{\theta} - x = 0 \Longrightarrow \frac{x-r}{\theta} = x$$
$$\Longrightarrow \frac{x-r}{x} = \hat{\theta}$$

Thus the ML estimator for θ is given by

$$\theta_{ML}(x) = \frac{x-r}{x}. \tag{4.2}$$

ALGORITHM 1: Bayes Estimator θ_n, ML Estimator θ_{ML}, and corresponding Risks

/* We replace ∞ with 100 to obtain a numerical approximation. */

Generate X=r, r+1, ..., 100 /* Vector X of current outbreak size */

for *x in r:100* **do**

    Compute $c_r(x) = \frac{rx^{x-r-1}}{(x-r)!}$ /* BT $c_r(x)$−coefficient */

    Compute $\theta_G(x) = \frac{\int_a^b \theta^{x+1-r} e^{-x\theta} d\theta}{\int_a^b \theta^{x-r} e^{-x\theta} d\theta}$ /* Bayes estimator θ_G */

    Compute $\theta_{MLE}(x) = \frac{x-r}{x}$ /* Maximum Likelihood estimator θ_{ML} */

end

Calculate minimum Bayes risk $r(G, \theta_G)$ /* Using (2.4) */

Calculate Bayes risk $r(G, \theta_{ML})$ /* Using (2.4) */

Calculate Regret risk $R(\theta_{ML})$ /* Using (2.5) */

Using R and the framework from Algorithm 1 to compute the Bayes risk of (4.1) and (4.2) correspondingly we obtain,

$$r(G, \theta_G) = \frac{1}{0.3} \sum_{x=3}^{\infty} c_r(x) \int_{0.5}^{0.8} (\theta_G(x) - \theta)^2 \theta^{x-3} e^{-x\theta} d\theta \approx 0.0069$$

and

$$r(G, \theta_{ML}) = \frac{1}{0.3} \sum_{x=3}^{\infty} c_r(x) \int_{0.5}^{0.8} (\theta_{ML}(x) - \theta)^2 \theta^{x-3} e^{-x\theta} d\theta \approx 0.1003.$$

Thus, by (2.5), for the ML estimator θ_{ML}, the regret risk $R(\theta_{ML}) \approx 0.0935$.

Now consider a sequence of past epidemics for which we have documented the epidemic size but the reproduction number of each instance remains unknown i.e., the EB setting. We simulate the data by following the framework in Algorithm 2. Considering that the current outbreak size is BT, we will take $x_{max} = 20$ as the maximum current outbreak size. Otherwise the epidemic is underway of becoming a pandemic in which case the model is no longer fit. As will be demonstrated in Table 4.1, the models better fit the data as more past epidemics feed into it, i.e., the estimators' risk decreases i.e., as n increases. However the models still provide valuable insight on the reproduction parameter especially when few past epidemics have been observed. We considered $n = 20, 40, 60, 80, 100$ number past observations per data set and $m = 10$ data sets at a time. Also, in order to simulate the data, we use $Uni(a = 0.5, b = 0.8)$ as prior G so that each randomly generated parameter value θ_i generates a corresponding past epidemic outbreak size X_i. The parameter values θ_i that the r.v. Θ_i assumes remain irrelevant since in actuality these remain unobserved. For the EB estimator θ_n and the monotonized EB estimator θ_n^* we only work with the the epidemic size X_i generated in the data simulation.

ALGORITHM 2: Data Simulation

for *j in 1:m* **do**
  for *i in 1:n* **do**
    Draw random θ from prior $G-$prior
    Generate $\Theta_{n\times m}$ /* Matrix of parameter values $\theta_i^{(j)}$ */
    Generate $\underline{X}_{n\times m}$ /* Matrix of past data $\underline{X} = (X_i^{(j)})$ parametrized by $\Theta_{n\times m}$ using BT pmf */
    Compute $c_r(X_i^{(j)}) = \frac{rX_i^{(j)X_i^{(j)}-r-1}}{(X_i^{(j)}-r)!}$ /* BT $c_r(x)-$coefficient for $\underline{X}$ */
  end
end

Algorithm 3 shows a construction for θ_n following [9]. The EB estimator (2.7) is a ratio of the functions (2.6) and is bounded from above by 1. In terms of epidemics, $q_n(x)$ is a weighted average of the instances a past epidemic size was identical to the current total outbreak size x, while the function $\psi_n(x)$ is a weighted average of the instances in which a past epidemic size was greater than x. The EB estimator, however, exhibited jumpiness behavior in all trial runs (see Fig. 4.1). As previously stated, due to the MLR property of BT distribution, monotonicity of the parameter estimator is desired.

ALGORITHM 3: Empirical Bayes Estimator θ_n

x=r
while x<=$xmax$ **do**
  for j *in 1:m* **do**
    for i *in 1:n* **do**
      Call indicator function $I_i^{(j)}(x) = \mathbb{1}(X_i^{(j)} = x)$ `/* Matrix indicating instances when */` `/* past outbreak size matched current outbreak size */`
      if $(X_i^{(j)} - x) > 0$ **then**
        Calculate $c_1(X_i^{(j)} - x)$
      else
        Set $c_1(X_i^{(j)} - x) = 0$ `/* Compute BT coefficient` $c_1(X_i^{(j)} - x)$; $r = 1$ `*/`
      end
    end
  end
  Compute ratio $\frac{c_1(X_i^{(j)} - x)}{c_r(X_i^{(j)})}$ `/* Used later to define` ψ_n`; it is component- */` `/* -wise subtraction creating an` $n \times m$ `matrix */`
  j=1
  while j<=m **do**
    Compute $q_n^{(j)}(x) = \frac{1}{n}\sum_{i=1}^{n} \frac{I_i^{(j)}(x)}{c_r(x)}$ `/* Vector of` $q_n(x)$ `values for` j^{th} `data set */`
    Compute $\psi_n^{(j)}(x) = \frac{1}{n}\sum_{i=1}^{n} \frac{c_1(X_i^{(j)} - x)}{c_r(X_i^{(j)})}$ `/* Vector of` $\psi_n(x)$ `values for` j^{th} `data set; */`
    Compute $\theta_n^{(j)}(x) = \min\left\{\frac{\psi_n^{(j)}(x)}{q_n^{(j)}(x)}, 1\right\}$ `/* EB estimator` $\theta_n(x)$ `for` j^{th} `data set; */`
    j=j+1 `/* Update of data set` j `*/`
  end
  x=x+1 `/* Update of current outbreak size` x `*/`
end

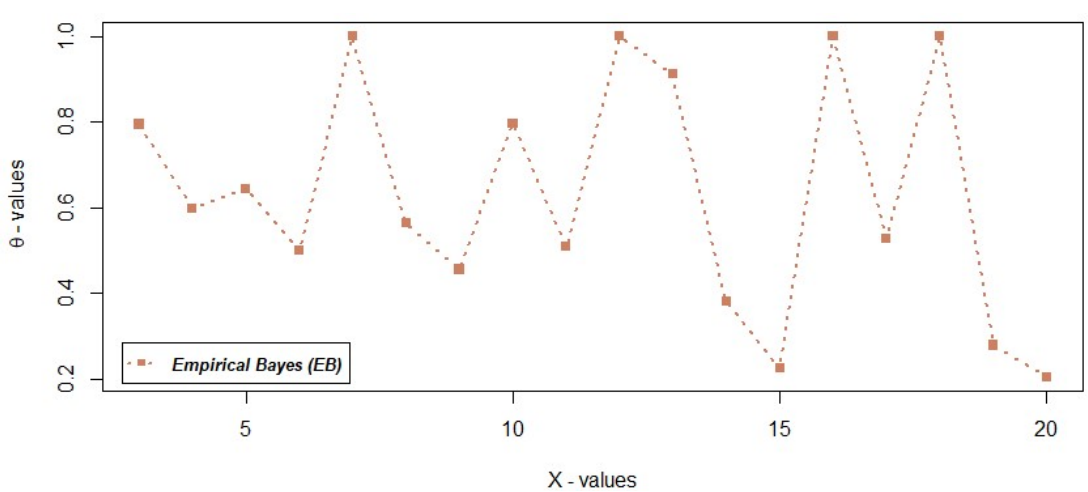

Figure 4.1: Empirical Bayes estimator for one simulation with $n = 60$.

We monotonized the EB estimator according to [14] (see Algorithm 4). The interval $(0,1)$ was partitioned into a grid of $na = 100$ equally spaced sub-intervals. The value a_i represents a point within $i^{th}-$partition of the interval and is used to construct a randomized estimator $D(a \mid x)$ for θ_n^*. We then use D for the construction of α, see (3.1). Next, we create a cdf $F(x \mid \theta)$ for the BT distribution and use this to construct a cdf $D^*(a \mid x)$, see (3.2). Lastly we construct a non-randomized monotone estimator θ_n^*, see (3.3) for θ.

The estimators θ_n, θ_n^* and θ_{ML} are assessed through their regret risks (see Appendix A). For each of the estimators, θ_n and θ_n^*, 100 simulations were generated. We average the regret risk for the 100 data sets of the EB estimator θ_n;

$$\overline{R}(\theta_n) = \frac{1}{100} \sum_{k=1}^{100} R(\theta_n^{(j)}),$$

where $j = 1, 2, \ldots, 100$. Similarly, we average the regret risk for the monotonized EB estimator θ_n^* $\overline{R}(\theta_n^*)$. The numerical results are reported in Table 4.1.

4.2 Concluding Remarks

In this paper we studied the estimation problem for the reproduction parameter θ of the BT distribution. A good quality of this model is its simplicity; the only information needed is the total outbreak size of past similar epidemics; with this data, under GW assumptions, we can produce estimators for the disease reproduction number θ and address the the current epidemic outbreak if necessary.

Fig. 4.2 shows one example trial comparison between θ_n and θ_n^*. The behavior is accordingly to that of an estimator whose distribution has MLR property. In Fig. 4.3, we display one trial run of all three Bayesian estimators. For further study, perhaps we could focus more attention to the seeing

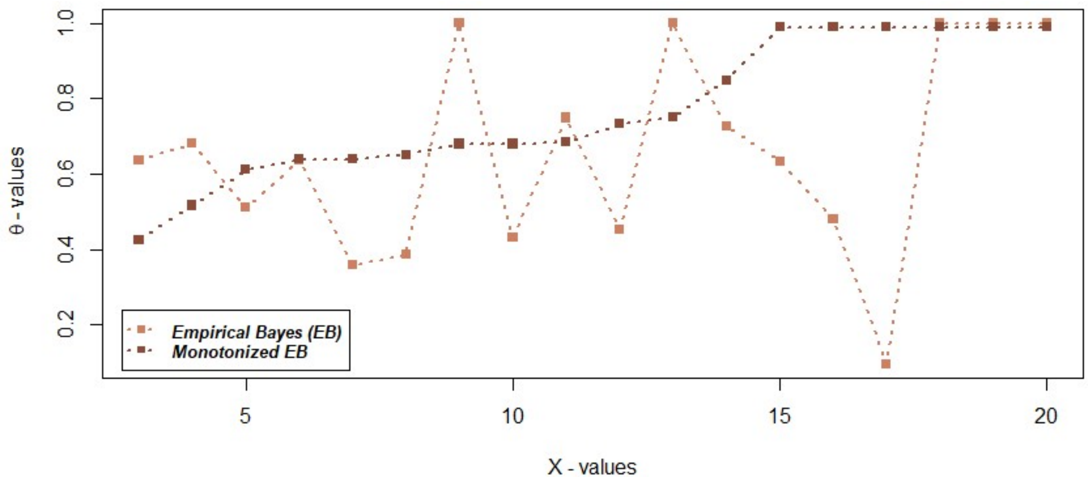

Figure 4.2: EB and Monotonized EB comparison for one simulation with $n = 40$.

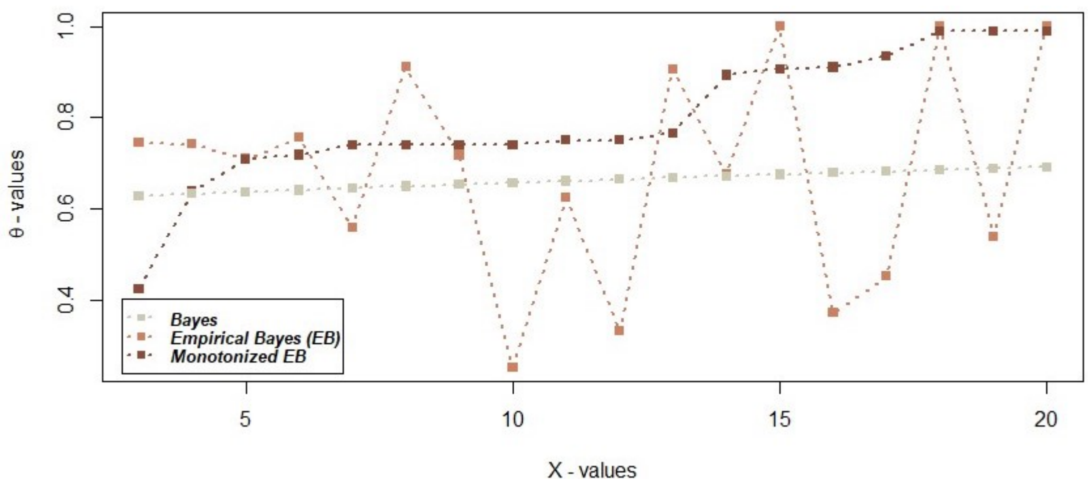

Figure 4.3: Bayesian estimates based on one simulation for $n = 100$.

if we can find an interval containing ideal $X-$values for which the model is best. For example, in Fig. 4.3 we see that for $x = 4$ to about $x = 14$, our estimator θ_n^* is closest to the Bayes estimator θ_G; thus risk is minimal throughout these points.

We constructed the ML estimator θ_{ML}, an EB estimator θ_n and a monotonized version θ_n^* for θ_n. The results demonstrate that not only does the monotonized EB estimator θ_n^* behave as desired, regardless of the number of past observed epidemics, but it also is a better estimator than the original EB estimator θ_n since the risk associated is smaller than that of θ_n and θ_{ML}.

ALGORITHM 4: Monotonized EB Estimator θ_n^*

```
for j in 1:m do
    for i in 1:na do
        for x in 1:xmax do
            if θ_n^(j)(x) < a_i then
                α^(j)(a_i) = α^(j)(a_i) + Σ_{i=1}^{na} p_r(x | a_i)   /* Construct D and calculate α from (3.1) */
            end
        end
    end
end
Initiate F_{x_max × na}(x | a_i) as zero matrix   /* Construct BT cdf */
for i in 1:na do
    F(r | a_i) = p_r(r | a_i)
    for x in r+1:xmax do
        F(x | a_i) = F(x − 1 | a_i) + p_r(x | a_i))
    end
end
j=1   /* Construct D* from (3.2) */
while j<=m do
    for i in 1:na do
        if α^(j)(a_i) > F(r | a_i)) then
                                                   /* case: x = r */
            D^{*(j)}(a_i | r) = 1
        else
            D^{*(j)}(a_i | r) = α^(j)(a_i) / F(r | a_i)
        end
                                                   /* case: x > r */
        for x in r+1:xmax do
            if F(x − 1 | a_i) > α^(j)(a_i) then
                D^{*(j)}(a_i | x) = 0
            else
                if F(x | a_i) < α^(j)(a_i) then
                    D^{*(j)}(a_i | x) = 1
                else
                    D^{*(j)}(a_i | x) = (α^(j)(a_i) − F(x−1 | a_i)) / (F(x | a_i) − F(x−1 | a_i))
                end
            end
        end
    end
    x=r   /* Construct θ_n^* from (3.3) */
    while x<=xmax do
        for i in 1:na do
            tail_i(x) = 1 − D^{*(j)}(a_i | x)
            θ_n^{*(j)}(x) = (1/na) Σ_{i=1}^{na} tail(x)
        end
        x=x+1   /* Update of current outbreak size x */
    end
    j=j+1   /* Update of data set j */
end
```

Table 4.1: Estimates for the regret risks of θ_n, θ_n^* and θ_{ML}

n	$R(\theta_{ML})$	$\overline{R}(\theta_n)$	$\overline{R}(\theta_n^*)$
20	0.0935	0.0969	0.0557
40	0.0935	0.0746	0.0402
60	0.0935	0.0632	0.0330
80	0.0935	0.0581	0.0311
100	0.0935	0.0500	0.0270

All standard errors are less than 10^{-4} and $r = 3$.

BIBLIOGRAPHY

[1] *Randomization*. `https://explorable.com/randomization`, 8 2010. Accessed: Aug. 5 2019.

[2] K. Albertsen, J. Steffensen, and E. Kirstensen, *THREE PAPERS ON THE HISTORY OF BRANCHING PROCESSES*, Tech. Rep. 242, Department of Statistics, GN-22 University of Washington, Seattle, Washington, 11 1992. Translated from Danish by Peter Guttorp.

[3] D. J. Aldous, *Deterministic and stochastic models for coalescence (aggregation and coagulation): a review of the mean-field theory for probabilists*, Bernoulli, 5 (1999), pp. 3–48.

[4] P. G. Gipps, *An Abbreviated Procedure for Estimating Equilibrium Queue Lengths in Rural Two Lane Traffic.*, Transportation Science, 10 (1976), p. 337.

[5] F. A. Haight and M. A. Breuer, *The Borel-Tanner distribution*, Biometrika, 47 no. 1/2 (1960), p. 143.

[6] T. Iešmantas and R. Alzbutas, *Bayesian assessment of electrical power transmission grid outage risk*, International Journal of Electrical Power & Energy Systems, 58 (2014), pp. 85 – 90.

[7] M. Kimmel and D. Axelrod, *Branching Processes in Biology*, Springer Publishing Company, Incorporated, 2nd ed., 2015.

[8] G. Koorey, *Passing Opportunities at Slow-Vehicle Bays*, Journal of Transportation Engineering, 133 (2007), pp. 129–137.

[9] T. Liang, *Empirical Bayes estimation for Borel–Tanner distributions*, Statistics & Probability Letters, 79 (2009), pp. 2212 – 2219.

[10] J. S. Maritz, *Empirical Bayes Methods*, vol. 4 of 10, Methuen's monographs on applied probability and statistics. Methuen and Co. Ltd.,, London, 1970.

[11] M. Nirei, T. Stamatiou, and V. Sushko, *Stochastic Herding in Financial Markets Evidence from Institutional Investor Equity Portfolios*, Tech. Rep. 371, Bank for International Settlements, 2 2012.

[12] S. H. Sellke, N. B. Shroff, and S. Bagchi, *Modeling and Automated Containment of Worms*, IEEE Transactions on Dependable and Secure Computing, 5 (2008), pp. 71–86.

[13] T. Stijnen, *On the asymptotic behaviour of monotonized empirical Bayes rules*, PhD thesis, University of Utrecht, The Netherlands, 1980.

[14] J. C. VAN HOUWELINGEN, *Monotonizing empirical Bayes estimators for a class of discrete distributions with monotone likelihood ratio*, Statistica Neerlandica, 31 (1977), pp. 95–104.

[15] G. P. YANEV, *Statistical Modeling of Epidemic Disease Propagation via Branching Processes and Bayesian Inference*, PhD thesis, University of South Florida, Tampa, Florida, 2001.

[16] G. P. YANEV AND R. COLSON, *Monotone Empirical Bayes Estimators for the Reproduction Number in Borel-Tanner Distribution*, Pliska Studia Mathematica, 27 (2017), pp. 115–122.

APPENDIX A

APPENDIX A

```r
library(VGAM)
```

```
## Loading required package: stats4

## Loading required package: splines
```

```r
#---A FEW PREDEFINED ITEMS-------------------------------------------------------------------
options(max.print = 10000)
a=.5                          #Lower bound on G-prior
b=.8                          #Upper bound on G-prior
r=3                           #Initial outbreak size (OB)
m=10                          #no. DataSets
n=100                         #no. Past Observations/set
xmax=21                       #max no. of current OB size
kmax=xmax-r                   #no. distinct CurrObservations
ra=mat.or.vec(n,m)            #rand. theta values from G-prior
dg=mat.or.vec(kmax,1)         #Bayes estimator for X
dmle=mat.or.vec(kmax,1)       #ML estimator for X
Xpast=mat.or.vec(n,m)         #past observed total OB size (tOBs)
cXpast=mat.or.vec(n,m)        #Borel-Tanner (BT) crx-coefficient for past tOBs
cX=mat.or.vec(kmax,1)         #BT crx-coefficient for X
EqI=mat.or.vec(n,m)           #Numerical Indicator function for past tOBs = curr tOBs
q=mat.or.vec(kmax,m)          #q-values using Liang's procedure
c1Xdiff=mat.or.vec(n,m)       #BT crx-coefficient for Xdiff>0
psi=mat.or.vec(kmax,m)        #psi-values using Liang's procedure
dn=mat.or.vec(kmax,m)         #Empirical Bayes estimator (EBE) using Liang's procedure
naG=100                       #no. partitions in aGrid
aG=seq(from = 0,to=naG, length=(naG+2))/naG   #partitioned grid a=[0,1]
aG=aG[-1]                     #update partitioned grid to (0,1]
aG=aG[-101]                   #update partitioned grid to (0,1)
FBT=mat.or.vec(kmax,naG)      #BT cummulative distribution function
alpha=mat.or.vec(naG,m)       #alpha funtion used in monotonization procedure
Dstar=mat.or.vec(kmax,naG)    #D*(ai,x) estimator used in monotonization procedure
taiil=mat.or.vec(kmax,naG)    #ai x D*(ai,x) equivalent used in monotonization procedure
dns=mat.or.vec(kmax,m)        #monotonized EBE construction using Houwalingen's procedure
Ldg=mat.or.vec(kmax,1)        #integral values for Bayes estimator under loss function (L)
Ldmle=mat.or.vec(kmax,1)      #integral values for Max Likelihood estimator (MLE)  under L
Ldn=mat.or.vec(kmax,m)        #integral values for EBE under L
Ldns=mat.or.vec(kmax,m)       #integral values for monotonized EB estimator under L
rdn=mat.or.vec(1,m)           #Bayes risk for EB estimator
rdns=mat.or.vec(1,m)          #Bayes risk for monotonized EB estimator
umax=10                       #no. of runs
uavgRdn=mat.or.vec(1,umax)    #avg regret risks for EBE
uavgRdns=mat.or.vec(1,umax)   #avg regret risks for monotonized EBE
```

```r
#seeds<-c(2,12,16,17,23,27,59,65,72,75)     #seeds used for n=20
#seeds<-c(50,51,52,53,55,63,64,65,66,68)    #seeds used for n=40
#seeds<-c(50,51,52,53,55,58,61,63,64,65)    #seeds used for n=60,80
#seeds<-c(50,51,52,53,55,58,59,61,63,64,77)  #seeds used for n=100
#-----------------------------------------------------------------------------------------
# u=1                                                    #initiatate 1st run
# while (u <=10) {                                       #limit for the no. of repetitions
# set.seed(seeds[u])                                     #these seeds were used in my study
#%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% BEGIN RUN
#################################### DATA SIMULATION #####################################
X=matrix(r:(xmax-1), kmax, 1, FALSE)                    #define X values, current tOBs
for(k in 1:kmax){
  cX[k]= r*X[k]^(X[k]-r-1)/factorial(X[k]-r)            #BT crx-coefficient for past tOBs
  dg[k]=integrate(function(theta){theta^(X[k]+1-r)* #compute Bayes Estimator, theta_G
      exp(-X[k]*theta)},lower = a, upper = b)$val/integrate(function(theta)
        {theta^(X[k]-r)*exp(-X[k]*theta)},lower = a, upper = b)$val
  dmle[k] = (X[k]-r)/X[k]                               #compute MLE Estimator, theta_{MLE}
}
for(j in 1:m){
  for(i in 1:n){
    raTemp=runif(1,min=a, max=b)                        #draws random theta from G: Uni(a,b)
    ra[i,j]=raTemp                                      #matrix of theta values
    Xpast[i,j]=rbort(1, Qsize = r, a = raTemp)          #Xpast: observed past total OB sizes
    cXpast[i,j]= r*Xpast[i,j]^
      (Xpast[i,j]-r-1)/factorial(Xpast[i,j]-r)          #BT crx-coefficient for Xpast
  }
}
############################### EMPIRICAL BAYES ESTIMATOR ################################
k=1
while(k<=kmax){
  for(j in 1:m){
    for(i in 1:n){
      EqI[i,j]=as.numeric(I(Xpast[i,j]==X[k]))          #Indicator fn: used for q numerator
      if((Xpast[i,j]-X[k])>0) {                         #verifies psi condition is met
          c1Xdiff[i,j]= ((Xpast[i,j]-X[k])^             #psi numeratro: BT crx-coeff for diffX
                    ((Xpast[i,j]-X[k])-1-1))/factorial((Xpast[i,j]-X[k])-1)} #, r=1
      else {c1Xdiff[i,j]=0}                             #psi cond. not met->assigns zero...
    }                                                   #...to psi numerator
  }
  cXratio= c1Xdiff/cXpast                               #further used to define psi
  j=1
  while(j<=m){
    q[k,j]=sum(EqI[,j])/(n*cX[k])                       #compute q values
    psi[k,j]=sum(cXratio[,j])/n                         #compute psi values
    dn[k,j]=min(psi[k,j]/q[k,j],1)                      #define EBE theta_n
    j=j+1
  }
  k=k+1
}
#################################### MONOTONIZED EBE #####################################
for (j in 1:m) {
  for (i in 1:naG) {
    for (k in 1:kmax) {
```

```
      if (dn[k,j]<=aG[i]) {                                  #verifies EBE<=aGrid val
        alpha[i,j]=alpha[i,j]+sum(dbort(X[k],r,aG[i]))  #computes alpha as in Houwalingen
      }else{alpha[i,j]=alpha[i,j]}
    }
  }
}
for (i in 1:naG) {
  FBT[1,i]=dbort(X[1],r,aG[i])                               #BT cdf for case x=r
  for (k in 2:kmax) {
    FBT[k,i]=(FBT[k-1,i]+dbort(X[k],r,aG[i]))                #BT cdf for case x>r
  }
}
j=1
while (j<=m) {                                               #define D^*(a;x)
  for (i in 1:naG) {
    if (alpha[i,j]> FBT[1,i])                                #case x=r
      {Dstar[1,i]=1}
    else {Dstar[1,i]=alpha[i,j]/FBT[1,i]}
    for (k in 2:kmax) {                                      #case x>r
      if (FBT[k-1,i]>alpha[i,j])
        {Dstar[k,i]=0}
      else if (FBT[k,i]<alpha[i,j])
        {Dstar[k,i]=1}
      else {Dstar[k,i]=(alpha[i,j]-FBT[k-1,i])/
        (FBT[k,i]-FBT[k-1,i])}
    }
  }
  k=1
  while (k<=kmax) {
    for (i in 1:naG){
      taiil[k,i]=1-Dstar[k,i]
      dns[k,j]=sum(taiil[k,])/naG                            #define monotonized EBE, theta_n^*
    }
    k=k+1
  }
  j=j+1
}
#################### RISKS UNDER SQUARED ERROR LOSS FUNCTION: L ##########################
for (k in 1:kmax) {
  Ldg[k]=integrate(function(theta) {(dg[k]-theta)^2*          #L values: Bayes estimator
      theta^(X[k]-r)*exp(-X[k]*theta)}, lower = a, upper = b)$val*cX[k]/(b-a)
  Ldmle[k]=integrate(function(theta) {(dmle[k]-theta)^2*      #L values: MLE
      theta^(X[k]-r)*exp(-X[k]*theta)},lower = a, upper = b)$val*cX[k]/(b-a)
  for (j in 1:m) {
    Ldn[k,j]= integrate(function(theta){(dn[k,j]-theta)^2*    #L values: EBE
        theta^(X[k]-r)*exp(-X[k]*theta)},lower=a, upper=b)$val*cX[k]/(b-a)
    Ldns[k,j]= integrate(function(theta){(dns[k,j]-theta)^2*  #L values: monotonized EBE
        theta^(X[k]-r)*exp(-X[k]*theta)},lower=a, upper=b)$val*cX[k]/(b-a)
    rdn[j]=sum(Ldn[,j])                                       #Bayes risk for EBE
    rdns[j]=sum(Ldns[,j])                                     #Bayes risk for monotone EBE
  }
}
rdg=sum(Ldg)                             #min Bayes risk
```

```
rdmle=sum(Ldmle)                          #Bayes risk for MLE
Rdmle=rdmle-rdg                           #regret risk for MLE
Rdn=rdn-rdg                               #regret risk for EBE
Rdns=rdns-rdg                             #regret risk for monotonized EBE
avgRdn=1/m*sum(Rdn)                       #avg EBE regret risk for m sets
avgRdns=1/m*sum(Rdns)                     #avg mEBE regret risk for m sets
Vdn=sum((rdn-avgRdn)^2)/(m-1)             #variance for EBE
Vdns=sum((rdns-avgRdns)^2)/(m-1)          #variance for monotonized EBE
SDdn=sqrt(Vdn)                            #standard deviation for EBE
SDdns=sqrt(Vdns)                          #standard deviation for monotonized EBE
SEdn=Vdn/sqrt(n)                          #standard error for EBE
SEdns=Vdns/sqrt(n)                        #standard error for monotonized EBE
# Rresults=rbind(Rdn,Rdns)                  #combines Rdn, rdns results; 2 by 10 mat.
# row.names(Rresults)<-c("Rdn","Rdns")      #adds corresponding names to the rows
# print(Rresults)                           #prints regret risk matrix for run u
#%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%END RUN
# uavgRdn[1,u]=avgRdn                       #avg EBE regret risk for run u
# uavgRdns[1,u]=avgRdns                     #avg mEBE regret risk for run u
# u=u+1                                     #update run u and repeat
# }
# print(paste("Rdmle = ",Rdmle))            #outputs label and value of MLE regret risk
# print(uavgRdn)                            #outputs corresponding regret risk avg for...
# print(uavgRdns)                           #...each run as a 1 by u vector
```

APPENDIX B

APPENDIX B

Table 2.1: References on notation

Notation	Description
Θ	unknown rv parametrizing X; the reproduction parameter
θ	a realization of the reproduction parameter Θ
$\hat{\theta}$	refers to any estimator
θ_{ML}	ML estimator for BT distribution
θ_n	Empirical Bayes estimator for BT based on Liang's [9] procedure
θ_n^*	Monotonized EB estimator for BT based on Houwalingen's [14] procedure
$\underline{X}$	the set of n past observations $X_1, X_2, \ldots, X_n$
BT	Borel–Tanner
cdf	cummulative distribution function
EB	Empirical Bayes
GW	Galton–Watson also known as Bienaymé–Galton–Watson
iid	independent identically distributed
ML	maximum likelihood
MLR	monotone likelihood ratio
pmf	probability mass function
$Poi(\theta)$	Poisson distribution with parameter (θ)
rv	random variable
$r(G, \hat{\theta})$	Bayes risk for estimator $\hat{\theta}$ under $G-$prior
$R(\hat{\theta})$	Regret risk for estimator $\hat{\theta}$
$\overline{R}(\hat{\theta})$	Average regret risk for estimator $\hat{\theta}$
$Uni(a, b)$	Uniform distribution with parameter values (a, b)

BIOGRAPHICAL SKETCH

Celestina Ruby Soltero was born in Houston, Texas and lived most of her childhood in Roma, Texas. She received her Associate of Science in Mathematics from South Texas College in August 2012. In August 2015 she was awarded a Bachelor of Science in Mathematics from University of Texas–PanAmerican and, in August 2017, completed a Master of Science in Mathematics at University of Texas–Rio Grande Valley.

Celestina may be contacted via email at crubysa@hotmail.com.